53
Г27

Р. Г. ГЕВОРКЯН

КУРС ФИЗИКИ

Р. Г. ГЕВОРКЯН

КУРС ФИЗИКИ

Допущено Министерством высшего
и среднего специального образования СССР
в качестве учебного пособия для студентов
высших учебных заведений



МОСКВА «ВЫСШАЯ ШКОЛА» 1979

ББК 22.3
Г 27
УДК 53

Рецензент — зав. кафедрой физики МИИГА, проф. А. Д. Суханов

Геворкян Р. Г.

Г 27 Курс физики: Учеб. пособие. — М.: Высш. школа, 1979. — 656 с., ил.

В пер.: 1 р. 40 к.

В пособии приводятся основные теоретические сведения, определенные действующей программой для высших учебных заведений. Рассмотрена физическая сущность явлений, описаны методы их изучения, формулируются физические понятия и законы. В книге нашли отражение новейшие достижения физики, получившие практическое применение.

Предназначается для студентов вузов, в основном для вечерних и заочных отделений.

Г $\frac{20401-211}{001(01)-79}$ 39—79 1704010000

53
ББК 22.3

© Издательство «Высшая школа», 1979

ПРЕДИСЛОВИЕ

Настоящее учебное пособие предназначено для студентов высших учебных заведений, прежде всего для их вечерних и заочных отделений. Автор поставил перед собой задачу — составить краткое, местами конспективное изложение сведений, предусмотренных действующей программой курса физики. В книге приведено в основном только «установившееся» содержание курса, которое практически мало изменяется со временем. Для сокращения объема книги было решено исключить исторические обзоры, детальные описания приборов, экспериментальных установок и методов измерений, с которыми студенты должны ознакомиться на лабораторных занятиях, а также описание демонстраций, обычно показываемых на лекциях.

Порядок размещения учебного материала в книге традиционный. Сделано лишь одно исключение: некоторые положения, обоснование которых приводится в курсе теоретической физики, в тексте книги даны в виде «готовых» утверждений, но их детальное обсуждение, которое может заинтересовать любознательных студентов, выделено отдельно в виде Заключения к курсу физики. В Заключении приведен также анализ некоторых важных предположений, используемых при объяснении физических явлений; этот анализ может рассматриваться как «материал для размышлений», полезный для творческого, не формального отношения к физике — одной из важнейших наук об окружающей нас природе.

При написании книги была использована некоторая часть ранее вышедшего «Курса общей физики», написанного автором совместно с В. В. Шепелем. Разделы, принадлежащие автору, подвергнуты некоторой переработке; разделы, написанные В. В. Шепелем, заменены новым текстом.

Автор приносит глубокую благодарность рецензенту проф. А. Д. Суханову и редактору С. А. Крылову за ценные замечания, которые были использованы при подготовке рукописи к изданию.

Автор

Часть I
МЕХАНИКА

Глава 1
ОСНОВНЫЕ ЗАКОНЫ МЕХАНИКИ

§ 1. СИСТЕМЫ ОТСЧЕТА; ЛИНЕЙНЫЕ И УГЛОВЫЕ СКОРОСТИ И УСКОРЕНИЯ

Для описания движения тел необходимо предварительно выбрать система отсчета, т. е. выбрать одно или несколько тел, которые условно полагаются неподвижными (*тела отсчета*), и связать с ними какую-нибудь *координатную систему*, например прямоугольную.

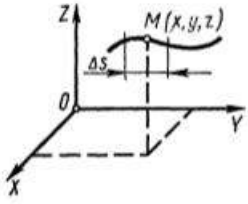


Рис. I.1

В качестве тел отсчета можно выбрать каркас измерительной установки, стены лаборатории, поверхность Земли, неподвижные звезды.

В частности, для описания движений, происходящих на поверхности Земли, оси OX и OY прямоугольной координатной системы располагают в горизонтальной плоскости, а ось OZ направляют по вертикали. Для учета кривизны поверхности Земли начало координатных осей помещают в центре Земли, а одну из этих осей совмещают с осью вращения Земли. Для описания движения космических объектов координатную систему связывают с определенными звездами, выделяющимися своей яркостью.

При движении тела каждая его точка описывает некоторую линию — траекторию. В прямоугольной системе координат (рис. I.1) траектория точки определяется тремя функциями:

$$x = x(t); \quad y = y(t); \quad z = z(t), \tag{1.1}$$

показывающими, как изменяются координаты x, y и z этой точки с течением времени. Различные точки тела могут описывать различные по форме и размерам траектории. Наиболее простым образом будет описываться движение тел, которые имеют очень малые размеры и поэтому могут рассматриваться как точечные тела (или как *материальные точки*). Реальные тела могут быть представлены как системы из взаимосвязанных точечных тел; описание движения таких тел сводится к определению траекторий их отдельных точечных частей. Движение тела называется *поступательным*, если траектории всех точек тела одинаковы и поэтому любая прямая, проведенная в теле, смещается параллельно самой себе, и *вращательным*, если точки тела

описывают концентрические окружности. Сложное движение тела можно представить как сочетание поступательных и вращательных движений.

СКОРОСТЬ И УСКОРЕНИЕ

Допустим, что траектория движения точечного тела известна. Выберем на этой траектории некоторый отрезок пути Δs, внутри которого находится точка M. Этот отрезок будет пройден телом за время Δt. Нас интересует скорость, с которой тело проходит через точку M. Отношение

$$\Delta s/\Delta t = v_{ср}$$

будет *средней скоростью* тела на участке Δs. При неравномерном движении эта скорость будет получаться различной в зависимости от выбираемых размеров Δs и Δt. Однако если участок траектории и, следовательно, промежуток времени Δt будут достаточно малыми, то скорость тела не успеет измениться на заметную величину и движение тела на участке Δs можно считать почти равномерным. Очевидно, что измеряемая скорость движения тела не будет зависеть от произвольного выбора размеров Δs и Δt, если эти размеры будут достаточно малыми. Ввиду этого предел отношения

$$\lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t} = \frac{ds}{dt} = v \qquad (1.2)$$

называют *скоростью тела* в данный момент времени (взятый в пределах промежутка dt) или скоростью тела в данной точке траектории (расположенной в пределах участка ds). Эти скорости называют также *мгновенными* или *истинными*.

Для криволинейной траектории любой формы бесконечно малый участок траектории можно полагать прямолинейным и представлять в виде вектора $\mathbf{ds}$, ориентированного по направлению движения. Тогда скорость тела

$$\mathbf{v} = \mathbf{ds}/dt \qquad (1.3)$$

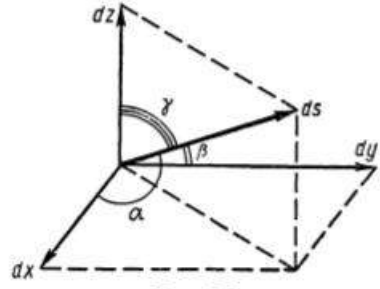


Рис. 1.2

также будет вектором, ориентированным в том же направлении. Вектор $\mathbf{ds}$ можно представить как сумму трех векторов (рис. 1.2):

$$\mathbf{ds} = \mathbf{dx} + \mathbf{dy} + \mathbf{dz},$$

поэтому скорость тела $\mathbf{v}$ может быть разложена на составляющие по координатным осям:

$$\mathbf{v} = \frac{\mathbf{ds}}{dt} = \frac{\mathbf{dx}}{dt} + \frac{\mathbf{dy}}{dt} + \frac{\mathbf{dz}}{dt} = \mathbf{v}_x + \mathbf{v}_y + \mathbf{v}_z. \qquad (1.4)$$

Если известны функции (1.1), определяющие траекторию движения тела, то проекции скорости на координатные оси будут равны:

$$v_x = \frac{dx(t)}{dt}; \quad v_y = \frac{dy(t)}{dt}; \quad v_z = \frac{dz(t)}{dt}, \tag{1.5}$$

а величина скорости

$$v = \sqrt{v_x^2 + v_y^2 + v_z^2}.$$

По этим данным можно вычислить также углы α, β и γ между вектором скорости $\mathbf{v}$ и координатными осями:

$$\cos\alpha = \frac{dx}{ds} = \frac{v_x}{v}; \quad \cos\beta = \frac{dy}{ds} = \frac{v_y}{v}; \quad \cos\gamma = \frac{dz}{ds} = \frac{v_z}{v}. \tag{1.6}$$

Допустим, что v_1 и v_2 — скорости тела (по величине) в начальной и конечной точках некоторого отрезка траектории Δs, внутри которого находится точка M, а Δt — время прохождения этого отрезка. Нас интересует ускорение, с которым тело проходит через точку M. Отношение

$$\frac{v_2 - v_1}{\Delta t} = \frac{\Delta v}{\Delta t} = a_{ср}$$

будет *средним ускорением* на участке Δs, и в общем случае оно может получиться различным в зависимости от выбранных размеров Δs и Δt. Однако если эти размеры выбрать достаточно малыми (в пределе — бесконечно малыми), то результат измерения уже не будет зависеть от размеров Δs или Δt, так как за малое время или на малом участке траектории ускорение не может измениться на заметную величину. Ввиду этого предел отношения

$$\lim_{\Delta t \to 0} \frac{\Delta v}{\Delta t} = a \tag{1.7}$$

называют *ускорением тела в данной точке траектории* или в данный момент времени. Если при движении тела его скорость уменьшается, то изменение скорости $\Delta v = v_2 - v_1$, а следовательно, и ускорение a будут отрицательными.

Так как изменение скорости $\Delta\mathbf{v} = \mathbf{v}_2 - \mathbf{v}_1$ — вектор, то ускорение будет вектором, ориентированным по направлению $\Delta\mathbf{v}$ (очевидно, что с р е д н е е ускорение не имеет смысла представлять в виде вектора, так как не только величина, но и направление $\Delta\mathbf{v}$ может заметно измениться при выборе различных размеров Δs и Δt). Следовательно, вектор ускорения в данной точке траектории (или в данный момент времени) должен определяться по вектору бесконечно малого изменения скорости:

$$\mathbf{a} = \frac{d\mathbf{v}}{dt}. \tag{1.8}$$

Составляющие (проекции) этого вектора по координатным осям будут равны:

$$a_x = \frac{dv_x}{dt} = \frac{d^2x}{dt^2}; \quad a_y = \frac{dv_y}{dt} = \frac{d^2y}{dt^2}; \quad a_z = \frac{dv_z}{dt} = \frac{d^2z}{dt^2}, \tag{1.9}$$

а полное ускорение

$$a = \sqrt{a_x^2 + a_y^2 + a_z^2}.$$

По этим данным можно вычислить углы α, β и γ, образованные вектором ускорения и координатными осями:

$$\cos\alpha = a_x/a; \quad \cos\beta = a_y/a; \quad \cos\gamma = a_z/a.$$

В общем случае эти углы не будут совпадать с указанными выше углами между вектором скорости и координатными осями. Это означает, что направление вектора ускорения **a**, вообще говоря, не совпадает с направлением вектора скорости **v**; такое совпадение имеет место только при прямолинейном движении.

При постоянном ускорении расчетные формулы для скорости и пройденного пути имеют вид

$$v = v_0 + at; \quad s = v_0 t + \frac{at^2}{2}, \tag{1.10}$$

где v_0 — скорость в начальный момент времени ($t = 0$). Если ускорение движения вдоль траектории не остается постоянным, то можно разбить время движения t на такие малые промежутки времени Δt_1, Δt_2, ..., чтобы в течение каждого из них изменением ускорения можно было пренебречь. Обозначим ускорение тела для промежутка времени Δt_i через a_i; тогда изменение скорости в течение этого времени будет равно $a_i \Delta t$. Суммируя эти изменения за истекшее время t, можно написать:

$$v = v_0 + a_1 \Delta t_1 + a_2 \Delta t_2 + \ldots = v_0 + \sum a_i \Delta t_i.$$

Чем меньше промежутки времени Δt_i, тем точнее этот расчет. В пределе, когда эти промежутки выбираются бесконечно малыми, сумма заменяется интегралом:

$$v = v_0 + \int_0^t a(t)\,dt. \tag{1.11}$$

Для вычисления этого интеграла необходимо знать функцию $a(t)$, показывающую, как изменяется ускорение с течением времени. Если известно, как изменяется со временем скорость движения тела, т. е. задан вид функции $v = v(t)$, то путь, пройденный телом за время t, можно рассчитать приближенно по формуле

$$s = v_1 \Delta t_1 + v_2 \Delta t_2 + \ldots = \sum v_i \Delta t_i$$

или точно по формуле

$$s = \int_0^t v(t)\,dt. \tag{1.12}$$

При определении скорости по формуле $v = s/t$ путь s, пройденный телом, обычно берется по абсолютной величине, поэтому скорость движения получается всегда положительной величиной. Однако во многих случаях путь s определяется как разность координат движу-

щегося тела в конечный и начальный моменты. Например, при движении тела вдоль оси OX

$$s = x_2 - x_1; \quad v = \frac{x_2 - x_1}{t}.$$

Тогда знак скорости движения будет зависеть от того, в каком направлении происходит перемещение тела. Если тело перемещается в направлении возрастания координаты x, то $x_2 > x_1$ и скорость тела положительная. Если же тело движется в направлении убывания координаты x, то $x_2 < x_1$ и скорость перемещения будет отрицательной.

В связи с этим знак ускорения $a = \frac{v_2 - v_1}{t}$ определяется не только возрастанием или убыванием скорости по величине, но и знаком самой скорости. Например, если тело движется ускоренно в направлении убывания координаты x, то скорости v_1 и v_2 будут отрицательными, а так как по величине $v_2 > v_1$, то и ускорение a будет отрицательной величиной: $a = \frac{-v_2 - (-v_1)}{t} = -\frac{v_2 - v_1}{t}$. Таким образом, если тело движется в направлении убывания координаты x, то возрастание скорости ($v = \Delta x/\Delta t$) происходит при отрицательном ускорении ($a = -\Delta v/\Delta t$), а уменьшение скорости — при положительном ускорении.

ДВИЖЕНИЕ ПО КРИВОЛИНЕЙНОЙ ТРАЕКТОРИИ

Рассмотрим движение точечного тела по некоторой криволинейной траектории (рис. I.3). Положение тела в моменты t_1 и t_2, обозначенные A и B, можно фиксировать радиус-векторами $\mathbf{r}_A$ и $\mathbf{r}_B$, проведенными из предварительно выбранной нами точки O. Вектор $\Delta\mathbf{r}$, соединяющий точки A и B, не определяет форму траектории на участке AB. Например, для сплошной и пунктирной траекторий, показанных на рис. I.3, форма и размеры участка Δs различны, тогда как вектор $\Delta\mathbf{r}$ одинаков по величине и направлению, поэтому отношение $\Delta r/\Delta t$ не равно средней скорости движения тела на участке AB. Однако если точка B очень близка к точке A, то Δr будет мало отличаться от дуги $\Delta s = AB$ и в пределе (когда $\Delta s \to 0$, угол $\Delta\alpha \to 0$) будет совпадать с ней. Ввиду этого предел отношения

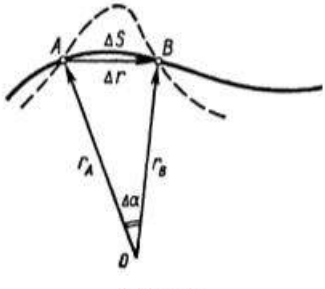


Рис. I.3

$$\lim_{\Delta t \to 0} \frac{\Delta\mathbf{r}}{\Delta t} = \frac{d\mathbf{r}}{dt} \tag{1.13}$$

будет вектором скорости тела на бесконечно малом участке траектории между точками A и B.

Известно, что очень малые участки любой кривой можно заменить дугой окружности; для этого необходимо подобрать подходящий ра-

диус R этой окружности и расположение ее центра O (R называется *радиусом кривизны кривой в данном месте*, а O — *центром кривизны*). На рис. I.4, *а* показаны радиус и центр окружности, дуга которой Δs совпадает с участком траектории AB. На рис. I.4, *б* показаны скорости $\mathbf{v}_1$ и $\mathbf{v}_2$ в точках A и B и изменения скорости $\Delta\mathbf{v} = \mathbf{v}_2 - \mathbf{v}_1$ на

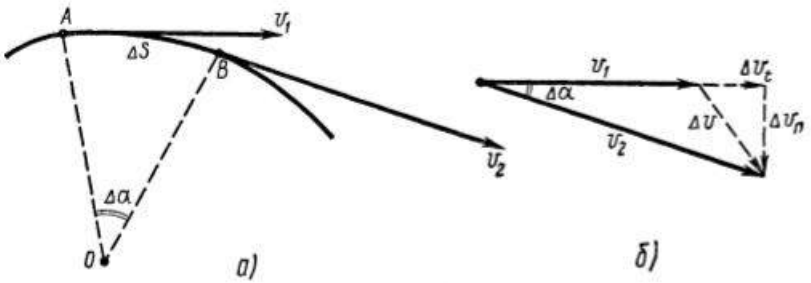


Рис. I.4

этом участке. Вектор $\Delta\mathbf{v}$ можно разложить на две составляющие: $\Delta\mathbf{v}_t$ — по направлению скорости $\mathbf{v}_1$ и $\Delta\mathbf{v}_n$ — перпендикулярно ей:

$$\Delta\mathbf{v} = \Delta\mathbf{v}_t + \Delta\mathbf{v}_n.$$

Разделим это равенство на Δt и перейдем к пределу при $\Delta t \to 0$ (при этом $\Delta\alpha \to 0$, т. е. направления векторов $\mathbf{v}_1$ и $\mathbf{v}_2$ будут почти совпадать):

$$\lim_{\Delta t \to 0} \frac{\Delta\mathbf{v}}{\Delta t} = \lim_{\Delta t \to 0} \frac{\Delta\mathbf{v}_t}{\Delta t} + \lim_{\Delta t \to 0} \frac{\Delta\mathbf{v}_n}{\Delta t}$$

или

$$\mathbf{a} = \mathbf{a}_t + \mathbf{a}_n. \tag{1.14}$$

Вектор $\mathbf{a}_t$ (*тангенциальное ускорение*) есть проекция вектора полного ускорения $\mathbf{a}$ на направление движения. Это ускорение изменяет скорость движения только по величине. Оно направлено по скорости или против нее в зависимости от того, является ли движение ускоренным или замедленным. Численное значение тангенциального ускорения определяется по изменению *величины* скорости:

$$a_t = \frac{dv}{dt}. \tag{1.15}$$

Если движение (по любой криволинейной траектории) является равномерным, то $a_t = 0$.

Вектор $\mathbf{a}_n$ есть проекция вектора полного ускорения $\mathbf{a}$ на направление, перпендикулярное скорости, т. е. на направление радиуса вектора R или нормали к траектории. Это ускорение называется *нормальным* или *центростремительным*. Оно вызывает изменение скорости только по направлению и поэтому отсутствует при прямолинейном движении. Так как $\Delta s = R\Delta\alpha$ (α в радианах), а $\Delta v_n \approx v\Delta\alpha = v(\Delta s/R)$

(индекс у скорости можно отбросить, так как $\mathbf{v}_1 \approx \mathbf{v}_2$), то

$$a_n = \lim_{\Delta t \to 0} \frac{\Delta v_n}{\Delta t} = \lim_{\Delta t \to 0} \frac{v}{R} \frac{\Delta s}{\Delta t} = \frac{v^2}{R}, \tag{1.16}$$

следовательно, величина нормального ускорения в каждой точке криволинейной траектории определяется скоростью движения и радиусом кривизны траектории в этой точке.

Величина полного ускорения в данной точке траектории может быть вычислена по значениям тангенциального и нормального ускорений:

$$a = \sqrt{a_t^2 + a_n^2} = \sqrt{(\mathrm{d}v/\mathrm{d}t)^2 + (v^2/R)^2}. \tag{1.17}$$

Можно также определить угол α между направлениями векторов скорости $\mathbf{v}$ и ускорения $\mathbf{a}$:

$$\cos\alpha = \frac{a_t}{a} = \frac{1}{a}\frac{\mathrm{d}v}{\mathrm{d}t}; \quad \sin\alpha = \frac{a_n}{a} = \frac{1}{a}\frac{v^2}{R}. \tag{1.18}$$

Рассмотрим частный случай, когда точечное тело движется по окружности. Выберем настолько малый промежуток времени Δt,

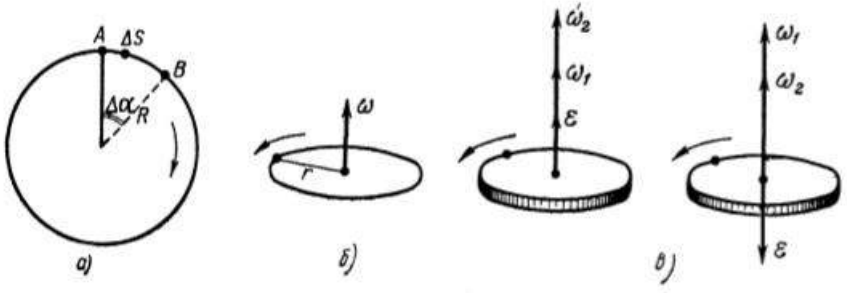


Рис. I.5

чтобы движение тела на участке окружности Δs можно было полагать почти равномерным. Отношение угла поворота $\Delta\alpha$ радиус-вектора R (рис. I.5, *а*) к прошедшему времени Δt есть *средняя угловая скорость* вращения

$$\omega_{\mathrm{ср}} = \frac{\Delta\alpha}{\Delta t}.$$

Единица угловой скорости: $[\omega] = 1$ рад/с.

В пределе при $\Delta t \to 0$ и $\Delta\alpha \to 0$ мы получаем *угловую скорость в данный момент времени*:

$$\omega = \lim_{\Delta t \to 0} \frac{\Delta\alpha}{\Delta t} = \frac{\mathrm{d}\alpha}{\mathrm{d}t}. \tag{1.19}$$

Аналогично, если ω_1 и ω_2 — угловые скорости вращения в моменты времени t_1 и t_2, то отношение

$$\varepsilon = \frac{\omega_2 - \omega_1}{t_2 - t_1} = \frac{\Delta\omega}{\Delta t}$$

будет *средним угловым ускорением* за время Δt, а предел отношения

$$\varepsilon = \lim_{\Delta t \to 0} \frac{\Delta \omega}{\Delta t} = \frac{d\omega}{dt} \tag{1.20}$$

будет *угловым ускорением в данный момент времени*. Так как $\alpha = s/R$, то между линейными и угловыми величинами, характеризующими движения точечного тела по окружности, существуют соотношения:

$$s = R\alpha; \quad v = \frac{ds}{dt} = R\frac{d\alpha}{dt} = R\omega; \quad a_t = \frac{dv}{dt} = R\frac{d\omega}{dt} = R\varepsilon; \tag{1.21}$$

$$a_n = \frac{v^2}{R} = R\omega^2.$$

При постоянном угловом ускорении расчетные формулы для угловой скорости вращения и угла поворота имеют простой вид:

$$\omega = \omega_0 + \varepsilon t; \quad \alpha = \omega_0 t + \frac{\varepsilon t^2}{2}, \tag{1.22}$$

где ω_0 — угловая скорость в начальный момент времени $t = 0$. Если за время t тело совершает несколько оборотов вокруг оси вращения, то их число можно рассчитать, разделив α на 2π.

Угловая скорость вращения изображается вектором ω, направленным по оси вращения; направление вектора угловой скорости условились определять по правилу «правого винта», т. е. в направлении поступательного движения винта, имеющего данное вращение (рис. I.5, *б*). Угловое ускорение есть также векторная величина:

$$\varepsilon = \frac{\omega_2 - \omega_1}{t}. \tag{1.23}$$

Если угловая скорость по величине возрастает, то разность векторов $\omega_2 - \omega_1$, а следовательно, и угловое ускорение будут иметь направление, совпадающее с направлением ω_1 и ω_2. Если же угловая скорость вращения убывает, то разность $\omega_2 - \omega_1$, а следовательно, и вектор углового ускорения ориентируются в направлении, противоположном направлению угловой скорости (рис. I.5, *в*).

§ 2. ЗАКОНЫ МЕХАНИКИ НЬЮТОНА. СИСТЕМЫ ЕДИНИЦ

СИЛА. ДЕФОРМАЦИЯ. МАССА

В механике для описания воздействия одних тел на другие пользуются понятием силы: *силой* называется всякое воздействие на данное тело, в результате которого тело получает ускорение или изменяет свою форму и размеры.

Рассмотрим сначала *деформацию* твердых тел под воздействием различных сил. Деформация называется *упругой*, если при прекращении действия сил тело принимает в точности первоначальную форму и размеры. Реальные тела обнаруживают в той или иной степени *остаточную деформацию*, т. е. после прекращения действия сил они не возвращаются в точности к первоначальным размерам. Если дефор-

мирующие силы слабы, то вызываемые ими небольшие деформации тел оказываются почти упругими.

Рассмотрим простейший вид деформации — удлинение стержня или проволоки под действием некоторой растягивающей силы F. Пусть длина стержня l под действием этой силы увеличивается на Δl. Опыт показывает, что при упругой деформации величина Δl пропорциональна растягивающей силе F, длине стержня l, обратно пропорциональна площади поперечного сечения стержня S и зависит от упругих свойств вещества, из которого сделан стержень. Эта зависимость записывается в виде

$$\Delta l = \frac{Fl}{ES}, \tag{1.24}$$

где величина E, называемая *модулем Юнга* (модулем продольной упругости), зависит от вещества стержня. Обозначим относительное удлинение $\Delta l/l$ через ε. Отношение $F/S = \sigma$ показывает, какая растягивающая сила приходится на единицу площади поперечного сечения стержня, и называется *нормальным механическим напряжением*. Тогда

$$\sigma = \varepsilon E; \quad \varepsilon = \sigma/E, \tag{1.25}$$

т. е. относительное удлинение стержня или проволоки пропорционально нормальному напряжению. Прямая пропорциональность между относительной деформацией и соответствующим напряжением наблюдается и для других видов упругой деформации и называется з а к о н о м Г у к а.

Силы, сообщающие телам ускорения или деформирующие их, можно разделить на две группы:

а) силы, *распределенные* по объему тела и пропорциональные массам отдельных частей этого тела. Такой «распределенной» силой является, например, сила тяготения. Заметим, что распределенные силы не зависят от того, каким образом связаны между собой отдельные части данного тела; если тело, притягивающееся к Земле, разрезать на части, то силы тяготения, действующие на эти части, остаются без изменения;

б) силы *сосредоточенные*, т. е. приложенные к определенному месту тела — в какой-нибудь точке, вдоль некоторой линии или поверхности. В этом случае действие внешней силы передается от одних частей тела к другим благодаря существованию взаимной связи между ними.

Рассмотрим действие распределенных и сосредоточенных сил. Если на тело действует распределенная сила, сообщающаяся каждому элементу этого тела одно и то же ускорение (по величине и направлению), то все тело будет двигаться с этим ускорением и при этом не будет деформировано. Например, тело малых размеров, свободно падающее в поле тяготения Земли, не деформировано, так как все его составные части движутся с одинаковыми ускорениями. Распределенные силы могут вызвать деформацию тела в том случае, если ускорения, которые сообщаются отдельным элементам тела, несколько отличаются по численному значению или направлению.

Если на тело действует сосредоточенная сила, то тело всегда деформируется. Допустим, что к сложной пружине, состоящей из четырех одинаковых элементов, в точке A приложена некоторая сосредоточенная сила, вследствие чего вся пружина движется с ускорением в направлении действия силы (рис. I.6, *а*). Пружина деформируется и притом тем сильнее, чем больше ускорение; кроме того, элементы пружины, расположенные ближе к точке приложения силы, растянуты больше, чем элементы, находящиеся дальше от этой точки. Действительно, деформация элемента *1* должна быть такой, чтобы сила, приложенная со стороны этого элемента к точке B, сообщала ускорение элементам *2, 3* и *4*. Деформация элемента *2* должна быть меньше, так как этот элемент сообщает ускорение только третьей и четвертой пружинам и т. д.

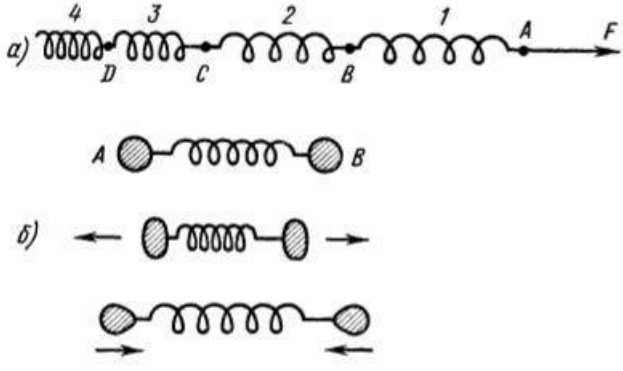


Рис. I.6

Из изложенного видно, что величину силы можно определить как по ускорению, которое эта сила может сообщить, так и по деформации, которую эта сила вызывает. Рассмотрим систему из двух тел, связанных упругой пружиной и совершающих колебательные движения (рис. I.6, *б*). При малых ускорениях деформации колеблющихся тел A и B незначительны и трудно измеряемы, поэтому силы, действующие на эти тела со стороны пружины, легче определить по величине ускорений, которые они приобретают, чем по степени деформаций этих тел. Наоборот, силы, действующие на пружину со стороны тел A и B, значительно удобнее определять по величине деформации самой пружины — изменению ее длины, а не по величине ускорений, которые сообщаются различным участкам пружины.

Рассмотрим еще один пример: допустим, к потолку кабины лифта (рис. I.7) подвешено на пружине некоторое тело. Если кабина неподвижна, то пружина деформирована настолько, чтобы уравновесить притяжение тела к Земле, т. е. сообщить телу ускорение, равное g и направленное вверх. Если кабина получает ускорение a вверх, то деформация пружины увеличивается, так как пружина должна сообщить телу ускорение, равное $g + a$. При этом сила, деформирующая пружину, равна $F = m(g + a) = P + ma$. При опускании лифта с

ускорением a деформация пружины уменьшается; в этом случае пружина должна сообщить телу ускорение, равное $g - a$. При помощи измерения можно установить связь между деформацией пружины и величиной ускорения $g \pm a$, сообщаемого этой пружиной телу.

Для измерения сил можно выбрать либо эталонные пружины и измерять силы по величине деформации этих пружин, либо эталонное тело и измерять силы по величине ускорений, которые сообщаются этому телу. Первый способ имеет преимущество перед вторым: измерение деформации эталонных пружин (их удлинения или изгиба) произвести значительно проще, чем определить ускорение эталонного тела. Однако существенным недостатком первого способа является изменение упругих свойств пружины под действием различных факторов, среди которых могут быть и не контролируемые. Кроме того, реальные пружины обнаруживают «остаточную деформацию».

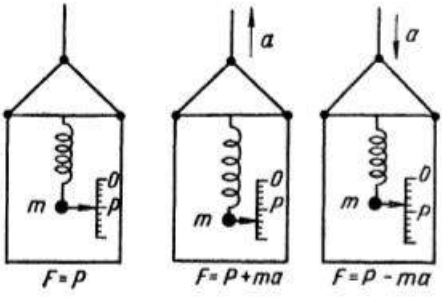


Рис. 1.7

Второй способ, основанный на определении ускорений, которые сообщаются силами эталонному телу, в этом отношении безупречен, но практически весьма громоздок и не всегда применим.

Если к двум различным местам тела приложены две сосредоточенные силы (или одна сила — сосредоточенная, а другая — распределенная), то возможно такое соотношение этих сил, когда суммарное ускорение тела равно нулю. Говорят, что в этом случае силы «уравновешивают» друг друга; ускорения, которые сообщают эти силы данному телу, равны по величине и противоположны по направлению. Однако, если векторная сумма ускорений, сообщаемых «уравновешенными» силами, равна нулю, то сумма деформаций, вызываемых этими силами, не равна нулю, т. е. уравновешенные силы, приложенные в различных местах тела, вызывают деформацию тела. Например, если к телу (пружине) приложены две сосредоточенные силы F_1 и F_2 (рис. I.8, *а*), то тело оказывается деформированным, даже если суммарное ускорение его движения равно нулю.

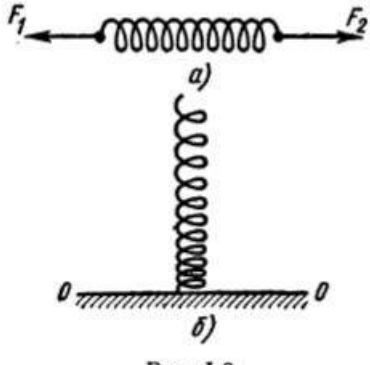


Рис. 1.8

Следует иметь в виду, что деформировано и всякое тело, покоящееся в поле тяготения Земли — подвешенное или опирающееся на поверхность Земли. Это легко заметить у длинной пружины, поставленной на

горизонтальную опорную поверхность (рис. I.8, *б*). Здесь распределенная сила тяжести уравновешивается сосредоточенной силой, приложенной к пружине со стороны опорной поверхности. Суммарное ускорение, сообщаемое всей пружине в целом или отдельным ее элементам, равно нулю, но тело деформировано. Подобно примеру, приведенному на рис. I.6, деформация участков пружины, расположенных близко к месту приложения сосредоточенной силы (т. е. к поверхности $O - O$), больше, чем деформация участков, расположенных выше.

Выбрав какой-нибудь способ измерения силы, можно установить, что одна и та же сила сообщает различным телам различное ускорение. Условились называть *массой* свойство тел, которое определяет величину ускорения, приобретаемого телами под действием данных сил. Таким образом, измерение массы связано с измерением сил.

Масса тела характеризует меру его инертности, т. е. свойство сохранять приобретенное движение или состояние покоя.

Допустим, что эталонная пружина, растянутая (или сжатая) на l сантиметров, сообщает измеряемому телу ускорение a, а эталонному телу — ускорение a_0. Для каждого тела отношение a/a_0 оказывается постоянным и не зависящим от l. Это отношение приравняется обратному отношению масс тела (m) и эталона (m_0) и тогда

$$\frac{a}{a_0} = \frac{m_0}{m}; \quad m = m_0 \frac{a_0}{a}.$$

Масса тела, определяющая величину ускорения, приобретаемого телом под действием сил, иногда называется *инертной массой* (о «весомой» массе см. § 5; о массе, не связанной с ускорением, а определяемой только энергией см. § 1 ч. IV).

ОСНОВНЫЕ ЗАКОНЫ МЕХАНИКИ

Приведем краткие формулировки законов механики Ньютона и некоторые следствия из них.

Тело, изолированное от всякого внешнего воздействия, либо покоится, либо движется прямолинейно и равномерно (з а к о н и н е р ц и и).

Это утверждение можно рассматривать как обобщение опытных данных, если предварительно выбраны способы измерения длины и времени (для установления равномерности движения) и если имеются какие-либо способы установить наличие или отсутствие внешних воздействий на рассматриваемое тело. Тело может покоиться или двигаться с постоянной скоростью и в том случае, если действующие на него силы взаимно уравновешены; но тогда тело будет деформировано. Следовательно, если на тело не действуют никакие другие тела, то оно может покоиться или двигаться прямолинейно и равномерно и при этом будет иметь определенную форму и размеры. Воздействие других тел на рассматриваемое тело заключается, таким образом, в изменении скорости движения и в изменении формы и размеров тела.

Ускорение, приобретаемое телом под действием данной силы, прямо пропорционально величине этой силы и обратно пропорционально массе тела.

Направление ускорения совпадает с направлением действующей силы:

$$\mathbf{a} = \mathbf{F}/m.$$

Здесь предполагается, что сила измеряется при помощи специально проградуированных пружин (динамометров), а масса — по отношению ускорений a и a_0, сообщаемых определенной силой данному телу и эталону: $m = m_0(a_0/a)$.

Таким образом, второй закон механики выражает полученное на опыте соотношение между силой, сообщающей ускорение, величиной этого ускорения и массой тела, к которому приложена эта сила:

$$\mathbf{F} = m\mathbf{a}. \tag{1.26}$$

Если на тело действует не одна сила, а несколько сил различных величин и направлений ($\mathbf{F}_1$, $\mathbf{F}_2$, ...), то в формулу (1.26) вместо силы $\mathbf{F}$ следует подставить *равнодействующую* (т. е. векторную сумму) всех приложенных к телу сил:

$$\mathbf{F} = \mathbf{F}_1 + \mathbf{F}_2 + \ldots = \sum \mathbf{F}_i.$$

Это означает, что каждая сила (любого происхождения или характера) оказывает свое действие на данное тело независимо от наличия или отсутствия других сил, приложенных к этому телу (закон независимости действия сил).

Если одно тело действует на второе с некоторой силой $\mathbf{F}_{12}$, (т. е. деформирует его или сообщает ему ускорение), то всегда имеет место обратное воздействие второго тела на первое с силой $\mathbf{F}_{21}$, равной $\mathbf{F}_{12}$ и противоположно направленной:

$$\mathbf{F}_{12} = -\mathbf{F}_{21}. \tag{1.27}$$

Измерения показывают, что взаимодействующие тела сообщают друг другу ускорения, отношение которых

$$\frac{a_1}{a_2} = \frac{m_2}{m_1}$$

обратно пропорционально отношению их масс. Выше было указано, что это соотношение можно использовать для измерения массы. Допустим, что m_1 — масса эталонного тела, а m_2 — масса измеряемого тела; заставив тела взаимодействовать между собой, можно определить сообщаемые ими друг другу ускорения a_1 и a_2 и, следовательно, по этой формуле найти m_2. Однако надо еще проверить, остается ли найденная нами масса m_2 постоянной по отношению к различным по

величине и характеру силам взаимодействия. Это постоянство массы данного тела должно быть установлено дополнительно.

Два точечных тела притягиваются друг к другу через пространство с силой, прямо пропорциональной их инертным массам и обратно пропорциональной квадрату расстояния между ними:

$$F = G\frac{m_1 m_2}{r^2}, \qquad (1.28)$$

где G — гравитационная постоянная.

Иногда вводят понятие *гравитационной* массы как свойства тел воздействовать на другие тела через пространство. (Очевидно, необходимо предположить, что тела взаимосвязаны с пространством, обладающим определенными физическими свойствами; эта связь проявляется в виде поля тяготения, окружающего каждое тело.) Силу тяготения между двумя телами можно представить в зависимости от их гравитационных масс M_1 и M_2 в виде

$$F = \frac{M_1 M_2}{r^2}.$$

Если выбрать эталонное тело с единичной весомой массой, то весомую массу любого другого тела можно вычислить по этой формуле; необходимо лишь измерить расстояние и силу тяготения между ними.

Таким образом, для характеристики каждого тела могут быть введены две величины: инертная и весомая массы. Измерения показали, что для всех тел эти величины пропорциональны друг другу, т. е.

$$M = \beta m,$$

где β — некоторая постоянная величина, одинаковая для всех тел. Подставив в выражение для силы тяготения $M_1 = \beta m_1$ и $M_2 = \beta m_2$ и обозначив β^2 через G, получим формулу (1.28). Благодаря этой пропорциональности инертные массы тел можно измерять путем взвешивания, так как отношение инертных масс двух тел равно отношению их весомых масс. Если два тела с инертными массами m_1 и m_2 притягиваются к Земле (в одном и том же месте) с силами F_1 и F_2, то

$$\frac{m_1}{m_2} = \frac{M_1}{M_2} = \frac{F_1}{F_2}.$$

Указанные выше законы механики Ньютона (лежащие в основе «классической механики») позволяют найти форму траектории тела, а также скорости и ускорения в различных точках траектории, если известна скорость тела в начальной точке траектории и заданы силы, действующие на это тело в каждый момент времени или в каждой точке, где может находиться тело.

Связь между проекциями силы и ускорения на каждую из трех координатных осей, согласно второму закону механики, выражается следующим образом:

$$F_x = ma_x = m\frac{\mathrm{d}^2x}{\mathrm{d}t^2}; \quad F_y = ma_y = m\frac{\mathrm{d}^2y}{\mathrm{d}t^2}; \quad F_z = ma_z = m\frac{\mathrm{d}^2z}{\mathrm{d}t^2}. \qquad (1.29)$$

Наукова бібліотека
Вінницького

В векторном виде эта связь имеет вид

$$\mathbf{F} = m\mathbf{a} = m\frac{d^2\mathbf{r}}{dt^2}, \qquad (1.30)$$

где $\mathbf{r}$ — радиус-вектор, проведенный от начала координат до той точки траектории, где в данный момент находится движущееся тело.

Соотношения (1.29) между силами и ускорениями вдоль координатных осей называются в механике *дифференциальными уравнениями движения тел*. Каждое из этих уравнений можно решать независимо от других и таким образом выяснять характер движения тела вдоль соответствующей оси. Если в каком-нибудь направлении, например вдоль оси OX, составляющая силы F_x равна нулю, то в этом направлении $d^2x/dt^2 = dv_x/dt = 0$ и движение тела будет равномерным ($v_x = \text{const}$); при этом движение по двум другим направлениям — вдоль осей OY и OZ может иметь любое ускорение. Таким образом, характер движения тела по каждому из трех взаимно перпендикулярных направлений не зависит от того, как движется тело по двум другим направлениям. Это утверждение связано с трехмерностью окружающего нас пространства.

ИМПУЛЬС. РАБОТА. МОЩНОСТЬ

В формулах (1.29) сила F, действующая на тело, разложена на составляющие F_x, F_y и F_z по координатным осям. Однако при решении некоторых задач удобно использовать другое разложение. В § 1 было указано, что полное ускорение тела a целесообразно разложить на два ускорения: тангенциальное a_t, вызывающее изменение скорости тела только по величине, и нормальное a_n, вызывающее изменение скорости только по направлению. В связи с этим силу, действующую на тело, также можно разложить на две составляющие:

1) тангенциальную силу F, ориентированную по направлению движения (по касательной к траектории движения), и

2) нормальную силу N, перпендикулярную направлению скорости движения (т. е. ориентированную по нормали к траектории).

Сила F (которую иногда называют *движущей силой*) сообщает телу тангенциальное ускорение

$$F = ma_t = m\frac{dv}{dt}.$$

Сила N сообщает телу нормальное (центростремительное) ускорение

$$N = ma_n = \frac{mv^2}{R}$$

(R — радиус кривизны траектории в данной точке M). На рис. 1.9 показано разложение сил и ускорений на тангенциальную и нормальную составляющие; сила F используется для расчета изменения величины скорости, а сила N — для расчета радиуса кривизны траектории.

Законы механики в виде (1.29) или (1.30) связывают между собой значения сил и ускорений в каждый рассматриваемый момент времени. Пользуясь ими, можно найти изменение скорости или кинетической

энергии тела за конечный промежуток времени или на конечном участке траектории. Допустим, что на тело массой m действует сила F, величина и направление которой изменяется со временем. Разделим время наблюдения t на такие малые промежутки Δt_i, в течение которых эту силу можно было бы с удовлетворительной точностью полагать постоянной (по величине и направлению). Тогда для каждого из этих промежутков времени можно написать $\mathbf{F} = m\Delta\mathbf{v}/\Delta t$ или $\mathbf{F}\Delta t = m\Delta\mathbf{v}$, т. е.

$$\mathbf{F}_1 \Delta t_1 = m\,\Delta\mathbf{v}_1;$$

$$\mathbf{F}_2 \Delta t_2 = m\,\Delta\mathbf{v}_2 \text{ и т. д.}$$

Суммируя эти соотношения, получим (при постоянной массе тела)

$$\sum \mathbf{F}_i \Delta t_i = \sum m\,\Delta\mathbf{v}_i = m\sum \Delta\mathbf{v}_i = m\mathbf{v} - m\mathbf{v}_0,$$

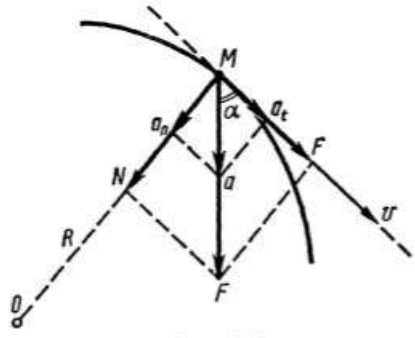


Рис. I.9

где $\mathbf{v}_0$ — начальная скорость тела в момент $t = 0$. Для точности расчета промежутки времени Δt_i должны быть бесконечно малыми, поэтому

$$\int_0^t \mathbf{F}\,dt = m\mathbf{v} - m\mathbf{v}_0. \tag{1.31}$$

Для постоянной по величине и направлению силы это соотношение можно написать и в скалярном виде:

$$Ft = mv - mv_0.$$

Формула (1.31) позволяет рассчитать изменение скорости тела за время t, если известно, как изменяется действующая на нее сила с течением времени.

Аналогичным образом можно получить формулы для расчета изменений скорости вдоль каждой из координатных осей:

$$\int_0^t F_x\,dt = mv_x - mv_{x0};\quad \int_0^t F_y\,dt = mv_y - mv_{y0};\quad \int_0^t F_z\,dt = mv_z - mv_{z0}, \tag{1.32}$$

где v_{x0}, v_{y0} и v_{z0} — проекции начальной скорости v_0 на координатные оси. Если на тело действуют несколько сил (включая силы тяготения, трения, упругости и т. д.), то в эти формулы должна быть подставлена их равнодействующая, т. е. их векторная сумма.

Для иллюстрации формулы (1.31) рассмотрим один пример. Допустим, что на тело действует сила F, всегда перпендикулярная скорости движения и постоянная по величине (с такой силой действует, например, однородное магнитное поле на движущийся электрический заряд). Очевидно, что такая сила будет сообщать телу только нормаль-

ное ускорение и не будет изменять величину скорости, т. е. тело будет двигаться по окружности с постоянной скоростью. Применим к такому движению формулу (1.31) и выберем промежуток времени t, в течение которого тело совершает четверть оборота. Разделим это время на элементарные промежутки Δt и проведем суммирование векторов $\mathbf{F}\Delta t$ (на рис. I.10, *а* они обозначены **f**). Эти векторы будут направлены к центру окружности. Для нахождения их равнодействующей совместим их начала в общую точку, как это показано на рис. I.10, *б*. Векторная сумма $\Sigma \mathbf{F}\Delta t$ будет вектором **I**, также направленным к центру окружности. Так как скорость тела, оставаясь постоянной по величине, изменяется по направлению, то изменение скорости за время наблюдения будет равно вектору $\Delta \mathbf{v} = \mathbf{v}_2 - \mathbf{v}_1$, показанному на рис. I.10, *в*.

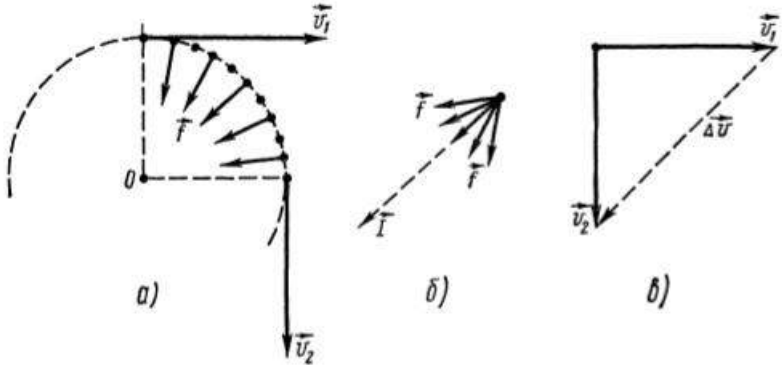


Рис. I.10

Формула (1.31) утверждает, что равнодействующая всех элементарных векторов $\mathbf{F}dt$ $\left(\text{т. е. } \mathbf{I} = \int_0^t \mathbf{F}\,dt\right)$ равна по величине и направлению вектору $m\Delta\mathbf{v} = m\mathbf{v}_2 - m\mathbf{v}_1$. Если тело сделает один оборот, то равнодействующая **I** будет равна нулю; так как вектор скорости тела при этом вернется к первоначальному направлению, то изменение скорости $\Delta\mathbf{v}$ также будет равно нулю.

Произведение массы тела на скорость его движения ($m\mathbf{v}$) называется *импульсом* (количеством движения) тела. Формула (1.31) показывает, что изменение импульса тела за время наблюдения равно интегралу $\int_0^t \mathbf{F}(t)\,dt$, где **F** — равнодействующая всех сил, действующих на данное тело.

Воздействие силы на тело можно определить в зависимости не только от времени наблюдения, но и от пути, пройденного движущимся телом. Будем интересоваться изменением скорости только по величине, т. е. выведем формулу, которая показывала бы воздействие силы на тело независимо от формы траектории. Так как величина скорости изменяется только тангенциальной силой, то второй закон механики

($\mathbf{F} = m\mathrm{d}\mathbf{v}/\mathrm{d}t$) перепишем в скалярном виде:

$$F\cos\alpha = m\frac{\mathrm{d}v}{\mathrm{d}t},$$

где α — угол между направлениями силы $\mathbf{F}$ и ее тангенциальной составляющей (см. рис. I.9). Выразим время dt через пройденный путь ($\mathrm{d}t = \mathrm{d}s/v$) и перепишем это уравнение в виде

$$F\cos\alpha\,\mathrm{d}s = mv\,\mathrm{d}v = \mathrm{d}(mv^2/2).$$

Если тело прошло по траектории (любой формы) путь s, то

$$\int_s F\cos\alpha\,\mathrm{d}s = \frac{mv^2}{2} - \frac{mv_0^2}{2}. \qquad (1.33)$$

Произведение силы на пройденный путь и на косинус угла между направлением силы и направлением движения (скорости) называется *работой силы* на этом участке траектории. Формула (1.33) показывает, что работа, совершаемая силой на некотором конечном участке траектории, равна изменению кинетической энергии тела, к которому эта сила приложена. Разумеется, F есть равнодействующая всех сил, действующих на данное тело.

Если вдоль траектории (любой формы) сила остается постоянной по величине и направлению (например, при движении тела в поле тяготения Земли по траекториям небольших размеров), то работа этой силы будет равна

$$\int_s F\cos\alpha\,\mathrm{d}s = F\int_s \cos\alpha\,\mathrm{d}s = Fs_0,$$

где s_0 — проекция пройденного пути на направление силы.

Работа, совершенная переменной силой на элементарном участке пути, может быть записана в виде скалярного произведения векторов силы $\mathbf{F}$ и перемещения $\mathrm{d}\mathbf{s}$: $\mathrm{d}A = \mathbf{F}\mathrm{d}\mathbf{s} = F\mathrm{d}s\cos\alpha$.

Нормальная составляющая действующей силы не изменяет численного значения скорости тела и поэтому работы не совершает. Работа совершается только тангенциальной составляющей действующей силы, ибо только она изменяет величину скорости тела. Отметим, что:

1) работа $A = Fs\cos\alpha$ может быть положительной, если $\cos\alpha > 0$, и отрицательной, если $\cos\alpha < 0$. Если сила совершает отрицательную работу, то скорость тела уменьшается: $v < v_0$;

2) при горизонтальном движении тела сила тяжести работы не совершает ($\cos\alpha = 0$);

3) работа, совершаемая силой тяжести при движении тела в поле тяготения, не зависит от формы траектории, а определяется только разностью высот h начальной и конечной точек траектории (рис. I.11). На малом отрезке траектории Δs_i

$$\Delta A_i = P\,\Delta s_i\cos\alpha_i = P\,\Delta h_i;$$

суммарная работа

$$A = \sum P\,\Delta h_i = P\sum \Delta h_i = Ph.$$

Здесь предполагается, что высота h невелика и поэтому сила тяжести остается в процессе движения тела постоянной;

4) если в поле тяготения тело описывает замкнутую траекторию, то суммарная работа силы тяжести равна нулю, так как положительная работа, совершаемая при опускании тела, компенсируется отрицательной работой при подъеме;

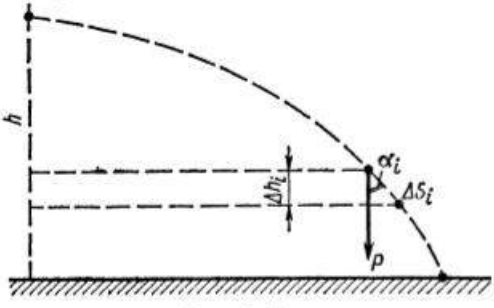


Рис. I.11

5) при движении тела по криволинейной траектории центростремительная сила работы не совершает ($\cos\alpha = 0$);

6) работа, совершаемая несколькими силами, приложенными к данному телу, равна работе равнодействующей этих сил, найденной путем их векторного сложения.

Работа и энергия выражаются в СИ в джоулях (Дж) и в СГС — в эргах: 1 джоуль = 1 ньютон · 1 метр = 1 кг · м^2/с^2 = 10^7 эрг.

Работу силы F на участке пути ds можно представить как сумму работ, совершаемых компонентами этой силы по осям координат, т. е.

$$F\,ds = F_x\,dx + F_y\,dy + F_z\,dz. \qquad (1.34)$$

Действительно, если в эту формулу подставить $F_x = F\cos\alpha$, $F_y = F\cos\beta$, $F_z = F\cos\gamma$, а также $dx = ds\cos\alpha$, $dy = ds\cos\beta$, $dz = ds\cos\gamma$ и учесть, что $\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$, то получим тождество. Формулу (1.34) можно получить также из выражения для работы: $Fds = d(mv^2/2) = (m/2)d(v_x^2 + v_y^2 + v_z^2) = mv_x dv_x + mv_y dv_y + mv_z dv_z = ma_x dx + ma_y dy + ma_z dz = F_x dx + F_y dy + F_z dz$.

Мощностью источника энергии (или работы) называется отношение работы dA, совершаемой этим источником за время dt, ко времени dt:

$$N = dA/dt. \qquad (1.35)$$

Для конечных промежутков времени это отношение определяет среднюю мощность. Мощность в СИ выражается в ваттах (Дж/с): 1 Вт = 10^7 эрг/с.

Подставим в формулу (1.35) вместо dA произведение Fds; тогда мощность

$$N = F\,ds/dt = Fv.$$

В общем случае, когда угол между векторами силы и скорости отличен от нуля, мощность равна скалярному произведению этих векторов:

$$N = Fv\cos\alpha = (\mathbf{Fv}).$$

ЕДИНИЦЫ ФИЗИЧЕСКИХ ВЕЛИЧИН

Каждая из величин, используемых при описании механических явлений (расстояние, время, скорость, ускорение, масса, сила и т. д.), а также любая физическая величина должны иметь отчетливое (однозначное) *определение,* из которого должно

быть совершенно ясно, что понимается под этой величиной, каким образом следует производить ее измерение и как находится результат измерения. Для каждой физической величины результат измерения выражается в определенных единицах.

Измерение некоторых физических величин требует использования заранее выбранных *эталонов* этой величины. Например, для измерения расстояний необходимо иметь эталон длины — метр; для измерения времени используются эталонные часы, согласованные с вращением Земли вокруг своей оси или Солнца; для измерения массы необходимо заранее выбрать эталон массы и т. д. Эталоны физических величин не должны изменяться с течением времени или в процессе измерений. Поэтому некоторые из них хранятся в стационарных условиях в Палате мер и весов; для практических измерений используются копии с этих эталонов, тщательно сверенные с оригиналами. Физические величины, для измерения которых выбраны соответствующие эталоны, называются *основными*. Ввиду того, что постоянство (неизменяемость) эталонов в процессе измерений невозможно контролировать и обеспечить (см. Заключение), число основных величин должно быть минимальным.

Кроме основных в физике применяется большое число *производных* величин, измерение которых сводится к измерению основных величин. Например, для измерения скорости и ускорения необходимо найти время и путь, пройденный движущимся телом. Таким образом, для каждой производной величины указывается определяющая формула, показывающая, каким образом должна рассчитываться эта величина по результатам измерений основных величин.

Определения, способы измерения и единицы основных и производных физических величин составляют так называемую систему единиц. Существует несколько различных систем единиц, которые отличаются друг от друга по составу основных величин. В системах типа **LMT** основными величинами выбраны: длина (расстояние) L, время T и масса M. В системах типа **LFT** основными являются длина, время и сила F. В настоящее время преимущественной системой единиц принята Международная система единиц (сокращенно называемая СИ, т. е. *система интернациональная*), относящаяся к системам типа **LMT**. Кроме того, в физике широко применяется другая система того же типа — СГС. Система типа **LFT** применяется в технических измерениях, в которых используется единица силы — килограмм (кгс); однако, в связи с введением Международной системы единиц системы типа **LFT** постепенно вытесняются из измерительной практики.

В Международной системе единиц физических величин (СИ) даны следующие определения для основных величин.

Метр — длина, равная 1650763,73 длины волны в вакууме излучения, соответствующего переходу между уровнями $2p_{10}$ и $5d_5$ атома криптона-86.

Это определение позволяет воспроизводить метр с гораздо большей точностью (10^{-9}), чем снятие копий от платиноиридиевого эталона с двумя штрихами. Указанное излучение соответствует оранжевой линии в спектре криптона с атомным весом 86, а символы $2p_{10}$ и $5d_5$ условно обозначают два стационарных состояния этого атома, переход между которыми сопровождается указанным излучением.

Килограмм — единица массы — представлен массой Международного эталона килограмма. Масса этого эталона на 0,000028 кг больше массы одного кубического дециметра химически чистой воды при температуре, соответствующей наибольшей ее плотности (3,98° С).

Секунда — 9 192 631 770 периодов излучения, соответствующего переходу между двумя сверхтонкими уровнями основного состояния атома цезия-133.

Пользуясь единицами этих основных физических величин, можно дать определение для единиц других величин. Например, в СИ *ньютон* — сила, сообщающая телу массой в 1 кг ускорение 1 м/с^2 в направлении силы (или сила, изменяющая скорость тела массой в 1 кг на 1 м/с за секунду):

$$1 \text{ ньютон} = 1 \text{ килограмм} \cdot 1 \text{ метр}/1 \text{ (секунда)}^2, \text{ или}$$

$$1\text{Н} = 1 \text{ кг} \cdot \text{м}/\text{с}^2.$$

В системе СГС длина выражается в сантиметрах, масса — в граммах, время — в секундах, сила — в динах:

$$1 \text{ дина} = 1 \text{ г} \cdot \text{см}/\text{с}^2 = 10^{-5} \text{ ньютон}.$$

§ 3. ДВИЖЕНИЕ ТЕЛА В ПОЛЕ ТЯГОТЕНИЯ ЗЕМЛИ

ГРАВИТАЦИОННОЕ ПОЛЕ. ВЕС ТЕЛА

Закон всемирного тяготения Ньютона (1.28) написан для *точечных* тел. Для расчета силы тяготения между телами, имеющими большие размеры, необходимо представить их в виде совокупности точечных тел. Сначала определяют векторную сумму сил тяготения, действующих на каждую частицу первого тела со стороны всех частиц второго тела; затем векторно складывают силы, приложенные ко всем частицам первого тела. Так находят величину и точку приложения силы тяготения F_1, действующей на первое тело со стороны второго тела. Аналогичный расчет определит силу F_2 и точку ее приложения для второго тела. *Силы F_1 и F_2 всегда равны по величине и направлены вдоль линии, соединяющей точки их приложения.*

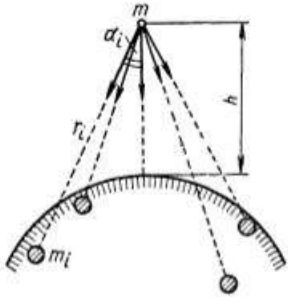


Рис. 1.12

При расчете силы притяжения к Земле точечного тела массой m, находящегося на высоте h (рис. 1.12), следует учесть симметричное расположение массы Земли относительно линии, соединяющей тело с центром Земли. Вследствие этого равнодействующая всех сил, приложенных к телу, будет направлена вертикально вниз и равна

$$F = \sum G \frac{m m_i}{r_i^2} \cos \alpha_i = mg,$$

где $g = \sum G \frac{m_i}{r_i^2} \cos \alpha_i$.

Величина $g = F/m$ есть ускорение, которое сообщается телу силой притяжения к Земле. Это ускорение зависит от высоты h и не зависит от массы m и вещества тела.

Расчет силы взаимного притяжения шарообразных тел, у которых массы распределены равномерно по объему (сферически симметрично), приводит к той же формуле (1.28), причем r означает расстояние между центрами шаров. Полагая Землю таким шаром, можно силу притяжения к ней тел рассчитать по формуле

$$F = G \frac{mM}{r^2},$$

где M — масса Земли; r — расстояние от тела до центра Земли; G — гравитационная постоянная, равная $6{,}672 \cdot 10^{-11}$ Н·м²/кг². Точка приложения силы притяжения данного тела к Земле называется центром тяжести этого тела.

Наличие тяготения между телами приводит к представлению о г р а в и т а ц и о н н о м п о л е (как особой форме материи), в пределах которого на каждое тело действует сила, прямо пропорциональ-

ная массе этого тела. В разделе «Электричество и магнетизм» будет указано на существование и других полей — электрических и магнитных. Гравитационное, электрическое и магнитное поля представляют собой разновидности с и л о в ы х п о л е й; на частицы, помещенные в каждой точке этих полей, действуют силы, прямо пропорциональные определенному физическому свойству этих частиц: массе, электрическому заряду и т. п. Земля окружена гравитационным полем (или полем тяготения), в котором на тела действуют силы, пропорциональные их массам.

В каждой точке поля можно определить отношение силы, действующей на точечное тело, к массе этого тела; это отношение не зависит от вещества и массы тела и равно ускорению, сообщаемому силой тяготения в данной точке поля:

$$\frac{F}{m} = g; \quad g = G\frac{M}{r^2}. \tag{1.36}$$

П о т е н ц и а л о м данной точки поля тяготения называется отношение работы A, которую нужно затратить для переноса точечного тела массой m из данной точки в бесконечность, к массе этого тела; для точки, расположенной на расстоянии r от центра Земли (вне ее объема), работа переноса и потенциал поля тяготения φ равны:

$$A = \int_r^\infty F\,\mathrm{d}r\cos\alpha = -\int_r^\infty G\frac{mM}{r^2}\,\mathrm{d}r = -G\frac{mM}{r}; \quad \varphi = \frac{A}{m} = -G\frac{M}{r}.$$

Работа и потенциал получились отрицательными, так как силы тяготения препятствуют этому перемещению ($\cos\alpha = -1$).

Работа переноса точечной массы из одной точки поля тяготения *1* в другую *2* равна

$$A = m(\varphi_1 - \varphi_2). \tag{1.37}$$

В гравитационном поле работа переноса не зависит от формы и размеров траектории, по которой происходит перемещение, а определяется только потенциалами начальной и конечной точек этой траектории. В таких полях (называемых потенциальными) можно провести *эквипотенциальные поверхности*, все точки которых имеют одинаковый потенциал. При перемещении точечного тела по эквипотенциальной поверхности силы тяготения, согласно формуле (1.37), работы не совершают. Это означает, что силы тяготения всегда ориентированы перпендикулярно эквипотенциальным поверхностям. У шаровых тел с равномерным распределением массы по объему эквипотенциальные поверхности в поле тяготения — сферические, а силы тяготения направлены по радиусам к центру шара.

Рассмотрим тело, покоящееся на поверхности Земли; на это тело действуют три силы (рис. 1.13, *а*): сила притяжения к Земле (по закону тяготения, F), сила Архимеда (F_A) и сила N, приложенная к телу со стороны опоры, на которой это тело лежит (или к которой оно подвешено). Так как тело участвует в суточном вращении Земли с угловой скоростью ω, то необходимая центростремительная сила ($F_{ц.с} =$

$= m\omega^2 r$) должна быть равна равнодействующей всех упомянутых выше сил, т. е.

$$\mathbf{F}_{цс} = \mathbf{F} + \mathbf{F}_A + \mathbf{N}.$$

Силы F и F_A всегда направлены по вертикали; сила же $F_{цс}$ будет иметь вертикальное направление только на экваторе (на полюсе $F_{цс} = 0$). Поэтому численное значение силы N для тел, покоящихся на поверхности Земли вблизи экватора, будет равно $N = F - F_A - F_{цс}$, а на полюсе $N = F - F_A$.

Вместо силы N, с которой опора действует на тело, будем рассматривать силу P, с которой тело действует на опору. Согласно третьему

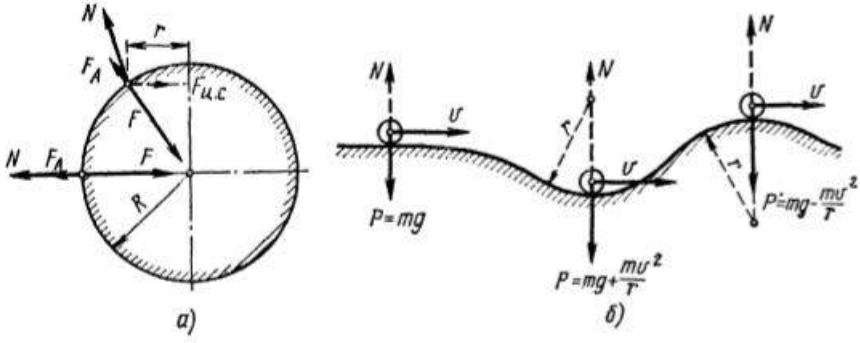


Рис. I.13

закону механики, сила P численно равна N и имеет противоположное направление. В частности, для тел, покоящихся на экваторе, пренебрегая силой Архимеда, мы получим

$$P = F - m\omega^2 R; \quad F = P + m\omega^2 R.$$

Эта формула показывает, что сила тяготения F для покоящегося тела может быть разложена на две составляющие: одна составляющая, равная $m\omega^2 R$, сообщает телу центростремительное ускорение $\omega^2 R$, а другая составляющая, равная P, действует на опору и деформирует ее.

Сила P называется *весом* тела. Так как $\omega^2 r \ll g$, то $P \approx F$, т. е. вес покоящегося тела можно приравнять силе, с которой это тело притягивается к Земле. Однако если бы угловая скорость вращения Земли вокруг своей оси была очень велика, то вес тела оказался бы значительно меньше силы притяжения к Земле. Например, при $\omega^2 R = g$ вес тела на экваторе был бы равен нулю. Заметим, что вследствие суточного вращения Земли ускорение свободного падения тел на Земле несколько меньше ускорения силы тяготения g.

Очевидно, имеет смысл различать:

1) силу притяжения тел к Земле по закону всемирного тяготения. Эта сила тяготения ($F_{тяг}$) не зависит от вращения Земли и от того, покоится или движется данное тело в поле тяготения Земли. Сила тяготения сообщает телам ускорение g;

2) силу тяжести ($F_{тяж}$) — разность между силой тяготения и центростремительной силой:

$$\mathbf{F}_{тяж} = \mathbf{F}_{тяг} - \mathbf{F}_{цс}.$$

Наблюдаемое на Земле ускорение свободного падения сообщается телам силой тяжести. Лишь пренебрегая $F_{цс}$ по сравнению с $F_{тяг}$, можно полагать, что ускорение свободного падения равно ускорению силы тяготения. На полюсе $F_{цс} = 0$; $F_{тяж} = F_{тяг}$; на экваторе разность между этими силами имеет максимальное значение;

3) вес тела (P) — сила, с которой тело действует на опоры, препятствующие его свободному падению. Для покоящегося тела вес равен силе тяжести: $P = F_{тяж}$. Пренебрегая $F_{цс}$, можно полагать, что для покоящегося тела $P \approx mg$.

Допустим теперь, что тело движется по поверхности Земли с постоянной по величине скоростью, причем траектория движения имеет сложный профиль (рис. I.13, *б*). Будем рассматривать только те точки траектории, в которых скорость движения тела направлена горизонтально; центростремительной силой, обусловленной вращением Земли, будем пренебрегать. Тогда на тело будут действовать две силы: сила тяжести $F = mg$ и сила N, приложенная со стороны несколько деформированной поверхности Земли, по которой движется тело. На горизонтальном участке пути ускорение движения равно нулю, поэтому $F = N = P$. При движении тела по вогнутому участку пути сила N будет больше силы F, так как центростремительное ускорение v^2/r сообщается силой N. Следовательно, вес тела, численно равный N, также будет больше силы тяжести:

$$P = F + mv^2/r = m(g + v^2/r).$$

Если же тело движется по выпуклому участку пути, то центростремительное ускорение сообщается силой F. Разность между $F = mg$ и центростремительной силой mv^2/r будет действовать на поверхность и вызовет ее реакцию N. В этом случае

$$P = F - mv^2/r = m(g - v^2/r).$$

При больших v и малых r можно получить $N = P = 0$; например, при $v = 180$ км/ч $= 50$ м/с и $r = 250$ м нормальное ускорение v^2/r, направленное вверх, равно ускорению силы тяжести, следовательно, вес тела P и сила N будут равны нулю. Если же $v^2/r > g$, то тело оторвется от поверхности Земли, так как сила тяжести F будет недостаточна для того, чтобы обеспечить движение тела с данной скоростью по дуге радиуса r. Сила тяжести сообщает ускорение, равное g, поэтому тело, имеющее очень большую скорость и удалившееся от поверхности Земли ($N = 0$), будет описывать траекторию, в каждой точке которой полное ускорение равно g.

Итак, *вес тела появляется в том случае, если в поле тяготения тело вынуждено двигаться с ускорением, отличным от g*. Это возможно, если на тело кроме силы тяготения действуют и другие силы. Вес тела есть та сила, с которой данное тело действует на другие тела, препятствующие его свободному движению в поле тяготения. Если же

тело движется только под действием силы тяготения, то оно является *невесомым* (при этом траектория движения может быть любой: прямолинейной, параболической, эллиптической или круговой). При любом состоянии движения (весомом или невесомом) сила тяготения действует всегда.

Для того чтобы в поле тяготения Земли данное тело двигалось с ускорением **a**, отличным от **g**, к нему должна быть приложена со стороны других тел дополнительная сила **N**, удовлетворяющая условию

$$\mathbf{N} + \mathbf{F} = m\mathbf{a}.$$

Тогда вес тела (т. е. сила, с которой данное тело действует на другие тела) будет равен

$$\mathbf{P} = -\mathbf{N} = m\mathbf{g} - m\mathbf{a} = m(\mathbf{g} - \mathbf{a}). \qquad (1.38)$$

Если тело покоится или движется прямолинейно и равномерно, то $\mathbf{a} = 0$ и $\mathbf{P} = m\mathbf{g}$. Если тело *свободно* движется в поле тяготения по любой траектории и в любом направлении, то $\mathbf{a} = \mathbf{g}$, $\mathbf{P} = 0$ и тело будет *невесомым*. Например, невесомы все тела, находящиеся внутри кабины космических кораблей, свободно движущихся в безвоздушном пространстве. Если же это движение не свободно, например существует сопротивление атмосферы или работают реактивные двигатели, установленные на корабле, то ускорение тел будет отличаться от **g**.

Физиологическое ощущение весомости для человека обусловлено совместным действием распределенных и сосредоточенных сил. Например, если человек стоит на опоре, то распределенная по объему его тела сила тяжести уравновешивается сосредоточенной силой, приложенной к его ногам со стороны опоры. При этом тело несколько деформируется и соответствующие мышцы оказываются напряженными. При свободном движении тела в поле тяготения (по любой траектории) сосредоточенные силы со стороны опор отсутствуют; распределенная же сила тяжести никакой деформации организма не вызывает, так как все части тела движутся с одинаковым ускорением, равным **g**. Это особое состояние организма и называется состоянием невесомости.

ДВИЖЕНИЕ КОСМИЧЕСКИХ ОБЪЕКТОВ

Закон всемирного тяготения позволяет определить массы Земли, Солнца, Луны и планет. Например, согласно формуле (1.36), масса Земли равна

$$M = gR^2/G. \qquad (1.39)$$

Постоянная тяготения была тщательно измерена в лабораторных условиях; ускорение свободного падения определяется по колебаниям маятника; на широте 45° оно равно $g = 9{,}806$ м/с^2. Расчет по формуле (1.39) дает массу Земли $M = 5{,}98 \cdot 10^{24}$ кг. Масса солнца M_C может быть рассчитана по периоду обращения (T) планет, например Земли, вокруг Солнца. Предположим, что планета вращается по круговой

орбите; так как сила тяготения является центростремительной силой, то

$$G\frac{MM_C}{r^2}=\frac{Mv^2}{r};\quad M_C=\frac{4\pi^2r^3}{GT^2}$$

(линейная скорость планеты по орбите $v=2\pi r/T$).

Для двух планет получим

$$\frac{r_1^3}{r_2^3}=\frac{T_1^2}{T_2^2}, \tag{1.40}$$

т. е. *квадраты времен обращения планет относятся как кубы их расстояний до Солнца* (закон Кеплера). Этим формулам подчиняются также и орбиты искусственных спутников.

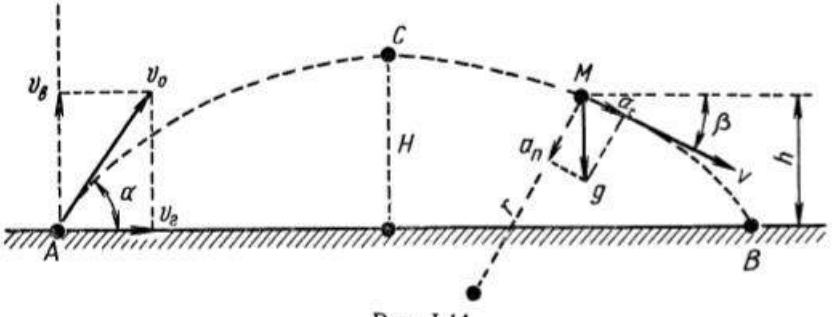


Рис. I.14

Рассмотрим движение тел малых размеров на небольших расстояниях от поверхности Земли. Вертикальное движение можно полагать равноускоренным (а при движении вверх — равнозамедленным). Если же изменение высоты тела соизмеримо с радиусом Земли, то ускорение свободного падения будет заметно изменяться; на высоте h оно равно

$$g_h=g_0\frac{R^2}{(R+h)^2}=\frac{g_0}{\left(1+\frac{h}{R}\right)^2}\approx g_0\left(1-2\frac{h}{R}\right).$$

Например, при подъеме на высоту $h=6{,}35$ км сила притяжения к Земле, а следовательно, и ускорение свободного падения уменьшаются приблизительно на 0,2%.

Задача о движении тел, брошенных под углом α к горизонту, решается просто только при отсутствии сопротивления воздуха и если кривизной линии AB (рис. I.14) можно пренебречь. В этом случае для определения траектории тела разлагают начальную скорость v_0 на горизонтальную $v_г=v_0\cos\alpha$ и вертикальную $v_в=v_0\sin\alpha$ составляющие и затем полагают, что горизонтальное движение происходит равномерно, а независимое от него вертикальное движение — равнозамедленно на участке AC и равноускоренно на участке CB.

Для произвольной точки траектории M имеем:

$$v=\sqrt{v_в^2+v_г^2};\quad v_в=\sqrt{2g(H-h)};\quad \operatorname{tg}\beta=\frac{v_в}{v_г};$$

нормальное и тангенциальное ускорения и радиус кривизны траектории в этой точке равны:

$$a_n = g\cos\beta = \frac{v^2}{r};\quad a_t = g\sin\beta;\quad r = \frac{v^2}{a_n} = \frac{v^2}{g\cos\beta}.$$

При увеличении начальной скорости тела (v_0) будут увеличиваться размеры описываемой им траектории; если эта скорость окажется достаточно большой, то тело уйдет в окружающее Землю космическое пространство. Допустим, что C есть та точка на траектории, в которой скорость тела v перпендикулярна прямой OC (рис. I.15). Скорость v может быть больше v_0, если на участке AC телу сообщалось ускорение, например, реактивными двигателями. Рассчитаем скорость v_1, необходимую для того, чтобы дальнейшее свободное движение тела происходило по окружности радиуса $r = R + h$, т. е. чтобы тело вышло

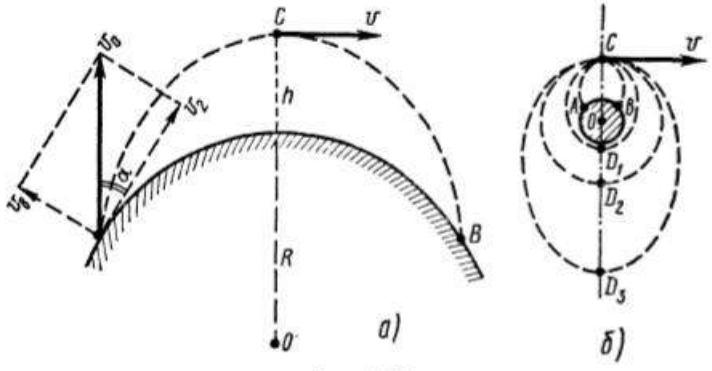


Рис. I.15

на круговую орбиту вокруг Земли. Вдоль орбиты сила тяготения производит только одно действие — сообщает телу центростремительное ускорение, поэтому искомая скорость равна

$$\frac{mv_1^2}{r} = G\frac{mM}{r^2};\quad v_1 = \sqrt{G\frac{M}{r}} = \sqrt{gr}.$$

Скорость v_1 называется *первой космической*; вблизи поверхности Земли ($h = 0, r = R = 6350$ км, $g = 9{,}81$ м/с^2) она равна $v_1 \approx 7800$м/с.

Допустим, что в точке C скорость тела v меньше v_1; тогда при данном радиусе кривизны траектории в этой точке $r = R + h$ сила тяготения будет больше необходимой центростремительной силы и тело в своем дальнейшем движении будет приближаться к Земле с возрастающей скоростью. В зависимости от величины v оно может либо ($v \ll v_1$) упасть на Землю (в точке B; рис. I.15, *а*), либо пройти мимо противоположной поверхности Земли на расстоянии OD_1 и вернуться в точку C. В этом случае тело будет двигаться по эллиптической траектории ($OD_1 < OC$; рис. I.15, *б*). Если $v = v_1$, то тело, как указывалось выше, описывает круговую орбиту. При $v > v_1$ сила тяготения будет недостаточна для удержания тела на круговой орбите и тело удаляется

от Земли; орбита примет эллиптическую форму ($OD_3 > OC$). По мере увеличения v точка D_3 будет удаляться от центра Земли и при очень большой скорости тело перестанет быть спутником Земли.

Рассчитаем минимальную скорость v_2, которую необходимо сообщить телу на поверхности Земли, чтобы оно (при отсутствии сопротивления атмосферы) могло преодолеть земное притяжение и уйти в космическое пространство. Для этого кинетическая энергия тела должна быть равна работе, совершаемой против силы тяготения:

$$\frac{mv_2^2}{2} = \int_R^\infty G\frac{mM}{r^2}\,dr = G\frac{mM}{R},$$

откуда

$$v_2 = \sqrt{2G\frac{M}{R}} = \sqrt{2gR}.$$

Эта скорость называется *второй космической*; у поверхности Земли она равна $\sim$ 11 200 м/с. Заметим, что $v_2 = \sqrt{2}\cdot v_1$.

Для того чтобы тело преодолело притяжение к Солнцу и ушло в межзвездное пространство, его начальная скорость у поверхности Земли должна быть еще больше. Эту скорость называют *третьей космической*; она оказывается приблизительно равной 42 000 м/с.

Первый искусственный спутник Земли был успешно запущен на орбиту 4 октября 1957 г. в Советском Союзе. Наибольшая высота орбиты над Землей была 950 км, наименьшая — 227 км. Начальный период обращения спутника был равен 96,2 мин. Спутник имел форму шара диаметром 58 см и массу 83,6 кг.

12 апреля 1961 г. был впервые осуществлен полет человека в космическое пространство. Летчик-космонавт Ю. А. Гагарин на космическом корабле «Восток-1» совершил полет по орбите, наибольшая высота которой равнялась 327 км, наименьшая — 181 км. Полет длился 1 ч 48 мин.

§ 4. ДВИЖЕНИЕ ТЕЛ С ПЕРЕМЕННОЙ МАССОЙ

Движение некоторых тел сопровождается непрерывным изменением их массы; например, масса движущейся капли может уменьшаться вследствие испарения или, наоборот, увеличиваться при конденсации паров на ее поверхности; масса ракеты изменяется при выбрасывании продуктов сгорания; по той же причине изменяется масса самолета, расходующего для своего движения запасы топлива, и т. д. Изменение массы тел приводит к некоторому усложнению формул, по которым рассчитывается их движение.

Если система выбрасывает часть своей массы в каком-нибудь определенном направлении, то она получает импульс (количество движения) в противоположном направлении. Это есть *принцип реактивного движения*, который имеет широкое применение; на нем основаны ракетная техника, расчеты реактивных двигателей самолетов и т. д.

Выведем уравнение движения тел с уменьшающейся массой при некоторых упрощающих предположениях. Допустим, что в начальный момент времени тело с массой M_0 покоилось относительно некоторой системы отсчета, связанной, например, с Землей. По истечении времени t масса тела сделалась равной M, а скорость — v. За каждый промежуток времени dt от тела отделяется масса dm, причем будем предполагать, что по окончании процесса отделения каждая из этих элементарных масс имеет одну и ту же конечную скорость u. Далее предположим, что на тело не действуют внешние силы, поэтому выбрасывание массы dm производится силами взаимодействия $\mathbf{f}_1$ и $\mathbf{f}_2$ между телом и отделяющимися частями его. Эти внутренние силы по третьему закону механики равны по величине и противоположны по направлению. За время dt масса тела уменьшается на dM, а скорость увеличивается на dv. Сила $\mathbf{f}_1$, действующая на массу $(M - dM)$, изменяет ее импульс на величину, равную

$$\mathbf{f}_1\,dt = (M - dM)(\mathbf{v} + d\mathbf{v}) - (M - dM)\,\mathbf{v}.$$

Пренебрегая бесконечно малыми второго порядка, получим

$$\mathbf{f}_1\,dt = M\,d\mathbf{v}. \tag{1.41}$$

Сила $\mathbf{f}_2$, действующая на выбрасываемую массу dm, изменяет скорость ее движения от начального значения $\mathbf{v}$ до конечного $\mathbf{u}$, т. е.

$$\mathbf{f}_2\,dt = (\mathbf{u} - \mathbf{v})\,dm. \tag{1.42}$$

Так как $\mathbf{f}_1 = -\mathbf{f}_2$, а отделяющаяся масса dm равна уменьшению массы тела, т. е. $dm = -dM$, то импульс (количество движения, приобретаемое телом за время dt), будет равен

$$M\,d\mathbf{v} = (\mathbf{u} - \mathbf{v})\,dM. \tag{1.43}$$

Разность скоростей $\mathbf{u} - \mathbf{v} = \mathbf{w}$ есть скорость отделяющихся масс относительно самого тела (по абсолютному значению $w = u + v$); для ракеты это есть средняя скорость выбрасываемых продуктов сгорания относительно корпуса ракеты. Так как $\mathbf{w}$ направлена противоположно скорости $d\mathbf{v}$, то при замене векторного уравнения (1.43) скалярным вместо $u - v$ следует написать $-w$; тогда

$$dv = -w\,\frac{dM}{M}. \tag{1.44}$$

Знак минус означает, что увеличение скорости тела (положительное dv) сопровождается уменьшением массы тела (отрицательное dM). Если дополнительно предположить, что скорость w отделяющихся масс относительно самого тела сохраняется в процессе движения постоянной, то уравнение (1.44) легко интегрируется:

$$v - v_0 = w \ln\frac{M_0}{M} = w \ln\left(1 + \frac{M_0 - M}{M}\right). \tag{1.45}$$

Из этой формулы, полученной для ракет выдающимся теоретиком космонавтики К. Э. Циолковским, следует, что приращение скорости ракеты $(v - v_0)$ за конечный промежуток времени определяется ско-

ростью истечения газов из выходного сопла ракеты (w) и отношением массы сожженного топлива ($\Delta M = M_0 - M$) к оставшейся массе ракеты (M). Например, если $w = 2500$ м/с, то для достижения конечной скорости $v = 12\,500$ м/с (при $v_0 = 0$) необходимо отношение массы горючего к массе ракеты, равное 89.

Для ракет и реактивных двигателей сила $\mathbf{f}_1$, приложенная к корпусу ракеты или двигателя со стороны продуктов сгорания, называется *силой тяги*. Для ракет с жидким и твердым топливом (не потребляющих атмосферного воздуха) отделяющиеся массы имеют начальную скорость (до сгорания), равную скорости $\mathbf{v}$ корпуса ракеты, и конечную скорость (вне ракеты), равную $\mathbf{u}$, поэтому

$$\mathbf{f}_1 = -\mathbf{f}_2 = (\mathbf{u} - \mathbf{v})\frac{dM}{dt} = \mathbf{w}\frac{dM}{dt}.$$

Например, если $w = 2500$ м/с, а ежесекундный расход топлива равен 200 кг/с, то сила тяги будет равна 500 000 Н. У воздушно-реактивных двигателей расход топлива мал по сравнению с количеством воздуха, проходящим через двигатель; расчет силы тяги производится по изменению импульса (количества движения) воздуха, прошедшего за секунду через двигатель.

В этих расчетах предполагалось, что внешние силы отсутствуют. Если же на тело с переменной массой действуют внешние силы (например, притяжение к Земле, сопротивление атмосферы и т. п.), то полное изменение импульса

$$M\,d\mathbf{v} = \mathbf{F}\,dt + \mathbf{f}_1\,dt.$$

Произведение $\mathbf{f}_1 dt$, согласно уравнениям (1.41) и (1.42), равно $(\mathbf{u} - \mathbf{v})dM$, поэтому

$$M\frac{d\mathbf{v}}{dt} = \mathbf{F} + (\mathbf{u} - \mathbf{v})\frac{dM}{dt}. \qquad (1.46)$$

Это одно из уравнений И. В. Мещерского, позволяющее решать ряд важных задач движения тел с переменной массой. Оно сохраняется и в том случае, когда масса тела не уменьшается, как предполагалось выше, а увеличивается (dM будет положительной величиной). В частности, если $u = 0$, т. е. отделяющиеся или присоединяемые к телу массы имеют нулевую скорость (относительно указанной вначале системы отсчета), то уравнение (1.46) принимает вид

$$M\frac{d\mathbf{v}}{dt} + \mathbf{v}\frac{dM}{dt} = \frac{d}{dt}(M\mathbf{v}) = \mathbf{F}. \qquad (1.47)$$

Выше рассматривалось движение тел, у которых изменение массы происходит вследствие выбрасывания его составных частей (или присоединения других тел, например частиц окружающей среды). Однако изменение массы как свойства тел может произойти также и вследствие изменения состояния этого тела, в частности изменения его скорости. Например, измерения показали, что масса электрона (а также других элементарных частиц) изменяется со скоростью

согласно формуле

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}, \tag{1.48}$$

где m_0 — масса покоящегося электрона; c — скорость распространения света в вакууме. Для таких тел второй закон механики в виде $\mathbf{F} = m\mathrm{d}\mathbf{v}/\mathrm{d}t$ непригоден и заменяется более общей формулой

$$\mathbf{F} = \frac{\mathrm{d}}{\mathrm{d}t}(m\mathbf{v}). \tag{1.49}$$

§ 5. ПРИНЦИП ОТНОСИТЕЛЬНОСТИ В МЕХАНИКЕ

Классическая механика основана не только на законах механики Ньютона, перечисленных в § 2, но и на некоторых предположениях и представлениях о том, как устроена окружающая нас природа. Обсуждение важнейших из них приведено в Заключении; здесь же приведем некоторые предварительные рассуждения.

При формулировке законов Ньютона предполагается, что:

1) любое ускорение создается силой; ускоренное движение тела, на которое не действуют никакие силы, невозможно;

2) всякая сила, приложенная к данному телу и сообщающая ему ускорение, есть воздействие на это тело со стороны других тел (и связанных с ними полей).

В механике Ньютона не допускается, чтобы рассматриваемое тело испытывало какое-либо внешнее воздействие, не исходящее от других тел, в частности чтобы оно имело ускорение, не вызванное воздействием со стороны других тел;

3) все силы носят характер взаимодействий. Это означает, что в механике Ньютона не допускается существование о д н о с т о р о н н и х сил, когда одно тело прилагает к другому телу некоторую силу, а обратное воздействие отсутствует.

ЗАКОНЫ МЕХАНИКИ В ПОКОЯЩИХСЯ И ДВИЖУЩИХСЯ СИСТЕМАХ ОТСЧЕТА

Допустим, что имеются две системы отсчета: S_1 и S_2, с которыми связаны координатные оси X, Y, Z и X', Y', Z', причем они некоторым образом движутся друг относительно друга. Мы можем любую из них полагать покоящейся, а другую — движущейся. Изучая механические явления, мы измеряем скорости, ускорения и силы относительно той системы отсчета, которую полагаем «покоящейся»; относительно другой системы отсчета эти величины могут иметь другие значения. В связи с этим возникает вопрос: сохраняются ли законы механики и перечисленные выше предположения по отношению к каждой из систем отсчета или, другими словами, применима ли классическая механика для любой системы отсчета, которую мы произвольно полагаем «покоящейся»?

Рассмотрим простой пример. Допустим, что на некоторое тело никакие другие тела (силы) не действуют и оно движется относительно S_1 прямолинейно и равномерно. Таким образом, относительно S_1 соблюдается закон инерции. Однако относительно S_2 движение рассматриваемого тела может быть криволинейным и неравномерным в зависимости от того, как движется система отсчета S_2 относительно S_1. Если, например, S_2 вращается вокруг S_1, то движущееся по инерции тело будет описывать относительно S_2 некоторую спиральную кривую. Следовательно, в системе отсчета S_2 закон инерции не будет соблюдаться. В связи с этим возникает проблема: как следует объяснить ускоренное движение тела относительно S_2, если известно, что это тело изолировано от воздействия других тел и поэтому должно двигаться прямолинейно и равномерно? Если исходить из указанного выше первого предположения, согласно которому «всякое ускорение создается силой», то придется допустить, что в системе отсчета S_2 на тело действует некоторая сила f, сообщающая ему тангенциальное и нормальное ускорения. Однако при этом будет нарушено второе предположение: сила f не может рассматриваться как воздействие со стороны других тел, так как по условию рассматриваемое тело изолировано и на него не действуют другие тела.

Допустим теперь, что на данное тело, имеющее массу m, действует сила F_1, сообщающая относительно S_1 ускорение $a_1 = \mathrm{d}v_1/\mathrm{d}t$:

$$F_1 = m\frac{\mathrm{d}v_1}{\mathrm{d}t}.$$

Сила F_1 обусловлена взаимодействием между данным телом и другими телами. Очевидно, что это взаимодействие не должно зависеть от того, какая система отсчета произвольно выбирается нами в качестве «покоящейся», поэтому при переходе от системы отсчета S_1 к системе отсчета S_2 величина и направление этой силы должны сохраниться. Кроме того, масса тела в механике Ньютона также рассматривается как величина, не зависящая от выбора системы отсчета. Однако если S_2 движется относительно S_1 со скоростью u, то рассматриваемое тело будет иметь относительно S_2 скорость $v_2 = v_1 - u$ и тогда, согласно второму закону механики,

$$F_2 = m\frac{\mathrm{d}v_2}{\mathrm{d}t} = m\frac{\mathrm{d}v_1}{\mathrm{d}t} - m\frac{\mathrm{d}u}{\mathrm{d}t}.$$

Таким образом, либо мы должны полагать, что $F_2 \neq F_1$, т. е. сила взаимодействия между телами различна для различных систем отсчета, либо же полагать, что $F_2 = F_1$, но в системе S_2 появляются новые силы $f = m\mathrm{d}u/\mathrm{d}t$, которые не являются воздействием на данное тело со стороны других тел. Эти силы необходимы для объяснения различий в форме траектории и характере движения тел относительно S_1 и S_2. Оба эти предположения несовместимы с основными положениями классической механики, указанными выше. Заметим также, что силы f (иногда называемые *силами инерции*) появляются при переходе от системы отсчета S_1 к системе отсчета S_2 и исчезают при обратном переходе, тогда как обычные силы взаимодействия F сохраняются при та-

ких переходах. Очевидно, что если бы система отсчета S_2 двигалась относительно S_1 прямолинейно и равномерно, то $du/dt = 0$ и как первый, так и второй законы механики соблюдались бы для обеих систем отчета. Этот вывод содержится в особом утверждении, называемом принципом относительности (Галилея):

законы механических явлений имеют одинаковый вид по отношению ко всем системам отсчета, которые движутся друг относительно друга прямолинейно и равномерно.

Системы отсчета, относительно которых соблюдаются законы механики Ньютона, называются *инерциальными*. Инерциальной системой отсчета является (с большой степенью точности) система, связанная с так называемыми «неподвижными» звездами. Система отсчета, связанная с Землей, не является инерциальной, главным образом вследствие вращения Земли вокруг своей оси; однако для большинства технических расчетов сила $f' = m\,du/dt$ очень мала и поэтому систему отсчета, в которой Земля условно принимается за покоящееся тело, полагают инерциальной. Неинерциальность этой системы замечается во всех тех явлениях, на которых сказывается вращение Земли вокруг своей оси (поворот плоскости колебаний маятника Фуко, размывание правого берега у рек, текущих в северном полушарии с севера на юг, и др.).

ПРЕОБРАЗОВАНИЯ КООРДИНАТ, СКОРОСТЕЙ И УСКОРЕНИЙ

Найдем связь между результатами измерений координат и времени, произведенных относительно двух инерциальных систем отсчета: S_1 и S_2. Обозначим через x, y, z, t координаты движущегося тела и время, измеренные относительно одной инерциальной системы отсчета S_1, а x', y', z', t' — координаты того же тела и время, измеренные относительно другой инерциальной системы отсчета S_2. Для простоты допустим, что оси OX и OX' этих систем отсчета совпадают. Тогда координаты y и z движущегося тела будут иметь относительно S_1 и S_2 одинаковые значения ($y = y'$; $z = z'$). Далее допустим, что начало координат O' системы S_2 в начальный момент времени ($t = 0$) находилось в точке O системы S_1 и с течением времени перемещается со скоростью v в положительном направлении оси OX. Так как S_1 и S_2 являются инерциальными системами отсчета, то, согласно принципу относительности, скорость v должна быть постоянной.

За время t расстояние между O и O' сделается равным vt. На такую же величину будут отличаться координаты x и x' движущегося тела, т. е. $x' = x - vt$. В механике Ньютона предполагается, что результат измерения времени не зависит от системы отсчета, поэтому $t' = t$. Следовательно, соотношения между координатами движущейся точки и временем в S_1 и S_2 имеют вид:

$$x' = x - vt; \quad y' = y; \quad z' = z; \quad t' = t. \tag{1.50}$$

Эти соотношения называются *преобразованиями Галилея*.

Обозначим теперь скорости тела относительно S_1 и S_2 через $dx/dt = w$, $dx'/dt = u$. Дифференцируя (1.50) по времени, получим закон сложения скоростей в механике Ньютона

$$w = u + v$$

(в теории относительности Эйнштейна закон сложения скоростей имеет иной вид). Иногда, полагая S_1 «покоящейся» системой, называют w — *абсолютной*, u — *относительной* и v — *переносной* скоростью. Вторичное дифференцирование по времени (полагая v переменной величиной) дает закон сложения ускорений

$$a_{абс} = a_{отн} + a_{пер}.$$

Рассмотрим теперь две другие системы отсчета, из которых одна — S_0 — является инерциальной, а вторая — S — вращается относительно S_0 с постоянной угловой скоростью ω вокруг общей оси Z.

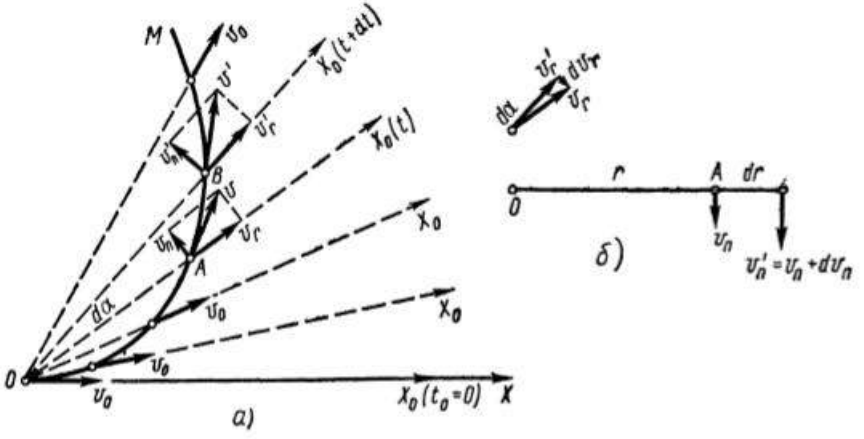


Рис. I.16

На рис. I.16, *а* оси OZ_0 и OZ обеих систем спроецированы в точку O, оси OY_0 и OY не показаны, а пунктирные линии обозначают расположение оси OX_0 относительно оси OX в различные моменты времени: $t_0 = 0$; t; $t + dt$ и т. д. Допустим, что изолированное тело движется вдоль оси OX_0 с постоянной скоростью v_0. Описывая это движение относительно S (т. е. полагая, что S является «покоящейся» системой отсчета), мы заметим, что относительно S тело движется с переменной скоростью v, описывая спиральную линию, небольшая часть которой OM изображена на рис. I.16, *а*. В каждый рассматриваемый момент времени (например, t и $t' = t + dt$) эта скорость состоит из радиальной составляющей v_r, численно всегда равной v_0, и нормальной составляющей v_n, изменяющейся по мере удаления тела от точки O. Таким образом, относительно S тело будет двигаться с ускорением.

За малое время dt тело перейдет из точки A в точку B и составляющие его скорости v в радиальном (v_r) и нормальном (v_n) направлениях изменятся на dv_r и dv_n. Радиальная скорость v_r в системе S (равная

скорости тела v_0 в системе S_0) по величине будет оставаться постоянной, но по направлению повернется на угол $d\alpha$ (рис. I.16, *б*). Вектор приращения радиальной скорости dv_r при малых углах $d\alpha$ можно считать перпендикулярным вектору v; численное значение dv_r будет равно $v_r d\alpha$. Нормальная же скорость v_n, равная ωr, будет изменяться вследствие удаления тела от точки O, поэтому

$$dv_n = d(\omega r) = \omega\, dr = \omega v_r\, dt.$$

По направлению $d\mathbf{v}_n$ будет также перпендикулярно радиус-вектору OA (рис. I.16, *б*). Таким образом, тело, которое в инерциальной системе отсчета S_0 движется прямолинейно и равномерно со скоростью $\mathbf{v}_0$, во вращающейся системе S будет иметь ускорение, равное по величине

$$\grave{a} = \frac{dv_r}{dt} + \frac{dv_n}{dt} = 2\omega v_r$$

и перпендикулярное направлению $\mathbf{v}$. Это ускорение называется *кориолисовым*; в векторной записи

$$\mathbf{a}_{кор} = 2[\boldsymbol{\omega}\mathbf{v}_r]. \qquad (1.51)$$

В общем случае, если система S имеет относительно S_0 и поступательное и вращательное движения, связь между ускорением тела относительно S_0 ($a_{абс}$) и $S(a_{отн})$ будет иметь вид

$$\mathbf{a}_{абс} = \mathbf{a}_{отн} + \mathbf{a}_{пер} + \mathbf{a}_{кор}. \qquad (1.52)$$

Произведение кориолисова ускорения на массу движущегося тела

$$m\mathbf{a}_{кор} = 2m[\boldsymbol{\omega}\mathbf{v}_r] = \mathbf{F}_{кор} \qquad (1.53)$$

называется *силой Кориолиса*. Эта сила относится к *силам инерции*; она появляется при переходе от инерциальных систем отсчета S_0 к вращающимся относительно них системам отсчета S и исчезает при обратном переходе.

Выше было предположено, что угол между векторами скорости тела ($\mathbf{v}_0$) и угловой скорости вращения системы $S(\boldsymbol{\omega})$ равен $\pi/2$. Однако полученные векторные выражения для $\mathbf{a}$ и $\mathbf{F}_{кор}$ применимы и в более общем случае, когда между векторами $\mathbf{v}_0$ и $\boldsymbol{\omega}$ существует любой угол φ. Ускорение и сила Кориолиса пропорциональны радиальной скорости движения тела $\mathbf{v}$, угловой скорости вращения системы отсчета S относительно инерциальной системы S_0 и зависят от угла $\boldsymbol{\varphi}$ между направлениями векторов $\mathbf{v}_r$ и $\boldsymbol{\omega}$

$$a_{кор} = 2\omega v_r \sin\varphi; \quad F_{кор} = 2m\omega v_r \sin\varphi.$$

При $\varphi = 0$ и $\varphi = \pi/2$ эти величины достигают наименьшего (нулевого) и наибольшего значения. Заметим также, что сила Кориолиса не совершает работы, так как она направлена перпендикулярно скорости движения тела.

Действием силы Кориолиса объясняется ряд эффектов, наблюдающихся на поверхности Земли, например поворот плоскости колебаний маятника Фуко относительно Земли. В инерциальной системе отсчета,

связанной с неподвижными звездами, плоскость колебаний маятника остается неизменной, но относительно поверхности Земли ввиду ее вращения относительно своей оси ориентация этой плоскости с течением времени регулярно изменяется. Полагая Землю «покоящейся» системой отсчета и исходя из требования, что «всякое ускорение вызывается силой», мы будем вынуждены для объяснения этого эффекта допустить, что на маятник действует особая сила, направленная перпендикулярно скорости движения тела и равная по величине $F_{кор}$.

Глава 2

ВРАЩЕНИЕ ТВЕРДОГО ТЕЛА

§ 6. КИНЕТИЧЕСКАЯ ЭНЕРГИЯ ВРАЩАЮЩЕГОСЯ ТЕЛА; МОМЕНТ ИНЕРЦИИ ТЕЛ

При вращении твердого тела относительно неподвижной оси отдельные его элементарные («точечные») части массами m_i описывают окружности различных радиусов r_i и имеют различные линейные скорости v_i. Однако угловая скорость вращения ω всех этих точек (если тело при вращении не деформируется) одинакова, т. е.

$$\omega = \frac{v_1}{r_1} = \frac{v_2}{r_2} = \ldots \tag{2.1}$$

Кинетическую энергию вращающегося тела $W_к$ подсчитаем как сумму кинетических энергий его составных элементарных частей:

$$W_к = \frac{m_1 v_1^2}{2} + \frac{m_2 v_2^2}{2} + \ldots = \frac{m_1 r_1^2 \omega^2}{2} + \frac{m_2 r_2^2 \omega^2}{2} + \ldots =$$
$$= \frac{\omega^2}{2}(m_1 r_1^2 + m_2 r_2^2 + \ldots).$$

Сумма произведений масс элементарных («точечных») частей тела на квадраты их расстояний до определенной оси называется моментом инерции тела относительно этой оси:

$$J = m_1 r_1^2 + m_2 r_2^2 + \ldots = \sum m_i r_i^2 \tag{2.2}$$

(выражается в кг·м² или г·см²). Более точно

$$J = \int_V r^2 \, dm, \tag{2.3}$$

где интегрирование производится по всему объему тела. Тогда выражение для кинетической энергии вращающегося твердого тела принимает простой вид:

$$W_к = \frac{J\omega^2}{2}. \tag{2.4}$$

Момент инерции тела зависит от распределения массы рассматриваемого тела относительно заданной оси (от формы и размеров тела

и от расположения оси, относительно которой определяется момент инерции).

Для тел, имеющих сложную форму, момент инерции можно определить путем измерений. Например, если измерить энергию $W_к$, необходимую для сообщения телу определенной угловой скорости ω относительно интересующей нас оси, то по формуле (2.4) можно вычислить момент инерции тела относительно этой оси.

Для практических применений важное значение имеет момент инерции тела относительно оси, проходящей через особую точку — *центр масс этого тела*. Выберем некоторую прямоугольную координатную систему и обозначим через x, y, z координаты элементарных масс

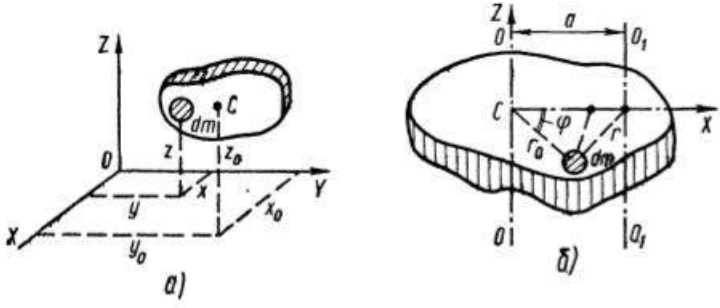


Рис. I.17

dm этого тела. По определению, центром масс данного тела называют точку, координаты которой равны (рис. I.17, *а*):

$$x_0 = \frac{1}{m}\int x\,dm; \quad y_0 = \frac{1}{m}\int y\,dm; \quad z_0 = \frac{1}{m}\int z\,dm, \qquad (2.5)$$

где m — масса всего тела, а интегрирование производится по всему объему тела. Произведения $x\,dm$, $y\,dm$, $z\,dm$ называют *статическими моментами* массы dm относительно координатных осей. В формулах (2.5) сумма этих моментов для каждой оси приравнивается статическому моменту всей массы тела, если полагать ее сосредоточенной в одной точке — в центре масс. Если начало координат O совместить с центром масс тела C, то $x_0 = y_0 = z_0 = 0$.

Прежде всего покажем, что *центр тяжести* тела совпадает с центром масс. По определению, центр тяжести данного тела есть точка приложения равнодействующей всех сил тяжести, приложенных к отдельным частям этого тела. Следовательно, относительно каждой из координатных осей момент силы тяжести P всего тела следует приравнять сумме моментов сил тяжести ΔP_i отдельных частей тела, т. е.

$$Px_0 = \sum \Delta P_i x_i; \quad Py_0 = \sum \Delta P_i y_i; \quad Pz_0 = \sum \Delta P_i z_i.$$

Подставим значения сил $P = mg$ и $\Delta P_i = \Delta m_i g$; сократив на g и заменив суммы на интегралы, получим соотношения (2.5).

Допустим, что нам известен момент инерции J_0 данного тела относительно оси OO, проходящей через центр масс C этого тела. Найдем

момент инерции J этого же тела относительно оси O_1O_1, параллельной OO и отстоящей от нее на расстоянии a. Как показано на рис. I.17, *б*, расстояние r элементарной массы dm до оси O_1O_1 можно выразить через расстояние r_0 до оси OO:

$$r^2 = r_0^2 + a^2 - 2r_0 a \cos\varphi.$$

Подставив в (2.3), получим

$$J = \int r^2\,dm = \int r_0^2\,dm + \int a^2\,dm - 2\int r_0 a\cos\varphi\,dm. \qquad (2.6)$$

Ориентируя координатную ось OZ вдоль оси OO и расположив ось OX в плоскости, проходящей через OO и O_1O_1, можем написать:

$$\int r_0\cos\varphi\,dm = \int x\,dm = x_0 m.$$

Так как мы предположили, что центр масс тела лежит на оси OO, то $x_0 = 0$ и, следовательно, последний интеграл в формуле (2.6) будет равен нулю. Таким образом, из (2.6) получается важный результат:

$$J = J_0 + ma^2, \qquad (2.7)$$

т. е. *момент инерции тела относительно любой оси равен моменту инерции относительно оси, параллельной заданной и проходящей через центр масс этого тела, плюс произведение массы тела на квадрат расстояния между осями* (теорема Штейнера).

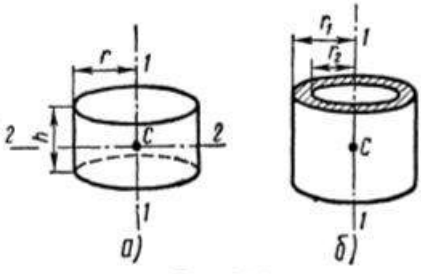


Рис. I.18

Приведем несколько формул для вычисления моментов инерции тел правильной формы относительно осей симметрии (проходящих через центр масс):

1) момент инерции точечного тела с массой m, находящегося на расстоянии r от заданной оси,

$$J = mr^2;$$

2) момент инерции однородного сплошного шара относительно оси, проходящей через центр шара,

$$J = \frac{2}{5}mR^2$$

(R — радиус шара);

3) моменты инерции однородного сплошного цилиндра относительно двух осей симметрии $1-1$ и $2-2$ (рис. I.18, *а*) соответственно равны:

$$J_1 = \frac{1}{2}mr^2; \quad J_2 = m\frac{3r^2+h^2}{12}.$$

Если $r \gg h$ (тонкий диск), то $J_2 \approx mr^2/4$; если же $r \ll h$ (тонкий стержень), то $J_2 \approx mh^2/12$;

4) момент инерции однородного полого цилиндра относительно оси I — I можно рассчитать как разность моментов инерции наружного и внутреннего (отсутствующего) цилиндров (рис. I.18, *б*):

$$J_1 = \frac{1}{2} m_1 r_1^2 - \frac{1}{2} m_2 r_2^2 = \frac{1}{2} m (r_1^2 + r_2^2),$$

где $m = m_1 - m_2$ есть масса полого цилиндра. Если цилиндр очень тонкий, то $r_1 \approx r_2 = r$ и тогда $J_1 = mr^2$.

§ 7. ЗАКОНЫ МЕХАНИКИ ДЛЯ ВРАЩАТЕЛЬНОГО ДВИЖЕНИЯ

Допустим, что твердое тело A (рис. I.19, *а*) может вращаться вокруг некоторой неподвижной оси. Для того чтобы вызвать вращение тела (изменить его угловую скорость), необходимо внешнее воздействие. Однако сила F', направление которой проходит через ось вращения,

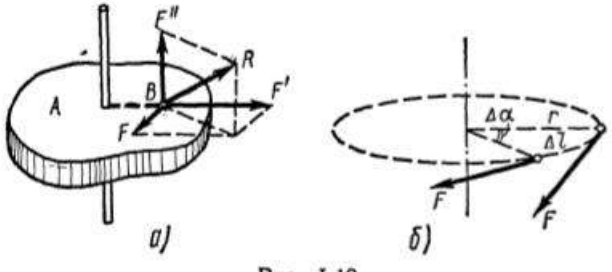


Рис. I.19

или сила F'', параллельная оси, не могут изменить угловую скорость тел. Поэтому из приложенной к телу внешней силы R необходимо выделить составляющие F' и F'', не вызывающие вращения. Вращение может быть вызвано только силой F (*вращающей силой*), лежащей в плоскости, перпендикулярной оси вращения и направленной по касательной к окружности, которую описывает точка ее приложения.

Заметим, что при вращении тела составляющие F' и F'' работы не совершают, так как точка приложения этих сил (B) перемещается перпендикулярно их направлениям. Работу совершает только вращающая сила F; она является проекцией действующей на тело силы R на направление движения точки приложения этой силы.

Определим величину работы ΔA, которую совершает вращающая сила, если точка приложения ее смещается по окружности радиуса r на Δl (рис. I.19, *б*). Предположим, что величина силы F при этом остается постоянной. Тогда

$$\Delta A = F \Delta l = Fr \Delta\alpha. \qquad (2.8)$$

Произведение вращающей силы F на радиус r есть момент вращающей силы, или *вращающий момент*, действующий на данное тело, и обозначается через M (напомним, что моментом данной силы относительно какой-нибудь оси называется произведение этой силы на ее плечо, т. е. на длину перпендикуляра, проведенного от указанной

оси до направления действия силы). Таким образом, в формуле (2.8)

$$F\dot{r} = M,$$

следовательно, работа, совершаемая вращающим моментом, равна произведению этого момента на угол поворота тела:

$$\Delta A = M\,\Delta\alpha. \tag{2.9}$$

Если вращающий момент (сила F или ее плечо r) с течением времени изменяется, то совершаемая работа определяется как сумма:

$$A = \sum M_i \Delta\alpha_i; \quad A = \int_0^{\alpha} M(\alpha)\,\mathrm{d}\alpha. \tag{2.10}$$

Момент вращающей силы представляется в виде вектора, совпадающего с осью вращения; положительную ориентировку этого вектора выбирают в том направлении, в котором перемещался бы правый винт, вращаемый этим моментом.

Вращающий момент M, приложенный к телу, сообщает ему некоторое угловое ускорение ε; согласно выбранным нами направлениям векторов $\mathbf{M}$ и $\boldsymbol{\varepsilon}$, они ориентированы по оси вращения в одну и ту же сторону. Связь между величиной вращающего момента и величиной сообщаемого им углового ускорения можно установить двумя способами:

а) можно воспользоваться тем, что работа движущей силы равна изменению кинетической энергии тела, к которому эта сила приложена: $\mathrm{d}A = \mathrm{d}W_к$. Для вращающегося тела, согласно формулам (2.9) и (2.4), имеем

$$M\,\mathrm{d}\alpha = \mathrm{d}\left(\frac{J\omega^2}{2}\right) = J\omega\,\mathrm{d}\omega. \tag{2.11}$$

Здесь мы предполагаем, что момент инерции тела при вращении не изменяется. Разделив это уравнение на $\mathrm{d}t$ и сократив на $\omega = \mathrm{d}\alpha/\mathrm{d}t$, получаем

$$M = J\frac{\mathrm{d}\omega}{\mathrm{d}t} = J\varepsilon; \tag{2.12}$$

б) можно воспользоваться тем, что момент вращающей силы M равен сумме моментов сил, которые сообщают отдельным составным частям тела m_1, m_2, ... тангенциальные ускорения a_1, a_2, ...; эти силы равны $m_i a_i$, а их моменты — $m_i a_i r_i$:

$$M = m_1 a_1 r_1 + m_2 a_2 r_2 + \ldots$$

Заменим тангенциальные ускорения на угловое ускорение, которое одинаково для всех частиц вращающегося тела (если тело при вращении не деформируется): $a_1 = r_1\varepsilon$; $a_2 = r_2\varepsilon$; ... Тогда

$$M = \varepsilon\,(m_1 r_1^2 + m_2 r_2^2 + \ldots) = J\varepsilon.$$

Формула (2.12) выражает основной закон динамики вращательного движения твердых (недеформирующихся) тел, для которых $J = \mathrm{const}$:

угловое ускорение, приобретаемое телом под действием данного вращающего момента M, прямо пропорционально величине этого момента и обратно пропорционально моменту инерции тела **J** *относительно оси вращения:*

$$\varepsilon = \frac{M}{J}; \quad M = J\varepsilon. \tag{2.13}$$

В векторной форме этот закон записывается в виде

$$\mathbf{M} = J\boldsymbol{\varepsilon}. \tag{2.14}$$

Если тело при вращении деформируется, то момент инерции его относительно оси вращения будет изменяться. Мысленно представим вращающееся тело состоящим из множества элементарных (точечных) частей; тогда деформация всего тела будет означать изменение расстояний r_i от этих частей тела до оси вращения. Однако изменение расстояния r_i при данной угловой скорости вращения ω будет сопровождаться изменением линейной скорости движения этой частицы $v_i = \omega r_i$, следовательно, и ее кинетической энергии. Таким образом, при постоянной угловой скорости вращения тела изменение расстояний r_i (следовательно, изменение момента инерции тела) будет сопровождаться изменением кинетической энергии вращения всего тела.

Из формулы (2.4), если полагать J переменным, можно получить

$$\mathrm{d}W_к = \mathrm{d}\,(J\omega^2/2) = J\omega\,\mathrm{d}\omega + \omega^2\,\mathrm{d}J/2. \tag{2.15}$$

Первое слагаемое ($J\omega\mathrm{d}\omega$) показывает изменение кинетической энергии вращающегося тела, которое произошло только вследствие изменения угловой скорости вращения (при данном моменте инерции тела), а второе слагаемое ($\omega^2\mathrm{d}J/2$) показывает изменение кинетической энергии, которое произошло только вследствие изменения момента инерции тела (при данной угловой скорости вращения).

Однако при изменении расстояния r_i от точечного тела до оси вращения внутренние силы, связывающие это тело с осью вращения, будут совершать работу: отрицательную, если тело удаляется, и положительную, если тело приближается к оси вращения; эта работа может быть рассчитана, если полагать, что сила, связывающая частицу m_i с осью вращения, численно равна центростремительной силе:

$$\mathrm{d}A_i = m_i\omega^2 r_i\,\mathrm{d}r_i = \frac{1}{2}\omega^2\,\mathrm{d}\,(m_i r_i^2).$$

Для всего тела, состоящего из множества частиц с массами m_i, получим

$$\mathrm{d}A = \frac{1}{2}\omega^2\mathrm{d}\,(m_1 r_1^2 + m_2 r_2^2 + \ldots) = \frac{1}{2}\omega^2\,\mathrm{d}J.$$

В общем случае, когда на тело действует внешний вращающий момент M, изменение кинетической энергии $\mathrm{d}W_к$ должно быть приравнено сумме двух работ: внешнего вращающего момента $M\mathrm{d}\alpha$ и внутренних сил $\mathrm{d}A = \omega^2\mathrm{d}J/2$. При ускоренном вращении величины $\mathrm{d}E_к$ и $M\mathrm{d}\alpha$ будут иметь положительные знаки, а $\mathrm{d}A$ — отрицательный

знак (так как частицы тела удаляются от оси вращения); тогда

$$dW_к = M\,d\alpha - \omega^2 \frac{dJ}{2}.$$

Подставив сюда значение $dW_к$ из выражения (2.15) и заменив $d\alpha$ на ωdt, получим

$$M\omega\,dt = J\omega\,d\omega + \omega^2\,dJ$$

или после сокращения

$$M = \frac{d}{dt}(J\omega); \quad \mathbf{M} = \frac{d}{dt}(J\boldsymbol{\omega}). \tag{2.16}$$

Это есть *общий вид основного закона механики для тел, вращающихся относительно неподвижной оси*; он применим и для деформирующихся тел. При $J = \text{const}$ формула (2.16) переходит в формулу (2.14).

Заметим, что у деформирующихся тел изменение угловой скорости вращения возможно и при отсутствии внешнего вращающего момента. Действительно, при $M = 0$ из формулы (2.16) получаем:

$$J\omega = \text{const}; \quad J_1\omega_1 = J_2\omega_2. \tag{2.17}$$

В этом случае угловая скорость вращения ω изменяется только вследствие изменения момента инерции тела, вызванного внутренними силами.



Рис. 1.20

Формула (2.17) наглядно иллюстрируется при помощи скамьи Н. Е. Жуковского (рис. 1.20) — круглой платформы, которая может с малым трением вращаться относительно вертикальной оси. Если грузы приближаются к оси вращения, то момент инерции системы уменьшается, а угловая скорость вращения увеличивается; при удалении грузов от оси вращения угловая скорость вращения уменьшается ($J_2 > J_1$; $\omega_2 < \omega_1$).

Получим два следствия из формулы (2.13), аналогичные выражениям (1.31) и (1.33). Допустим, что $M = \text{const}$ и $J = \text{const}$, тогда и $\varepsilon = \text{const}$, т. е. вращение будет равноускоренным; для некоторого промежутка времени t можно написать:

$$\varepsilon = \frac{\omega - \omega_0}{t}; \quad t = \frac{\alpha}{\omega_{ср}}; \quad \omega_{ср} = \frac{\omega + \omega_0}{2}.$$

Подставив эти величины в формулу (2.13), получим:

$$Mt = J\omega - J\omega_0;$$

$$M\alpha = \frac{J\omega^2}{2} - \frac{J\omega_0^2}{2}. \tag{2.18}$$

Основной закон динамики вращательного движения (см. формулу (2.13)) аналогичен второму закону Ньютона. Можно указать также и

на закон, аналогичный третьему закону Ньютона: если одно тело действует на другое тело с некоторым вращающим моментом $\mathbf{M}_{12}$, то второе тело всегда оказывает обратное воздействие на первое с вращающим моментом $\mathbf{M}_{21}$, равным и противоположно направленным $\mathbf{M}_{12}$:

$$\mathbf{M}_{12} = -\mathbf{M}_{21}. \tag{2.19}$$

Произведение момента инерции вращающегося тела на угловую скорость вращения $J\omega$ называется *моментом импульса* или моментом количества движения этого тела.

Так как J — скаляр, а $\boldsymbol{\omega}$ — вектор, то момент импульса $J\boldsymbol{\omega}$ есть векторная величина, ориентированная по направлению вектора угловой скорости $\boldsymbol{\omega}$.

Для точечной массы m, вращающейся по окружности радиуса r с линейной скоростью $v = \omega r$, момент импульса равен

$$J\omega = mr^2 (v/r) = mvr \tag{2.20}$$

и направлен по оси вращения (по правилу «правого винта»).

Момент силы выражается в ньютон-метрах (Н·м), момент импульса (количества движения) — в килограмм-метрах в квадрате на секунду (кг·м²/с); в системе СГС единицами этих величин являются дин·см и г·см²/с.

§ 8. СВОБОДНЫЕ ОСИ. БАЛАНСИРОВКА РОТОРОВ. ГИРОСКОПЫ

Рассмотрим поведение тела, которое может изменять свое расположение относительно оси вращения. В простейшем случае допустим, что система из двух тел A и B, соединенных стержнем (массой которого будем пока пренебрегать), приводится

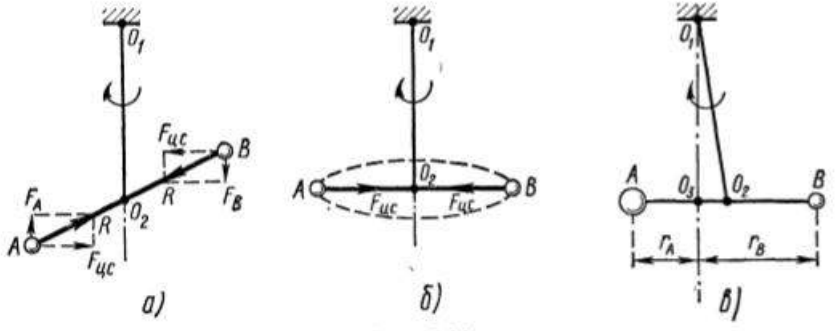


Рис. I.21

во вращение при помощи скручиваемой струны или резиновой нити O_1O_2 (рис. I.21, *а*). Для того чтобы тела A и B описывали окружности, к ним должны быть приложены центростремительные силы $F_{цс}$. Однако эти силы могут быть вызваны только в результате взаимодействия между телами и стержнем. Обозначим через R равные по величине силы, приложенные к телам со стороны деформированного (растянутого) стержня. Разложим эти силы на составляющие: силы $F_{цс}$, перпендикулярные оси вращения и сообщающие телам центростремительные ускорения, и силы F_A и F_B, параллельные оси вращения и вызывающие поворот системы вокруг точки O_2. Силы F_A и F_B будут существовать до тех пор, пока система не примет положения, показанного на рис. I.21, *б*; в этом состоянии $R = F_{цс}$ и поворачивающий момент отсутствует. Если

путем кратковременного внешнего воздействия, например удара по одному из тел, вывести систему из этого состояния, то она (уже без внешнего воздействия) снова вернется в это состояние. Таким образом, состояние вращающейся системы, показанное на рис. I.21, *б*, оказывается *устойчивым*. Заметим, что в этом состоянии векторная сумма центростремительных сил равна нулю и они не имеют момента; кроме того, момент инерции системы относительно оси вращения имеет наибольшее значение.

Допустим теперь, что массы тел m_A и m_B не равны. При вращении эта система также переходит в устойчивое состояние, при котором поворачивающие силы F_A и F_B и их моменты отсутствуют (рис. I.21, *в*). Так как стержень действует на тела с одинаковой силой, которая в установившемся положении равна центростремительным силам, то из равенства

$$R = m_A\omega^2 r_A = m_B\omega^2 r_B$$

(ω — угловая скорость вращения) следует, что $r_A/r_B = m_B/m_A$, т. е. тела будут описывать окружности различных радиусов. Ввиду этого ось вращения будет проходить через точку O_3, и если стержень закреплен не в этой точке, а в точке O_2, то струна O_1O_2 будет описывать конус. Заметим, что точка O_3 оказывается центром масс нашей системы.

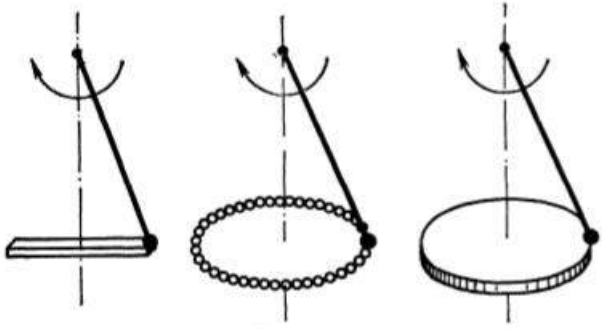


Рис. I.22

Подобное поведение наблюдается и у других вращающихся тел, имеющих возможность деформироваться и поворачиваться. На рис. I.22 показаны устойчивые положения стержня, цепочки и обруча или диска при достаточно больших угловых скоростях вращения (момент силы тяжести должен уравновешиваться моментом поворачивающих сил). Устойчивость этих состояний можно показать на примере вращающейся цепочки; если ее освободить от нити, то она падает в воздухе (или катится, подобно обручу, по поверхности демонстрационного стола), сохраняя свою форму и ориентировку оси вращения в пространстве.

Итак, во вращающейся системе существует тенденция (вызванная только внутренними силами) к переходу в устойчивое состояние, при котором:

1) центр масс системы лежит на оси вращения;

2) момент инерции тела относительно оси вращения (проходящей через центр масс) имеет наибольшее значение;

3) силы, действующие со стороны связей на каждую составную часть системы, равны по величине и направлению центростремительным силам, необходимым для вращения этих частей.

Таким образом, среди осей, проходящих через центр масс данного тела или механической системы, можно выделить оси, относительно которых вращение является устойчивым; их называют *свободными осями*. У твердых тел произвольной формы среди множества осей, проходящих через центр масс, существуют три «особые» оси: относительно одной из них момент инерции тела имеет наибольшее, относительно второй — наименьшее и относительно третьей — промежуточное значение. Эти оси взаимно перпендикулярны и называются *главными осями инерции*. Наиболее устойчивое вращение соответствует оси с наибольшим моментом инерции.

В связи с этим рассмотрим балансировку вращающихся роторов (гидравлических и паровых турбин, динамомашин, электромоторов, насосов и т. п.). В идеальном случае центр масс ротора должен лежать на оси вращения (фиксируемой подшипниками), а главная ось инерции ротора должна совпадать с осью вращения. Допустим, что вследствие технологической неточности изготовления ротора и его подшипников центр масс C ротора оказался смещенным относительно оси вращения на некоторую малую величину e (рис.1.23. $а$), но главная ось инерции параллельна оси вращения. В этом случае, для того чтобы центр масс ротора вращался с угловой скоростью ω, к нему должна быть приложена центростремительная сила $F = m\omega^2 e$, где m — масса ротора. Направление этой силы будет изменяться вслед за изменением расположения

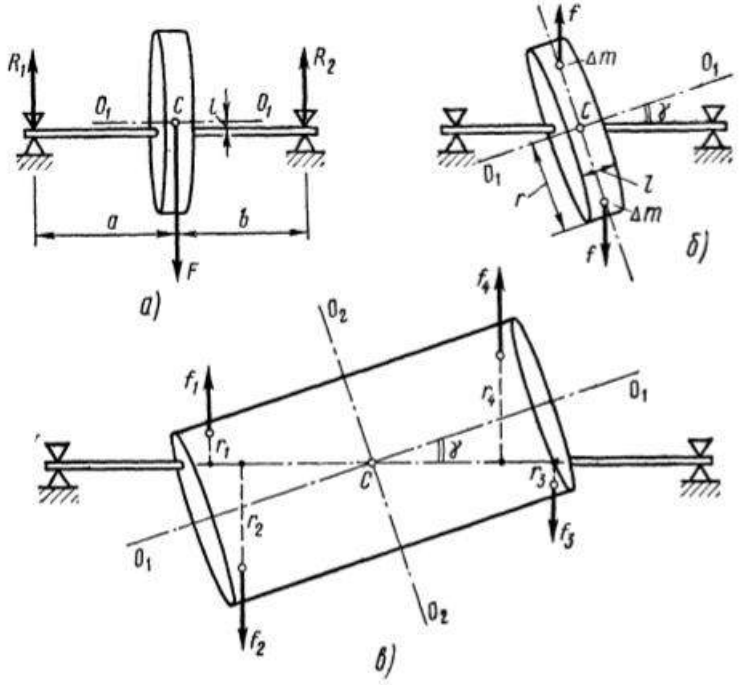


Рис. 1.23

центра масс в пространстве, т. е. вектор **F** будет вращаться в плоскости, перпендикулярной оси вращения и проходящей через центр масс. Чтобы уравновесить эту силу, на опорах возбуждаются реакции $\mathbf{R}_1$ и $\mathbf{R}_2$, равные по величине:

$$R_1 = F\frac{b}{a+b} = m\omega^2 e\frac{b}{a+b};\quad R_2 = m\omega^2 e\frac{a}{a+b}$$

(направление векторов $\mathbf{R}_1$ и $\mathbf{R}_2$ изменяется вслед за изменением направления силы **F**). При больших m, ω и e эти реакции могут быть значительными, а вызываемые ими вибрации — недопустимыми. Балансировка такого ротора (т. е. совмещение его главной оси с осью вращения) достигается либо обработкой тех частей ротора (цапф), которые входят в подшипники и определяют ось вращения, либо прикреплением к поверхности ротора некоторых «корректирующих» масс, которые вызвали бы параллельное смещение главной оси инерции ротора до ее совпадения с осью вращения.

Другой тип неуравновешенности ротора появляется, когда его главная ось инерции O_1O_1 составляет с осью вращения OO некоторый малый угол γ. На рис. 1.23, $б$ показаны две элементарные (почти точечные) массы Δm, симметрично расположенные относительно оси короткого ротора ($l \ll r$; l — длина, r — радиус ротора). Центростремительные силы, действующие на эти массы, появляются в результате взаимодей-

ствия этих масс с соседними частями ротора; в роторе возникают внутренние напряжения и деформации. Части ротора, связанные с массой Δm, прилагают к ней центростремительную силу $f_{цс} = \Delta m\omega^2 r$ (r — расстояние от этой массы до оси вращения). Согласно третьему закону механики, со стороны массы Δm на остальную часть ротора действует такая же по величине, но противоположная по направлению сила $\mathbf{f}$. Нам нужно выяснить суммарное действие на ротор всех этих элементарных сил $\mathbf{f}$. Из рис. 1.23, *б* видно, что силы, приложенные к двум симметрично расположенным массам Δm, имеют поворачивающий момент; суммируя эти моменты, мы получим общий момент, стремящийся повернуть ротор в сторону уменьшения угла γ. Этот момент должен уравновешиваться моментом сил реакций $\mathbf{R}$, возбуждаемых на подшипниках. Заметим, что в данном случае тенденция к уменьшению угла γ соответствует указанной выше тенденции к совмещению главной оси инерции (с максимальным моментом инерции) с осью вращения.

Однако если ротор «длинный» ($l \gg r$), то суммарный момент сил f будет стремиться повернуть ротор в сторону увеличения угла γ. На рис. 1.23, *в* показаны четыре симметрично расположенные элементарные массы $\Delta m_1 = \Delta m_2 = \Delta m_3 = \Delta m_4$ и силы $\mathbf{f}_1$, $\mathbf{f}_2$, $\mathbf{f}_3$ и $\mathbf{f}_4$, с которыми эти массы действуют на остальные части ротора. Массы Δm_2 и Δm_4 оказались расположенными дальше от оси вращения, чем массы Δm_1 и Δm_3,

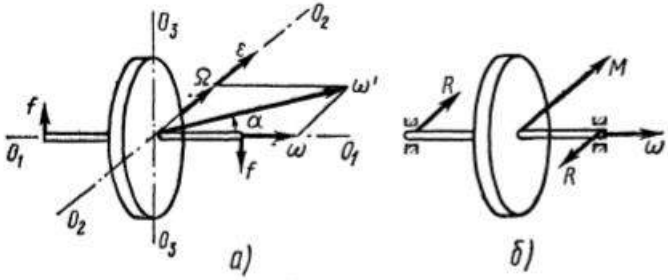


Рис. 1.24

поэтому силы $\mathbf{f}_2$ и $\mathbf{f}_4$ и их момент, стремящиеся повернуть ротор в сторону увеличения угла γ, оказываются больше сил $\mathbf{f}_1$ и $\mathbf{f}_3$ и их момента, поворачивающего ротор в противоположном направлении. Однако у длинного ротора осью с наибольшим моментом инерции является не O_1O_1, а O_2O_2, поэтому увеличение угла γ по-прежнему означает тенденцию к совмещению главной оси инерции с осью вращения.

Неуравновешенности, вызванные существованием угла γ, также можно устранить (или ослабить) добавлением к ротору «корректирующих» масс, которые создают момент, направленный против суммарного поворачивающего момента, вызванного силами $\mathbf{f}$. Этим будут ослаблены силы реакции на опорах.

Важное применение в технике получили г и р о с к о п ы — приборы, основной частью которых являются массивные тела, вращающиеся с большой угловой скоростью вокруг оси с максимальным моментом инерции. Конструкции гироскопов зависят от их назначения. Выясним их основное свойство, воспользовавшись рис. 1.24. Допустим, что в начальный момент времени $t_0 = 0$ к оси гироскопа, имеющего угловую скорость вращения ω и момент инерции J, приложена пара внешних сил $\mathbf{f}$, момент которых $\mathbf{M} = [\mathbf{fl}]$ (l — плечо пары) стремится повернуть ось вращения O_1O_1 вокруг перпендикулярной ей оси O_2O_2 (рис. 1.24, *а*). Этот внешний момент будет сообщать гироскопу угловое ускорение ε, вектор которого ориентирован по оси O_2O_2. Через некоторое время t под действием момента $\mathbf{M}$ гироскоп получит угловую скорость вращения вокруг оси O_2O_2, равную $\Omega = \varepsilon t$. Для того чтобы найти результирующую угловую скорость вращения гироскопа (по величине и направлению), необходимо сложить векторы ω и Ω. Оба эти вектора лежат в плоскости, проходящей через O_1O_1 и O_2O_2; в этой же плоскости будет лежать и вектор результирующей скорости ω'. Однако, для того чтобы к моменту t гироскоп вращался в соответствии с вектором угловой скорости ω', ось гироскопа за это время должна повернуться вокруг оси O_3O_3 на угол α.

Итак, если бы гироскоп не вращался, то внешний момент $\mathbf{M}$ привел бы его во вращение относительно O_2O_2; если же гироскоп вращается, то тот же момент $\mathbf{M}$ вызовет

поворот гироскопа вокруг оси O_3O_3. Это явление называется *гироскопическим эффектом*. Если ось гироскопа закреплена подшипниками (рис. I.24, *б*), то внешний момент **M** вызовет появление реакций опор **R**. При отсутствии вращения векторы этих сил лежали бы в плоскости, проведенной через ось O_1O_1, перпендикулярно вектору **M**; у вращающегося же гироскопа силы **R** лежат в одной плоскости с вектором **M** и осью O_1O_1. Весьма важно, что величина реакций опор **R** (называемых *гироскопическими силами*) пропорциональна угловой скорости вращения гироскопа. Эти силы появляются, например, в подшипниках паровых турбин, установленных на кораблях, при изменении курса плавания и т. п.

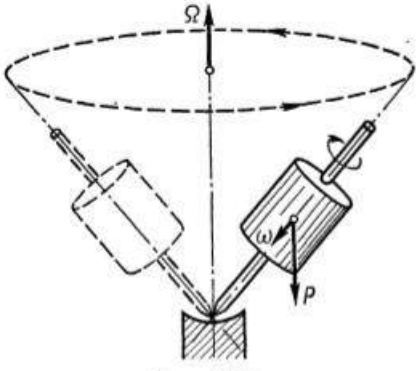


Рис. I.25

Гироскопический эффект используется для навигационных целей. Гироскоп содержит массивное тело, которому сообщено вращение с очень большим числом оборотов. Реакция этого прибора на всякое изменение направления оси вращения, соответствующим образом усиленная, может передаваться на рулевое устройство корабля или самолета.

Выше предполагалось, что при повороте гироскопа под действием внешнего момента **M** направление самого вектора **M** в пространстве не изменялось и все время оставалось ориентированным вдоль оси O_2O_2. Поэтому как только ось гироскопа повернется на угол $\alpha = \pi/2$ и будет совпадать с осью O_2O_2, векторы $\boldsymbol{\omega}$, **M** и $\boldsymbol{\varepsilon}$ будут по направлению совпадать и действие внешнего момента **M** будет заключаться только в изменении величины ω. Однако в некоторых случаях при повороте оси гироскопа вместе с ним поворачивается и направление внешнего момента **M**. Тогда вектор углового ускорения $\boldsymbol{\varepsilon}$ все время будет оставаться перпендикулярным вектору $\boldsymbol{\omega}$ угловой скорости вращения гироскопа. В этом случае величина ω останется постоянной, а его ориентация в пространстве будет изменяться с некоторой угловой скоростью Ω (при $M = \text{const}$, $\varepsilon = \text{const}$ будет постоянной также и Ω). Например, ось демонстрационного гироскопа (рис. 1.25), имеющего одну точку опоры, под действием момента силы тяжести описывает конус вокруг вертикали. Такое движение гироскопа называется *прецессией*.

Глава 3

ЗАКОНЫ СОХРАНЕНИЯ В СИСТЕМЕ ВЗАИМОДЕЙСТВУЮЩИХ ТЕЛ

§ 9. ЗАКОН СОХРАНЕНИЯ ИМПУЛЬСА (КОЛИЧЕСТВА ДВИЖЕНИЯ)

Рассмотрим систему взаимодействующих тел, в частности совокупность тел, взаимодействие и относительное движение которых определяется законами механики. Простейшими системами являются, например, два тела, притягивающиеся по закону всемирного тяготения; два тела, связанные упругой пружиной, и др.

Силы, действующие в системе, условно делят на две группы:

1) *внутренние силы* — это силы взаимодействия между составными частями системы. Согласно третьему закону механики, эти силы попарно равны и противоположно направлены, поэтому если произвести

векторное сложение всех внутренних сил, действующих между составными частями системы, то в сумме получится нуль;

2) *внешние силы* — это силы, приложенные к составным частям системы со стороны тел, не включенных в систему. Систему, на которую внешние силы не действуют, называют *изолированной*.

Рассмотрим сначала простейшую механическую систему из двух тел, изображенную на рис. 1.26, где f_{12} и f_{21} — внутренние силы взаимодействия между телами, F_1 и F_2 — внешние силы (или равнодействующие внешних сил, если их несколько). Допустим, что Δt — промежуток времени, в течение которого каждую из этих сил можно с удовлетворительным приближением считать постоянной (по величине и направлению). Тогда на основании формулы (1.31) для каждого из тел можно написать:

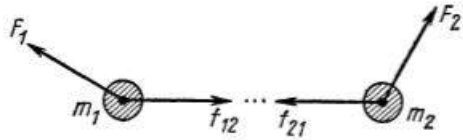


Рис. 1.26

$$\begin{aligned}(\mathbf{F}_1+\mathbf{f}_{12})\,\Delta t &= m_1\mathbf{v}_1 - m_1\mathbf{v}_{10},\\ (\mathbf{F}_2+\mathbf{f}_{21})\,\Delta t &= m_2\mathbf{v}_2 - m_2\mathbf{v}_{20},\end{aligned} \tag{3.1}$$

где $\mathbf{v}_{10}$ и $\mathbf{v}_{20}$ — начальные скорости тел, $\mathbf{v}_1$ и $\mathbf{v}_2$ — их скорости по истечении времени Δt. Правые части этих уравнений выражают изменение импульса тел за время Δt. Чтобы найти полное изменение импульса в системе за это время, произведем сложение правых и левых частей уравнений (3.1):

$$(\mathbf{F}_1+\mathbf{F}_2)\,\Delta t + (\mathbf{f}_{12}+\mathbf{f}_{21})\,\Delta t = (m_1\mathbf{v}_1+m_2\mathbf{v}_2) - (m_1\mathbf{v}_{10}+m_2\mathbf{v}_{20}). \tag{3.2}$$

Согласно третьему закону механики, векторная сумма сил взаимодействия равна нулю, т. е. $\mathbf{f}_{12} + \mathbf{f}_{21} = 0$. Если внешние силы отсутствуют или их векторная сумма (равнодействующая) равна нулю, то из выражения (3.2) получается важный результат:

$$m_1\mathbf{v}_1 + m_2\mathbf{v}_2 = m_1\mathbf{v}_{10} + m_2\mathbf{v}_{20}, \tag{3.3}$$

т. е. векторная сумма импульсов в системе остается постоянной. При этом возможно, что увеличение скорости одного тела сопровождается уменьшением скорости другого тела, но возможно, что обе скорости одновременно увеличиваются или уменьшаются; постоянным остается не сумма абсолютных значений импульсов, а сумма соответствующих векторов. Для примера рассмотрим два тела, притягивающихся друг к другу по закону тяготения; с течением времени их скорости увеличиваются, но изменение импульса первого тела в одном направлении компенсируется равным изменением импульса второго тела в противоположном направлении.

Другой пример: при выстреле под действием внутренних сил снаряд приобретает в одном направлении такой же импульс, который сообщается орудию в противоположном направлении; векторная сумма импульсов в этой системе, равная нулю до выстрела, остается равной нулю и после выстрела.

Если в системе имеется много тел, то можно рассматривать их взаимодействие между собой попарно и найти суммарное изменение импульса каждого из тел, входящих в состав системы:

$$(\mathbf{F}_1 + \mathbf{f}_1)\,\Delta t = m_1\mathbf{v}_1 - m_1\mathbf{v}_{10},$$
$$(\mathbf{F}_2 + \mathbf{f}_2)\,\Delta t = m_2\mathbf{v}_2 - m_2\mathbf{v}_{20},$$
$$\cdots\cdots\cdots\cdots\cdots$$

Здесь $\mathbf{F}_1$ — равнодействующая всех внешних сил, $\mathbf{f}_1$ — равнодействующая всех внутренних сил, действующих на первое тело, и т. д. Сложив левые и правые части этих уравнений, найдем суммарное изменение импульса в системе:

$$(\Sigma\,\mathbf{F}_i + \Sigma\mathbf{f}_i)\,\Delta t = \Sigma\, m_i\mathbf{v}_i - \Sigma\, m_i\mathbf{v}_{i0}. \qquad (3.4)$$

Векторная сумма всех внутренних сил в системе равна нулю по третьему закону механики, поэтому суммарный импульс в системе может быть изменен только под действием внешних сил. Обозначим векторную сумму всех внешних сил, действующих на систему, через $\mathbf{R} = \Sigma\,\mathbf{F}_i$; тогда

$$\mathbf{R}\,\Delta t = \Sigma\, m_i\mathbf{v}_i - \Sigma\, m_i\mathbf{v}_{i0}. \qquad (3.5)$$

Важным следствием из этого уравнения является з а к о н с о х р а н е н и я и м п у л ь с а (количества движения) в системе:

если на систему внешние силы не действуют или их векторная сумма $\mathbf{R}$ *равна нулю, то суммарный импульс в системе с течением времени не изменяется*:

$$\Sigma\, m_i\mathbf{v}_i = \Sigma\, m_i\mathbf{v}_{i0}; \quad \Sigma\, m_i\mathbf{v}_i = \text{const}. \qquad (3.6)$$

Заметим, что внутренние силы, так же как и внешние, могут изменять величины скоростей тел, составляющих систему, но при отсутствии внешних сил это изменение происходит таким образом, что векторная сумма импульсов в системе остается постоянной. Например, при разрыве шрапнельного снаряда или авиационной бомбы в воздухе сумма векторов импульсов всех осколков равна вектору импульса снаряда или бомбы до разрыва (разумеется, при условии, что время разрыва очень мало и поэтому силы тяжести не вызовут заметного изменения импульсов осколков в процессе разрыва).

Векторное уравнение (3.6), выражающее закон сохранения импульса механической системы, может быть представлено в виде совокупности трех скалярных уравнений:

$$\Sigma\, m_i v_{ix} = \text{const}_1; \quad \Sigma\, m_i v_{iy} = \text{const}_2; \quad \Sigma\, m_i v_{iz} = \text{const}_3,$$

т. е. не только сумма векторов импульсов, но и *сумма проекций этих векторов на координатные оси* (а вообще говоря, и на любое направление) *остаются постоянными*.

Рассмотрим частный случай, когда внешние силы, действующие на составные части системы, перпендикулярны некоторому направлению, например оси OX. Тогда вдоль этого направления внешние силы

не могут сообщить телам тангенциального ускорения; следовательно, проекции скоростей составных частей системы на это направление v_{1x}, v_{2x}, ... могут быть изменены только силами взаимодействия. Однако внутренние силы взаимодействия не могут изменить как общий импульс в системе, так и суммарный импульс составных частей системы в любом выделенном нами направлении. Ввиду этого вдоль указанной выше оси OX сумма импульсов составных частей системы с течением времени изменяться не будет:

$$m_1v_{1x}+m_2v_{2x}+\ldots=\text{const}.$$

В этой формуле скорости, направленные вдоль положительной ориентации оси OX, должны быть записаны со знаком плюс, а скорости, направленные в противоположную сторону, — со знаком минус.

§ 10. ЗАКОН СОХРАНЕНИЯ МОМЕНТА ИМПУЛЬСА (МОМЕНТА КОЛИЧЕСТВА ДВИЖЕНИЯ)

Рассмотрим простейшую систему из двух взаимодействующих тел, прилагающих друг к другу, согласно выражению (2.19), равные по величине и противоположно направленные вращающие моменты $\mathbf{M}_1$ и $\mathbf{M}_2$. Для каждого из этих тел можно написать уравнение:

$$\begin{aligned}\mathbf{M}_1\,\Delta t &= J_1\boldsymbol{\omega}_1 - J_1\boldsymbol{\omega}_{10},\\ \mathbf{M}_2\,\Delta t &= J_2\boldsymbol{\omega}_2 - J_2\boldsymbol{\omega}_{20}.\end{aligned}\qquad(3.7)$$

Сложив эти уравнения и учитывая, что $\mathbf{M}_1 = -\mathbf{M}_2$, получим

$$J_1\boldsymbol{\omega}_1 + J_2\boldsymbol{\omega}_2 = J_1\boldsymbol{\omega}_{10} + J_2\boldsymbol{\omega}_{20}.\qquad(3.8)$$

т. е. за время взаимодействия Δt суммарный момент импульса в системе не изменился. При этом существенно отсутствие внешних вращающих моментов. Этот результат можно получить и для любой механической системы, состоящей из многих вращающихся тел, взаимодействующих между собой с соблюдением условия (2.19); соотношения (3.7) и (3.8) будут иметь место для каждого отдельного взаимодействия любой пары составных частей системы. Таким образом,

---

если на механическую систему внешние вращающие моменты не действуют, то векторная сумма моментов импульсов составных частей системы с течением времени не изменяется:

$$\sum J_i\boldsymbol{\omega}_i = \text{const}.\qquad(3.9)$$

Другими словами, внутренние вращающие моменты, действующие между составными частями системы, не могут изменить суммарного момента импульса системы. Это утверждение называется законом сохранения момента импульса (момента количества движения).

Однако внутренние вращающие моменты могут изменять кинетическую энергию вращения системы. Рассмотрим пример: допустим,

что в начальный момент времени изолированная вращающаяся система имеет момент инерции J_1 и угловую скорость ω_1. Через некоторое время вследствие перераспределения масс относительно оси вращения под действием внутренних (центростремительных и упругих) сил момент инерции системы становится равным J_2, а угловая скорость — ω_2. Согласно закону сохранения момента импульса, $J_1\omega_1 = J_2\omega_2$, но тогда $J_1\omega_1^2/2 \neq J_2\omega_2^2/2$. Таким образом, при отсутствии внешних моментов кинетическая энергия вращающейся системы изменяется на величину, равную $J_2\omega_2^2/2 - J_1\omega_1^2/2$; очевидно, это происходит за счет работы внутренних сил. Заметим, что аналогичный результат имеет место и при поступательном движении. Например, общая кинетическая энергия системы, состоящей из винтовки и пули, до выстрела равна нулю, а после выстрела отлична от нуля, тогда как векторная сумма импульсов до и после выстрела не изменяется. Появление кинетической энергии составных частей изолированной системы объясняется *работой внутренних сил*.

Допустим, что вращающееся тело изолировано от внешних воздействий, поэтому $J\omega = \text{const}$. Если под действием внутренних сил тело деформируется и его момент инерции изменяется, то, дифференцируя это равенство по времени, получим

$$J\frac{d\omega}{dt} + \omega\frac{dJ}{dt} = 0,$$

следовательно, тело будет вращаться с угловым ускорением, величина которого

$$\frac{d\omega}{dt} = -\frac{\omega}{J}\frac{dJ}{dt}$$

пропорциональна скорости изменения момента инерции. Знак минус показывает, что при увеличении момента инерции (dJ/dt — положительное) угловая скорость будет уменьшаться ($d\omega/dt$ — отрицательное), и наоборот.

§ 11. ЗАКОН СОХРАНЕНИЯ МЕХАНИЧЕСКОЙ ЭНЕРГИИ

Вычислим суммарную работу внешних и внутренних сил; допустим, что за время Δt, в течение которого эти силы можно полагать постоянными, тело m_1 сместилось на Δs_1, тело m_2 — на Δs_2 и т. д.

На основании выражения (1.33) для каждого тела, входящего в состав системы, можно написать:

$$F_1\Delta s_1\cos\varphi_1 + f_1\Delta s_1\cos\alpha_1 = \frac{1}{2}m_1v_1^2 - \frac{1}{2}m_1v_{10}^2,$$

$$F_2\Delta s_2\cos\varphi_2 + f_2\Delta s_2\cos\alpha_2 = \frac{1}{2}m_2v_2^2 - \frac{1}{2}m_2v_{20}^2,$$

. . . . . .

где φ_i и α_i — углы между направлениями сил и смещений.

Сложив левые и правые части этих уравнений, получим

$$\sum F_i\Delta s_i\cos\varphi_i + \sum f_i\Delta s_i\cos\alpha_i = \sum\frac{1}{2}m_iv_i^2 - \sum\frac{1}{2}m_iv_{i0}^2. \quad (3.10)$$

Обозначив суммарную работу всех внешних сил $\Sigma F_i \Delta s_i \cos \varphi_i = A_{\text{внеш}}$, суммарную работу всех внутренних сил $\Sigma f_i \Delta s_i \cos \alpha_i = A_{\text{внут}}$, суммарную кинетическую энергию всех составных частей системы $\sum \left(\frac{1}{2} m_i v_i^2 \right) = W_{\text{к}}$, можем переписать уравнение (3.10) в виде

$$A_{\text{внеш}} + A_{\text{внут}} = W_{\text{к}} - W_{\text{к0}}, \tag{3.11}$$

где $W_{\text{к0}}$ — суммарная кинетическая энергия в системе в начальный момент времени.

Если внешние силы отсутствуют или их работа равна нулю, то изменение суммарной энергии в системе вызывается только работой внутренних сил, причем эта работа может быть положительной или отрицательной. Совершение работы внутренними силами сопровождается переходом системы из одного состояния в другое, т. е. изменениями в относительном расположении составных частей системы. Заметим, что внутренние силы могут совершать только ограниченное количество отрицательной или положительной работы.

Условимся называть *потенциальной энергией* системы величину $W_{\text{п}}$, зависящую от относительного расположения составных частей системы, уменьшение которой равно положительной работе внутренних сил:

$$W_{\text{п0}} - W_{\text{п}} = A_{\text{внут}}, \tag{3.12}$$

где $W_{\text{п0}}$ — начальное значение потенциальной энергии. Таким образом, согласно нашему условию, если внутренние силы совершают положительную работу, то потенциальная энергия системы уменьшается ($W_{\text{п}} < W_{\text{п0}}$); если же внутренние силы совершают отрицательную работу, то потенциальная энергия системы увеличивается ($W_{\text{п}} > W_{\text{п0}}$). С другой стороны, согласно уравнению (3.11), при $A_{\text{внеш}} = 0$ положительная работа внутренних сил сопровождается увеличением кинетической энергии в системе, т. е. $W_{\text{к}} > W_{\text{к0}}$. Следовательно,

$$A_{\text{внут}} = W_{\text{п0}} - W_{\text{п}} = W_{\text{к}} - W_{\text{к0}},$$

откуда получаем важный результат

$$W_{\text{к0}} + W_{\text{п0}} = W_{\text{к}} + W_{\text{п}}, \tag{3.13}$$

выражающий закон сохранения механической энергии:

если на систему внешние силы не действуют или их суммарная работа равна нулю, то сумма кинетической и потенциальной энергии в системе с течением времени не изменяется:

$$W_{\text{к}} + W_{\text{п}} = \text{const}.$$

Для соблюдения этого закона помимо условия $A_{\text{внеш}} = 0$ необходимо еще одно условие — чтобы внутренние силы не производили в системе иного действия, кроме сообщения телам ускорения, т. е. чтобы в системе *не происходило превращения механической энергии*

в другие виды энергии. Только в этом случае работа внутренних сил будет в точности равна изменению кинетической энергии системы. Если же такой переход имеет место, то работу внутренних сил следует приравнять изменению не только кинетической энергии, но и изменению других видов энергии:

$$A_{\text{внут}} = W_{\text{к}} - W_{\text{к0}} + W,$$

где W — изменение других видов энергии, которое произошло в системе за время Δt.

Закон сохранения механической энергии является, таким образом, частным случаем всеобщего закона сохранения энергии, согласно которому **в изолированной системе сумма всех видов энергии с течением времени не изменяется.**

Если в системе существует трение, то при относительном движении трущихся тел происходит разрушение их поверхностей и часть механической энергии системы постепенно превращается в теплоту. Упорядоченное движение составных частей системы превращается при трении в беспорядочное движение атомов и молекул.

Для измерения и расчета убыли механической энергии вследствие трения вводят понятие *силы трения.* Полагают, что при относительном движении двух трущихся тел между ними возникает некоторая сила $F_{\text{тр}}$, обусловленная характером трения; работа этой силы $A_{\text{тр}} = F_{\text{тр}}s$ (или $\Sigma F_{\text{тр}}\Delta s$) приравнивается той части механической энергии этих тел, которая превращается в теплоту. В качестве примера рассмотрим скольжение тела по наклонной плоскости под действием силы тяжести P. При этом нормальная составляющая действующей силы $N = P\cos\alpha$ (α — угол между N и P) не может изменять скорости движения тела по величине и поэтому работы не совершает. Однако эта сила, прижимая движущееся тело к плоскости скольжения, вызывает между ними силу трения $F_{\text{тр}}$, которая оказывается пропорциональной величине N:

$$F_{\text{тр}} = fN = fP\cos\alpha,$$

где f называют *коэффициентом трения* (f зависит от материала трущихся тел, состояния поверхностей трения, наличия смазки, температуры, а также от скорости относительного движения). Тангенциальная составляющая силы тяжести $F = P\sin\alpha$ совершает работу, идущую на увеличение кинетической энергии тела. Однако при наличии трения часть кинетической энергии, которую сила F сообщает телу, переходит в теплоту; поэтому при движении с трением необходимо из работы движущей силы F вычесть работу силы трения:

$$\frac{1}{2}mv^2 - \frac{1}{2}mv_0^2 = Fs - A_{\text{тр}} = (F - F_{\text{тр}})\,s.$$

В частности, если на пути s вся кинетическая энергия, полученная телом, переходит в теплоту, то $v = v_0$ и $F = F_{\text{тр}}$; можно сказать, что положительная работа движущей силы равна отрицательной работе силы трения.

Согласно выражению (3.12), *потенциальная энергия системы есть максимальная положительная работа, которую могут совершить действующие в системе внутренние силы.*

Рассчитаем потенциальную энергию сжатой или растянутой упругой пружины; внутренние силы f, действующие на концах пружины, направлены против внешних деформирующих сил и по величине пропорциональны деформации x (рис. 1.27, *а*):

$$f = kx,$$

где k — *коэффициент упругости пружины*. Вычислим работу, которую могут совершать внутренние силы при переходе пружины из деформированного состояния в нормальное (недеформированное); эта

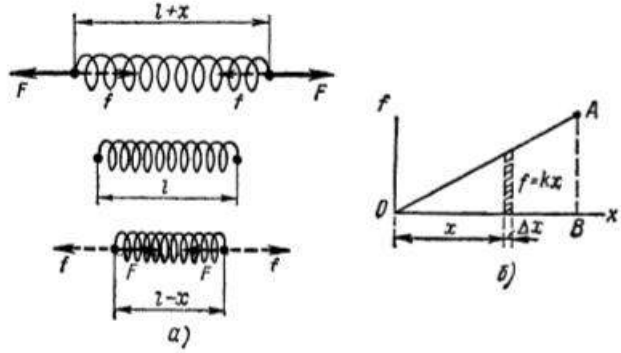


Рис. 1.27

работа всегда положительная. При изменении длины пружины на очень малую величину Δx силы f можно считать почти постоянными, а их работа будет равна $\Delta A = f\Delta x = kx\Delta x$. Графически эта работа изображается площадкой, заштрихованной на рис. 1.27, *б*. Полная работа внутренних сил при переходе в нормальное состояние представлена, очевидно, площадью треугольника OAB: $A = fx/2 = kx^2/2$. Эта работа и есть потенциальная энергия деформированной пружины

$$W_п = kx^2/2. \qquad (3.14)$$

Для скрученной спиральной пружины аналогичный расчет дает $E_п = D\alpha^2/2$, где D — *коэффициент упругости на кручение*, α — угол скручивания. Заметим, что в состоянии с нулевой потенциальной энергией внутренние силы равны нулю.

Для упругодеформированного сплошного тела (или среды) можно рассчитать потенциальную энергию, содержащуюся в единице объема. Применим формулу (3.14) к стержню длиной l и сечением S. Тогда $x/l = \varepsilon$ будет относительным удлинением тела под действием силы f, а $f/S = \sigma$ — нормальным механическим напряжением. Воспользуемся законом Гука (см. формулу (1.24)) и выразим потенциальную энергию всего стержня (см. (3.14)) в зависимости или от относительного удли-

нения ε, или от напряжения p:

$$W_{\text{п}} = \frac{E\varepsilon^2}{2} V = \frac{\sigma^2}{2E} V,$$

где $V = Sl$ — объем тела; E — модуль продольной упругости.

Таким образом, величина

$$w = E\varepsilon^2/2 = \sigma^2/2E$$

есть энергия упругой деформации (растяжения или сжатия), содержащаяся в единице объема тела или упругой среды. Для данной среды с определенным модулем продольной упругости E энергия в единице объема w прямо пропорциональна квадрату относительной деформации ε или квадрату напряжения σ. При данной относительной деформации эта энергия пропорциональна модулю продольной упругости. Аналогичные выражения получаются и для других видов упругих деформаций.

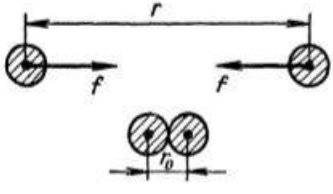


Рис. I.28

Рассчитаем потенциальную энергию двух тел с массами m_1 и m_2, притягивающихся друг к другу по закону тяготения. Силы взаимного притяжения $f = G\frac{m_1 m_2}{r^2}$ будут совершать положительную работу, если тела приближаются, и отрицательную работу, если тела удаляются. Допустим, что в начальном состоянии эти тела находятся на расстоянии r, а при сближении наименьшее возможное расстояние между ними (при соприкосновении) равно r_0 (рис. I.28). Тогда положительная работа, которая совершается силами тяготения при этом сближении, может быть подсчитана как сумма элементарных работ $\mathrm{d}A = f\mathrm{d}r$, т. е.

$$A = -\int_r^{r_0} G\frac{m_1 m_2}{r^2}\,\mathrm{d}r = G\frac{m_1 m_2}{r_0} - G\frac{m_1 m_2}{r} \qquad (3.15)$$

(знак минус перед интегралом появился вследствие того, что при уменьшении расстояния между телами величина $\mathrm{d}r$, как разность между новым и начальным значениями r, отрицательная, в то время как работа $\mathrm{d}A = f\mathrm{d}r$ положительная, так как перемещение тел происходит в направлении действия сил). В частном случае, когда расстояния r и r_0 велики, а их разность $(r - r_0)$ мала, как это имеет место при падении тела на поверхность Земли с небольшой высоты, можно произведение rr_0 заменить на r^2 и тогда

$$A \approx G\frac{m_1 m_2}{r^2}(r - r_0) = Ph.$$

Работу, которую может при своем падении совершить тело весом P, расположенное на высоте h, называют *потенциальной энергией этого тела в поле тяготения* Земли ($W_{\text{п}} = Ph = mgh$). По мере приближе-

ния тела к Земле сила тяжести совершает положительную работу и потенциальная энергия тела уменьшается.

Однако если потенциальную энергию системы из двух притягивающихся тел приравнивать работе A, рассчитанной по формуле (3.15), то потенциальная энергия системы будет равна нулю при $r = r_0$. Но наименьшее расстояние между телами r_0 не всегда есть определенная величина. Это обстоятельство побудило выбрать другое, более определенное состояние системы, при котором ее потенциальная энергия равна нулю; а именно условились считать, что *потенциальная энергия любой системы равна нулю, если ее составные части удалены друг от друга на бесконечно большие расстояния;* при этом силы взаимодействия между телами равны нулю.

Из этого условия следует, что потенциальная энергия системы *притягивающихся* тел должна быть отрицательной величиной (а системы *отталкивающихся* тел — *положительной* величиной). Действительно, если из начального состояния, когда расстояние между телами бесконечно большое и $W_п = 0$, тела начнут сближаться, например до расстояния r, то силы притяжения совершают положительную работу и поэтому потенциальная энергия должна уменьшаться и, следовательно, станет меньше нуля. Таким образом,

$$A = W_{п0} - W_п;\ W_{п0} = 0;\ W_п = -A = \int_{\infty}^{r} G\frac{m_1 m_2}{r^2}\,\mathrm{d}r.$$

Следовательно,

$$W_п = -G\frac{m_1 m_2}{r}. \qquad (3.16)$$

Для системы, состоящей из двух взаимодействующих электрических зарядов q_1 и q_2, потенциальная энергия выражается такой же формулой:

$$W_п = k_0\frac{q_1 q_2}{r},$$

где k_0 — постоянная величина. Если заряды имеют различные знаки, то потенциальная энергия получается отрицательной; для системы зарядов одинакового знака потенциальная энергия — положительная величина. На рис. I.29 приведены графики функции $W_п(r)$, показывающие изменение потенциальной энергии системы притягивающихся (*1*) и отталкивающихся (*2*) тел с изменением расстояния между ними.

Для более сложной системы, содержащей, например, N взаимодействующих тел, потенциальная энергия будет представлять собою функцию от координат всех этих тел: $W_п(x_1, y_1, z_1; x_2, y_2, z_2, \ldots)$. Как и в случае простой системы из двух тел, эта функция подбирается таким образом, чтобы при изменении координат взаимодействующих тел работа сил взаимодействия была равна $-\mathrm{d}W_п$. Для двух тел $\mathrm{d}W_п = -f\mathrm{d}r$, следовательно,

$$f = -\mathrm{d}W_п/\mathrm{d}r.$$

Для сложной системы, состоящей из многих взаимодействующих между собой тел, сила, действующая на i-е тело в направлении, напри-

мер, оси OX, равна частной производной:

$$f_{ix} = -\partial W_n / \partial x_i.$$

Допустим, что при графическом изображении функции $W_n(x_i, y_i, z_i)$ для некоторой физической системы начало координат O поме-

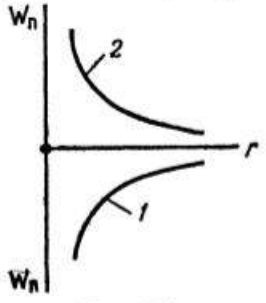


Рис. I.29

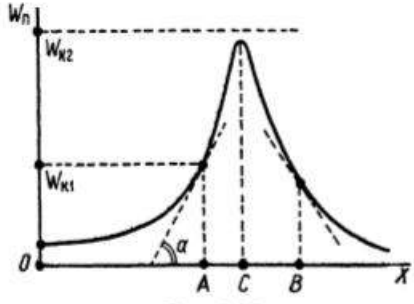


Рис. I.30

щено внутри системы, а ось OX ориентирована в каком-нибудь интересующем нас направлении (рис. I.30). Кривая $W_n(x)$ показывает, как изменяется потенциальная энергия системы, если одна частица перемещается вдоль оси OX, а остальные остаются неподвижными. Если $W_n(x)$ резко возрастает при удалении одной из частиц системы от остальных, то говорят, что система окружена *потенциальным барьером*. Заметим, что величина и направление силы, действующей на

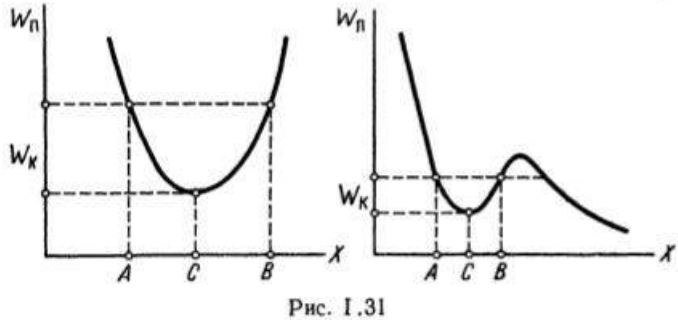


Рис. I.31

частицу, определяются величиной и знаком тангенса угла наклона касательной к кривой $W_n(x)$:

$$f = -\mathrm{d}W_n(x)/\mathrm{d}x = -k \operatorname{tg} \alpha,$$

где k — коэффициент, определяемый масштабами для W_n и x. Например, в точке A сила f имеет отрицательный знак, т. е. она направлена к точке O и будет препятствовать удалению частицы из системы; в точке B сила имеет противоположное направление.

При удалении от точки O потенциальная энергия будет возрастать, следовательно, кинетическая энергия частицы должна убывать. Если в точке O кинетическая энергия частицы была равна $W_{к1}$, то она сде-

лается равной нулю в точке A. Здесь частица остановится, после чего действующие на нее силы сообщат ей обратное движение; частица не сможет преодолеть потенциальный барьер и удалиться из системы. Это будет возможно только при достаточно большой кинетической энергии, например равной $W_{к2}$.

В сложных системах взаимодействующих тел могут образоваться также *потенциальные ямы* (рис. I.31). Если частица оказалась на участке ACB, то при перемещении в любом направлении потенциальная энергия возрастает, следовательно, кинетическая энергия $W_к$ убывает (сила, действующая на частицу, направлена к точке C). Поэтому если кинетическая энергия частицы внутри «ямы» мала, то она не сможет покинуть ее и будет совершать колебательное движение в окрестности точки C.

§ 13. СТОЛКНОВЕНИЕ ШАРОВ

Применим законы сохранения импульса и энергии для изучения столкновения шаров. Рассмотрим сначала простейший случай, когда шары движутся вдоль прямой, соединяющей их центры (рис. I.32), причем скорости их могут быть направлены навстречу или в одну сторону ($v_1 > v_2$). Полагая, во-первых, что на шары внешние силы не действуют *, и, во-вторых, что в процессе столкновения не происходит превращения механической энергии в тепловую, можно применить оба закона сохранения — импульса и механической энергии. Обозначим скорости шаров после удара через u_1 и u_2; тогда

$$m_1\mathbf{v}_1 + m_2\mathbf{v}_2 = m_1\mathbf{u}_1 + m_2\mathbf{u}_2, \qquad (3.17)$$

$$\frac{1}{2}m_1v_1^2 + \frac{1}{2}m_2v_2^2 = \frac{1}{2}m_1u_1^2 + \frac{1}{2}m_2u_2^2. \qquad (3.18)$$

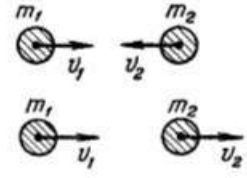


Рис. I.32

Для совместного решения этих уравнений необходимо первое векторное уравнение заменить скалярным; при этом должно быть учтено направление скоростей; например, скорости шаров, направленные вправо, можно полагать положительными, а влево — отрицательными. Ввиду этого для шаров, двигающихся навстречу, в левой части уравнения (3.17) должно быть написано $m_1v_1 - m_2v_2$. Однако нам неизвестны направления скоростей после удара; это затруднение можно обойти следующим образом: сделаем какое-либо предположение о направлении скоростей шаров после удара и решим совместно уравнения (3.17) и (3.18). Если наши предположения были правильными, то расчет даст для скоростей шаров положительные знаки; если же у какой-нибудь из этих скоростей получится отрицательный знак, то предварительно выбранное направление этой скорости следует изменить на обратное.

---

* Или они действуют перпендикулярно направлению движения тел и поэтому не изменяют величины скоростей тел и не совершают работу.

Для определения скоростей шаров, которые до удара двигались в одном направлении (рис. I.32), один догоняя другого, заменим векторное уравнение (3.17) скалярным:

$$m_1v_1 + m_2v_2 = m_1u_1 + m_2u_2. \tag{3.19}$$

Здесь предположено, что после удара оба шара по-прежнему движутся слева направо. Совместное решение уравнений (3.19) и (3.18) приводит к следующему результату:

$$\begin{aligned} u_1 &= \frac{(m_1 - m_2)\,v_1 + 2m_2v_2}{m_1 + m_2}, \\ u_2 &= \frac{(m_2 - m_1)\,v_2 + 2m_1v_1}{m_1 + m_2}. \end{aligned} \tag{3.20}$$

Исследуем эти формулы:

1) если массы шаров равны ($m_1 = m_2$), то $u_1 = +v_2$ и $u_2 = +v_1$. Это означает, что после удара оба шара движутся вправо, но обменявшись скоростями. Первоначально выбранное в уравнении (3.19) направление скоростей после удара оказалось правильным;

2) допустим, что второй шар покоится ($v_2 = 0$), причем $m_1 > m_2$; тогда

$$u_1 = \frac{m_1 - m_2}{m_1 + m_2}v_1; \quad u_2 = \frac{2m_1}{m_1 + m_2}v_1. \tag{3.21}$$

В этом случае обе скорости после удара имеют опять-таки положительные знаки, т. е. направлены, как было предположено в уравнении (3.19), вправо. Численные значения этих скоростей зависят от соотношения масс шаров и скорости первого шара. Заметим, что отношение кинетической энергии первого тела после столкновения $W_{к1} = m_1u_1^2/2$ к первоначальной $W_{к10} = m_1v_1^2/2$ равно

$$\frac{W_{к1}}{W_{к10}} = \left(\frac{m_1 - m_2}{m_1 + m_2}\right)^2;$$

3) допустим, что второй шар покоится, но он имеет бóльшую массу, чем первый. В этом случае очевидно, что формулы для расчета скоростей после удара будут совпадать с выражением (3.21), но ввиду того что $m_1 < m_2$, скорость первого шара u_1 получается отрицательной. Это означает, что после удара первый шар будет двигаться не вправо, как было предположено в выражении (3.19), а влево.

Если масса второго шара m_2 настолько велика, что по сравнению с ней массой первого шара m_1 можно пренебречь, то $u_1 \approx -v_1$, $u_2 \approx 0$, т. е. первый шар будет отскакивать от второго с той же скоростью, с какой ударяется. Такой же результат, очевидно, мы получим при столкновении шара с неподвижной плоскостью.

Столкновение шаров, движущихся навстречу, может быть рассмотрено аналогичным образом. При этом получаются те же формулы (3.20), но с заменой скорости v_2 на $-v_2$.

Мы предполагали, что шары движутся вдоль прямой, проведенной через их центры. Однако если это условие не соблюдается, то задача по-прежнему может быть решена при помощи уравнений (3.17)

и (3.18). Для этого необходимо произвести разложение векторов скоростей (рис. I.33) соударяющихся шаров w_1 и w_2 на составляющие, направленные вдоль линии центров (v_1 и v_2) и перпендикулярно ей (v_1' и v_2'). Затем следует решать задачу столкновения шаров, оперируя только скоростями, направленными по линии центров; скорости, перпендикулярные этой линии, остаются при ударе без изменений. Таким приемом можно доказать, что при ударе шара о плоскость под некоторым углом соблюдается равенство углов падения и отражения.

В процессе удара оба шара действуют друг на друга с упругими силами, которые, меняясь со временем по величине, остаются равными между собой и противоположно направленными в любой момент времени, пока шары находятся в соприкосновении. Эти внутренние силы взаимодействия сообщают соударяющимся шарам равные по величине и противоположные по направлению импульсы.

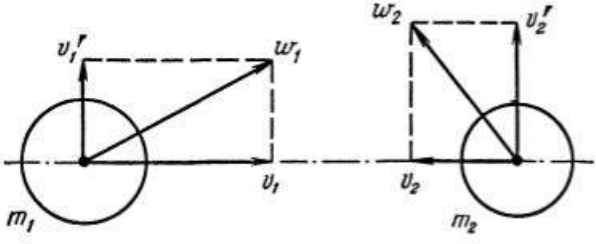


Рис. I.33

Однако это условие обеспечивает только выполнение закона сохранения импульса, но не гарантирует соблюдение закона сохранения механической энергии. При столкновении тел их деформация может происходить с внутренним трением, вследствие чего часть кинетической энергии тел перейдет в теплоту. Тогда вместо выражения (3.18) следует написать более общий вид закона сохранения энергии:

$$\frac{1}{2} m_1 v_1^2 + \frac{1}{2} m_2 v_2^2 = \frac{1}{2} m_1 u_1^2 + \frac{1}{2} m_2 u_2^2 + Q. \qquad (3.22)$$

Если количество выделившейся теплоты Q неизвестно, то определить скорости тел после удара невозможно, так как два используемых уравнения (3.17) и (3.22) будут содержать три неизвестных. Задача решается, если: 1) $Q = 0$ (*абсолютно упругий удар*), 2) $Q \neq 0$, но $u_1 = u_2$, т. е. тела после удара движутся вместе с одинаковыми скоростями (*абсолютно неупругий удар*). Однако во втором случае скорости тел после удара можно найти только из одного уравнения, выражающего закон сохранения импульса; тогда закон сохранения энергии (см. формулу (3.22)) можно использовать для вычисления количества выделяющегося тепла.

Если неупругий удар используется для каких-нибудь технологических целей (ковка изделий, забивание свай и т. п.), то можно оценить коэффициент полезного действия процесса удара как отношение полезно использованной энергии $W_{\text{пол}}$ к полной энергии W соударяю-

щихся тел до удара:

$$\eta = W_{\text{пол}}/W.$$

Рассмотрим простейший случай, когда $v_2 = 0$, $u_1 = u_2 = u$, $W = W_1 = m_1 v_1^2/2$. Тогда из (3.19) следует, что $m_1 v_1 = (m_1 + m_2)\, u$, а из соотношения (3.22)

$$Q = \frac{1}{2} m_1 v_1^2 - \frac{1}{2}(m_1 + m_2)\, u^2 = \frac{1}{2} m_1 v_1^2 \frac{m_2}{m_1 + m_2}.$$

Допустим, что полезной является работа, затрачиваемая на деформацию тел; эта работа при ударе превращается в теплоту. В этом случае коэффициент полезного действия

$$\eta = Q/W = m_2/(m_1 + m_2)$$

будет бóльшим при $m_1 \ll m_2$ (масса ударяющего тела значительно меньше массы деформируемого тела). Если же удар используется для передачи кинетической энергии, то Q должно быть минимальным, т. е.

$$\eta = \frac{W - Q}{W} = 1 - \frac{m_2}{m_1 + m_2} = \frac{m_1}{m_1 + m_2}$$

будет бóльшим при $m_1 \gg m_2$.

При столкновении реальных шаров результат удара определяется физическими свойствами вещества соударяющихся шаров. Измерения показали, что для шаров одинакового диаметра отношение

$$\frac{u_2 - u_1}{v_2 - v_1} = \varepsilon \tag{3.23}$$

имеет разные значения для различных веществ, из которых сделаны шары. Например, для стальных шаров $\varepsilon \approx 0{,}7$, бронзовых — 0,4, свинцовых — 0,20, деревянных — 0,6, стеклянных — 0,95 и т. д. Для абсолютно упругого удара ($Q = 0$) из выражения (3.20) можно получить, что $u_2 - u_1 = v_1 - v_2$, следовательно, $\varepsilon = 1$. Для абсолютно неупругого удара $u_1 = u_2$ и поэтому $\varepsilon = 0$. Величина ε, характеризующая столкновение, называется *коэффициентом восстановления*. Его легко определить, наблюдая уменьшение высоты подскоков шара, падающего на горизонтальную плоскость (из того же вещества). Чем ближе значение ε для данного вещества к единице, тем ближе столкновение к абсолютно упругому.

§ 14. О ЦЕНТРЕ МАСС СИСТЕМЫ

Законы механики для точечного тела, а также законы сохранения для системы взаимодействующих тел установлены относительно и:: циальных систем отсчета; таковой является система отсчета S_0, связанная с неподвижными звездами. Однако при изучении механических явлений (например, в лабораторных условиях) возникает необходимость в практически более удобных системах отсчета. Разумеется. можно любую составную часть изучаемой системы условно полагать неподвижной и с ней связать координатные оси, но полученная таким

образом система отсчета может оказаться неинерциальной. Например, в изолированной системе из двух притягивающихся тел можно одно из них принять за неподвижное; тогда второе тело будет с ускорением приближаться к «неподвижному» первому и общий импульс (состоящий в данном случае только из импульса второго тела) с течением времени увеличивается. Следовательно, при таком выборе «неподвижного» тела закон сохранения импульса в системе (3.6) не будет соблюдаться, несмотря на то что система изолирована и внешние силы отсутствуют; точно так же не соблюдается и закон сохранения механической энергии (см. выражение (3.13)).

Для нахождения инерциальных систем отсчета поступим следующим образом. Допустим, что S_0 — известная инерциальная система отсчета, относительно которой изучается взаимодействие двух тел

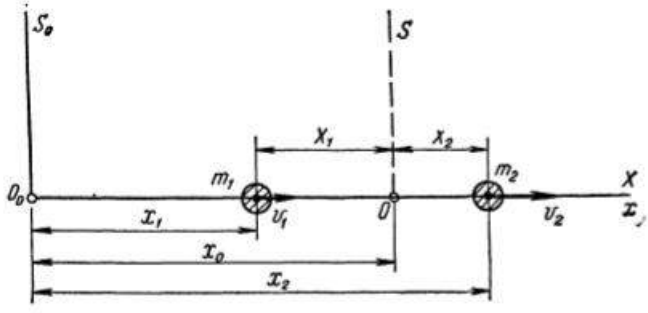


Рис. 1.34

(рис. 1.34). Для простоты рассуждений предположим, что движение тел происходит по оси OX. Найдем точку O, координата которой x_0 удовлетворяет условию

$$x_0 = \frac{m_1 x_1 + m_2 x_2}{m_1 + m_2}, \tag{3.24}$$

скорость движения этой точки

$$v_0 = \frac{dx_0}{dt} = \frac{m_1 v_1 + m_2 v_2}{m_1 + m_2}. \tag{3.25}$$

Если движение тел происходит только под действием внутренних сил взаимодействия (т. е. внешние силы отсутствуют и система является изолированной), то числитель выражения (3.25), представляющий собой суммарный импульс в системе, с течением времени изменяться не будет. Следовательно, скорость v_0 перемещения точки O относительно S_0 будет постоянной величиной. Если теперь связать с точкой O новую систему отсчета S, то она будет перемещаться относительно S_0 прямолинейно и равномерно и поэтому также будет инерциальной системой отсчета.

Приведенные выше рассуждения можно повторить и для любой изолированной системы взаимодействующих тел. Для каждой изучаемой системы тел можно найти точку O, координаты которой относи-

тельно S_0 равны:

$$x_0=\frac{m_1x_1+m_2x_2+\dots}{m_1+m_2+\dots};\quad y_0=\frac{m_1y_1+m_2y_2+\dots}{m_1+m_2+\dots};\quad z_0=\frac{m_1z_1+m_2z_2+\dots}{m_1+m_2+\dots}. \tag{3.26}$$

Если связать с этой точкой координатные оси, то получим новую инерциальную систему отсчета. Точка O называется *центром масс* (или *центром инерции*) данной системы (см. также § 6). Таким образом, *для каждой изолированной системы взаимодействующих тел система отсчета, связанная с ее центром масс, является инерциальной.*

Покажем, что суммарный импульс системы тел относительно системы отсчета S, связанной с центром масс, равен нулю. В частном случае для системы двух тел m_1 и m_2 (рис. I.34), координаты которых относительно S обозначены через X_1 и X_2, имеем:

$$X_1=x_1-x_0;\quad X_2=x_2-x_0;$$

$$m_1\frac{dX_1}{dt}+m_2\frac{dX_2}{dt}=-(m_1+m_2)\frac{dx_0}{dt}+\left(m_1\frac{dx_1}{dt}+m_2\frac{dx_2}{dt}\right).$$

Подставив в эту формулу значение dx_0/dt из выражения (3.25), получим

$$m_1\frac{dX_1}{dt}+m_2\frac{dX}{dt}=0.$$

В более общем случае можно получить:

$$\sum m_i\frac{dX_i}{dt}=0;\quad \sum m_i\frac{dY_i}{dt}=0;\quad \sum m_i\frac{dZ_i}{dt}=0,$$

где dX_i/dt, dY_i/dt и dZ_i/dt — проекции скоростей составных частей рассматриваемой системы тел на координатные оси системы отсчета S.

Допустим теперь, что на составные части системы действуют внешние силы; тогда ускорение каждого из этих тел относительно инерциальной системы отсчета S_0 определяется суммой внутренних и внешних сил, например $m_1\mathbf{a}_1=\mathbf{f}_1+\mathbf{F}_1$; $m_2\mathbf{a}_2=\mathbf{f}_2+\mathbf{F}_2$ и т. д. При суммировании этих уравнений внутренние силы, согласно третьему закону механики, исчезнут, поэтому

$$m_1\mathbf{a}_1+m_2\mathbf{a}_2+\dots=\mathbf{F}_1+\mathbf{F}_2+\dots$$

Обозначим векторную сумму внешних сил через $\mathbf{R}$. Разложим эту силу и ускорения тел на составляющие по координатным осям; для оси OX получим

$$m_1\frac{d^2x_1}{dt^2}+m_2\frac{d^2x_2}{dt^2}+\dots=R_x. \tag{3.27}$$

Из выражения для x_0, взяв вторую производную по времени, получим

$$(m_1+m_2+\dots)\frac{d^2x_0}{dt^2}=m_1\frac{d^2x_1}{dt^2}+m_2\frac{d^2x_2}{dt^2}+\dots \tag{3.28}$$

Обозначим полную массу системы через m; тогда, сопоставляя уравнения (3.27) и (3.28), получим

$$m\frac{d^2x_0}{dt^2}=R_x.$$

Такие же выражения получаются для составляющих по другим осям. Следовательно, полное ускорение центра масс относительно инерциальной системы отсчета S_0 определится из уравнения

$$m\mathbf{a} = \mathbf{R}; \quad \mathbf{a} = \mathbf{R}/m, \tag{3.29}$$

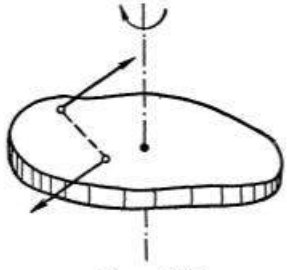


Рис. 1.35

т. е. *центр масс системы перемещается с ускорением, равным отношению равнодействующей всех внешних сил к массе всей системы.* Например, допустим, что на тело действуют две равные и противоположно направленные силы («пара сил»), стремящиеся сообщить ему вращательное движение (рис. 1.35). Так как равнодействующая пары сил равна нулю, то центр масс останется неподвижным, следовательно, тело будет вращаться вокруг оси, проходящей через центр масс, независимо от того, в каких точках приложены силы.

Глава 4

КОЛЕБАТЕЛЬНОЕ ДВИЖЕНИЕ ТЕЛ

§ 15. ПЕРИОДИЧЕСКОЕ ДВИЖЕНИЕ; ГАРМОНИЧЕСКИЕ КОЛЕБАНИЯ

Периодическими процессами называются такие изменения состояния системы, при которых она многократно, через некоторые промежутки времени, возвращается в одно и то же состояние. Простейшее периодическое движение — это вращение тел; к ним относятся также многократно повторяющиеся движения тел по любым замкнутым кривым, например движения планет по эллиптическим орбитам и т. д. Периодическими процессами являются также *колебательные процессы*, когда система последовательно отклоняется от своего положения равновесия то в одну, то в противоположную сторону. Простейшим примером колебательного движения является движение точечной массы m, подвешенной на нити или пружине, около положения равновесия — точки O (рис. 1.36).

Периодические процессы характеризуются последовательностью состояний, через которые проходит система в течение одного периода. Если эта последовательность точно повторяется через равные промежутки времени, то колебания называются *незатухающими*. При нарастающих или затухающих колебаниях периодически повторяются только определенные состояния системы, например прохождение колеблющегося тела через положение равновесия и т. п.

Среди множества различных незатухающих колебаний простейшим является *гармоническое* колебательное движение, описываемое функцией синуса или косинуса:

$$x = x_0 \sin(\omega t + \varphi), \quad x = x_0 \cos(\omega t + \varphi), \tag{4.1}$$

где x — колеблющаяся величина (смещение, скорость, сила и т. д.); t — время, x_0, ω и φ — некоторые постоянные величины. Величина x_0 называется *амплитудой*, аргумент синуса или косинуса $(\omega t + \varphi)$ — *фазой* колебания, а величина φ — *начальной фазой*. Фаза колебания определяет значение колеблющейся величины в данный момент времени. Начальная фаза определяет значение x в начальный момент времени: для

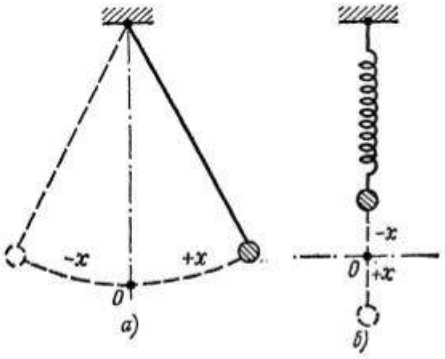


Рис. 1.36

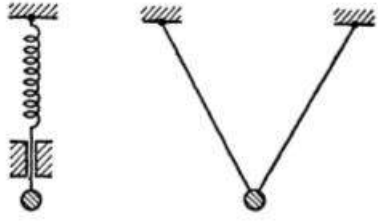


Рис. 1.37

синусоидального колебания при $t = 0$ $x = x_0 \sin \varphi$. Если, изучая колебательное движение, начинать отсчет времени ($t = 0$) при $x = 0$, то φ окажется равным нулю. Во всех случаях, когда рассматривается одно колебание, можно выбрать начало отсчета времени так, чтобы $\varphi = 0$; однако при одновременном существовании нескольких колебаний (например, при сложении колебаний) начальные фазы каждого

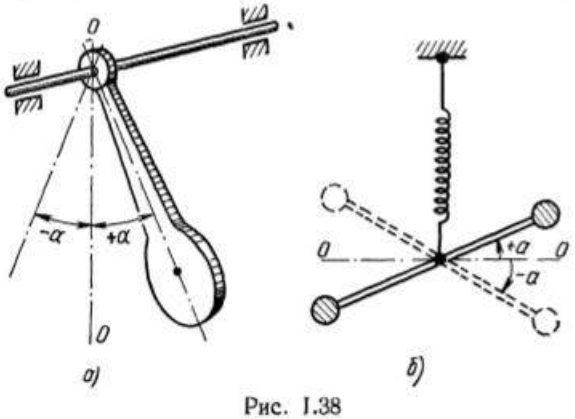


Рис. 1.38

колебания отличаются друг от друга и лишь в частных случаях эти фазы могут одновременно равняться нулю.

Формула (4.1) описывает гармонические колебательные движения, происходящие вдоль какой-нибудь линии — отрезка прямой или кривой. В этом случае для определения положения колеблющегося тела достаточно задать только расстояние x от тела до положения равновесия. Колебательные системы, в которых возможно только одно

колебательное движение (вдоль одной линии), изображены на рис. I.37; их называют колебательными системами с одной *степенью свободы*. Простой маятник (см. рис. I.36, *а*) может совершать два независимых друг от друга колебания в двух взаимно перпендикулярных направлениях, поэтому его относят к колебательным системам, обладающим двумя степенями свободы. Пружинный маятник, изображенный на рис. I.36, *б*, может колебаться в трех независимых направлениях и поэтому является колебательной системой с тремя степенями свободы.

Для описания колебательного движения сплошного твердого тела (рис. I.38, *а*) удобнее измерять углы поворота α от равновесного состояния; углы, отсчитываемые по одну сторону от OO, принимаются положительными, по другую сторону — отрицательными. Аналогичное правило знаков выбирается и для тел, совершающих так называемые *крутильные колебания* (рис. I.38, *б*). Гармонические колебания для углов поворота имеют вид $\alpha = \alpha_0 \sin(\omega t + \varphi)$, где α_0 — амплитуда угла поворота.

§ 16. СКОРОСТЬ И УСКОРЕНИЕ КОЛЕБЛЮЩЕГОСЯ ТЕЛА. ДИФФЕРЕНЦИАЛЬНОЕ УРАВНЕНИЕ ГАРМОНИЧЕСКИХ КОЛЕБАНИЙ

Допустим, что x — смещение колеблющегося тела относительно положения равновесия, причем начало отсчета времени выбрано так, чтобы $\varphi = 0$. Тело совершает гармоническое колебание, если

$$x = x_0 \sin \omega t, \tag{4.2}$$

причем x_0 и ω должны быть постоянными величинами. Скорость тела найдем как производную от x по времени:

$$v = \frac{dx}{dt} = x_0 \omega \cos \omega t; \quad v = v_0 \cos \omega t = v_0 \sin\left(\omega t + \frac{\pi}{2}\right), \tag{4.3}$$

где $v_0 = x_0\omega$ — максимальное значение (т. е. амплитуда) скорости. Формула (4.3) показывает, что скорость тела, так же как и смещение, изменяется со временем по гармоническому закону с той же частотой ω, но имеет фазу, отличающуюся от фазы смещения на $\pi/2$; в моменты времени, когда смещение x равно нулю, скорость тела приобретает наибольшие значения.

Так как скорость тела при гармоническом колебании непрерывно изменяется, то это движение является ускоренным; при этом величина ускорения изменяется со временем также по гармоническому закону:

$$a = \frac{dv}{dt} = \frac{d}{dt}(x_0 \omega \cos \omega t) = -x_0 \omega^2 \sin \omega t = a_0 \sin(\omega t + \pi), \tag{4.4}$$

где $a_0 = x_0\omega^2 = v_0\omega$ — максимальное (амплитудное) значение ускорения. Фаза ускорения отличается от фазы смещения на π, а от фазы скорости — на $\pi/2$. Заменив в выражении (4.4) $x_0 \sin \omega t$ на x, можно переписать формулу ускорения в виде

$$a = -\omega^2 x, \tag{4.5}$$

т. е. *при гармоническом колебательном движении ускорение тела прямо пропорционально смещению от положения равновесия и имеет*

противоположный ему знак. На рис. I.39 показаны изменения смещения x, скорости v и ускорения a с течением времени.

Периодом гармонического колебательного движения называется наименьшее время T, по истечении которого *все* величины, характеризующие это движение (x, v, a), в точности принимают первоначальные значения. Для того чтобы все тригонометрические функции (4.2), (4.3) и (4.4) одновременно приняли первоначальные значения, их аргументы (т. е. фазы) должны измениться на $2\pi n$, где n — целое число.

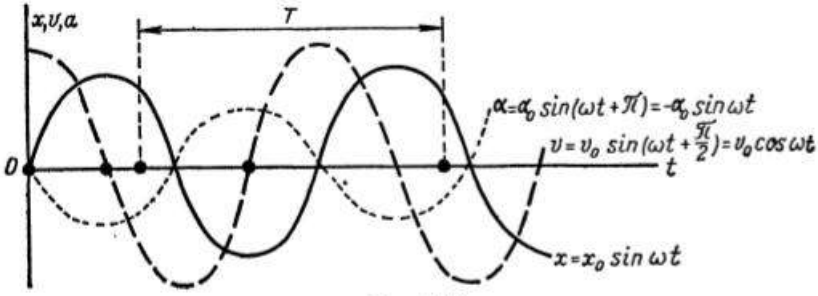


Рис. 1.39

Период колебания соответствует изменению фазы на 2π. Если в момент t фаза колебаний какой-нибудь величины была равна $\omega t + \varphi$, то через время T фаза оказывается равной $\omega (t + T) + \varphi$.

Приравняв изменение фазы 2π, получим:

$$[\omega (t+T)+\varphi]-(\omega t+\varphi)=2\pi;$$

$$\omega T=2\pi; \quad \omega=\frac{2\pi}{T}.$$

Величина

$$\nu=\frac{1}{T}=\frac{\omega}{2\pi}$$

называется *частотой колебаний* и показывает, сколько раз в единицу времени повторяется одно и то же состояние колеблющегося тела. Величина ω, также имеющая размерность 1/с, называется *угловой* (или круговой) частотой колебаний. В формуле гармонического колебания фаза может быть выражена не только через ω, но и через ν или T:

$$x=x_0 \sin(\omega t+\varphi)=x_0 \sin\left(2\pi \frac{t}{T}+\varphi\right)=x_0 \sin(2\pi\nu t+\varphi). \qquad (4.6)$$

Формула (4.5), переписанная в виде

$$\frac{d^2x}{dt^2}+\omega^2 x=0, \qquad (4.7)$$

есть *дифференциальное уравнение гармонических колебаний*; оно связывает колеблющуюся величину $x(t)$ с ее второй производной по времени. Можно утверждать, что если $x(t)$ изменяется со временем согласно формуле (4.1), то она удовлетворяет дифференциальному уравнению (4.7). Верно и обратное утверждение: если какая-либо пере-

менная физическая величина $x(t)$ удовлетворяет дифференциальному уравнению (4.7), то, следовательно, она изменяется со временем по гармоническому (синусоидальному или косинусоидальному) закону (4.1) с постоянными значениями x_0, ω и φ.

Для описания гармонических колебаний часто употребляется аналитически очень удобная комплексная функция

$$x = x_0 e^{i(\omega t + \varphi)}, \tag{4.8}$$

которая в своей действительной части содержит функцию косинуса, а в мнимой — функцию синуса:

$$e^{i\alpha} = \cos\alpha + i\sin\alpha,$$

где $i = \sqrt{-1}$. Легко показать, что функция (4.8) — периодическая, т. е. $x(t+T) = x(t)$, где $T = 2\pi/\omega$, и, кроме того, она удовлетворяет дифференциальному уравнению (4.7).

§ 17. СИЛА И ЭНЕРГИЯ ПРИ ГАРМОНИЧЕСКИХ КОЛЕБАНИЯХ. ПРОСТЕЙШИЕ МЕХАНИЧЕСКИЕ КОЛЕБАТЕЛЬНЫЕ СИСТЕМЫ

Всякое колебательное движение есть движение, происходящее с ускорением, поэтому на колеблющиеся тела должны действовать силы, сообщающие им эти ускорения. В частности, если точечное тело массой m совершает гармоническое колебание, то, согласно второму закону механики, на него должна действовать сила, равная

$$F = ma = -m\omega^2 x = -kx, \tag{4.9}$$

где $k = m\omega^2$. Направление силы совпадает с направлением ускорения, а вектор ускорения при гармонических колебаниях, согласно формуле (4.5), всегда направлен к положению равновесия. Таким образом, *для того чтобы тело совершало гармоническое колебательное движение, на него должна действовать сила, всегда направленная к положению равновесия, а по величине — прямо пропорциональная смещению от этого положения.* При исследовании колебательных систем можно легко найти коэффициент пропорциональности $k = F/x$ между действующей на тело силой F и смещением x этого тела от положения равновесия; тогда, зная еще и массу колеблющегося тела, можно вычислить частоту и период колебания; из соотношения $m\omega^2 = k$ следует:

$$\omega = \sqrt{\frac{k}{m}}, \quad T = \frac{2\pi}{\omega} = 2\pi\sqrt{\frac{m}{k}}. \tag{4.10}$$

Силы, всегда направленные к положению равновесия, называются *возвращающими*. Рассмотрим несколько примеров:

1. Колебательная система, состоящая из массы m и пружины (см. рис. 1.36, *б*). Возвращающей силой является упругая сила, действующая на тело со стороны деформированной пружины. Эта сила F при малых деформациях прямо пропорциональна изменению длины пружины $F = kx = k(l - l_0)$. Приложив к пружине внешние силы и измерив вызванные ими удлинения

(или сжатия) пружины, можно найти коэффициент упругости пружины $k = F/x$ и по формуле (4.10) рассчитать частоту колебаний тел, прикрепленных к концам пружины. При этом колебания будут гармоническими (т. е. x_0 и ω постоянны) только в том случае, если на колеблющееся тело не действуют никакие другие силы, кроме возвращающей $F = -kx$, причем коэффициент k, от которого, согласно формуле (4.10), зависит частота колебаний, должен все время сохраняться постоянным. В частности, если температура пружины изменяется, то k, а следовательно, и частота колебаний также изменяются; колебания не будут гармоническими.

2. Система, совершающая крутильные (поворотные) колебания (см. рис. I.38, *б*). При крутильных колебаниях на тело действует *возвращающий момент,* приостанавливающий отклонение тела от состояния равновесия и затем сообщающий ему обратное движение. Возвращающий момент возникает при деформации (кручении) пружины (или стержня), к которой прикреплено колеблющееся тело. При малых углах отклонения этот момент прямо пропорционален углу отклонения.

Если крутильные колебания гармонические, т. е.

$$\alpha = \alpha_0 \sin \omega t, \tag{4.11}$$

то угловая скорость и угловое ускорение ε при повороте также изменяются по гармоническому закону:

$$\frac{d\alpha}{dt} = \alpha_0 \omega \cos \omega t; \quad \varepsilon = -\alpha_0 \omega^2 \sin \omega t = -\omega^2 \alpha. \tag{4.12}$$

Возвращающий момент найдем как произведение углового ускорения на момент инерции колеблющегося тела:

$$M = J\varepsilon = -J\omega^2 \alpha = -D\alpha, \tag{4.13}$$

где $D = J\omega^2$ — постоянная величина (если момент инерции тела при колебаниях не изменяется). Этот коэффициент можно найти, приложив к пружине (или стержню) внешние скручивающие моменты M и измеряя углы скручивания α:

$$D = \frac{M}{\alpha};$$

тогда частота и период колебаний определяются по формулам:

$$\omega = \sqrt{\frac{D}{J}}; \quad T = \frac{2\pi}{\omega} = 2\pi \sqrt{\frac{J}{D}}. \tag{4.14}$$

Согласно выражению (4.13), при гармонических крутильных колебаниях возвращающий момент должен быть точно пропорционален углу отклонения; если эта пропорциональность не соблюдается (например, при очень больших углах поворота), то колебания не будут гармоническими (хотя при отсутствии трения будут незатухающими).

3. Физический маятник (рис. I.40). Возвращающим моментом является момент силы тяжести, имеющий знак, противопо-

ложный знаку угла отклонения α и равный

$$M = -Px = -mgl_0 \sin\alpha,$$

где l_0 — расстояние от точки опоры до центра тяжести тела.

При малых углах отклонения $\sin\alpha \approx \alpha$ (угол α — в радианах); тогда возвращающий момент

$$M = -mgl_0\alpha = -D\alpha$$

пропорционален углу отклонения и колебания маятника будут гармоническими.

Сравнивая с выражением (4.13), получим $J\omega^2 = mgl_0$ и, следовательно,

$$\omega = \sqrt{\frac{mgl_0}{J}}; \quad T = 2\pi\sqrt{\frac{J}{mgl_0}}. \qquad (4\ 15)$$

При больших углах отклонения, а также при деформации тела во время колебаний (переменные J и l_0) колебания оказываются негармо-

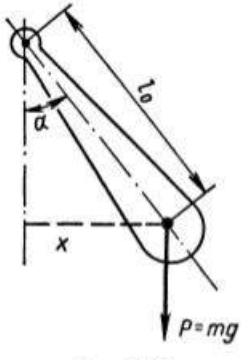


Рис. I.40

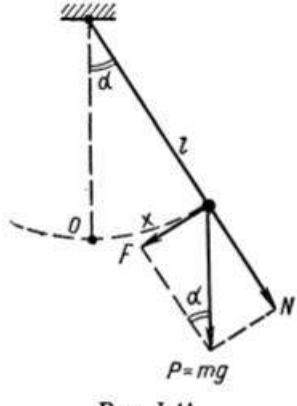


Рис. I.41

ническими, хотя они при отсутствии или компенсации трения могут быть незатухающими.

4. Математический маятник представляет собой точечное тело массой m, подвешенное к невесомой и нерастяжимой нити длиной l (рис. I.41). Возвращающей силой является проекция силы тяжести $P = mg$ на направление движения тела; имеем:

$$\alpha = \frac{x}{l}; \quad F = P\sin\alpha = mg\sin\frac{x}{l}$$

(α — в радианах). Замечаем, что условие пропорциональности между возвращающей силой F и смещением от положения равновесия x здесь также не соблюдается, поэтому колебания этого маятника не являются гармоническими. Но если углы α малы, так что $\sin\alpha \approx \alpha$, то

$$F = mg\sin\frac{x}{l} \approx mg\frac{x}{l};$$

так как эта сила всегда направлена к положению равновесия и поэтому имеет знак, противоположный знаку x, то

$$F = -\frac{mg}{l}x = -kx.$$

В этом случае колебания можно полагать гармоническими; сравнивая с выражением (4.9), получаем:

$$\frac{mg}{l} = m\omega^2; \quad \omega = \sqrt{\frac{g}{l}}; \quad T = 2\pi\sqrt{\frac{l}{g}}, \tag{4.16}$$

т. е. частота и период колебаний не зависят от массы колеблющегося тела, а определяются только длиной нити и ускорением силы тяжести (колебаниями маятников пользуются для определения g). Для постоянства коэффициента k, а следовательно, и частоты колебаний ω необходимо постоянство l. Между тем сила $N = P\cos\alpha$, действующая вдоль нити, может вызвать ее удлинение, которое будет минимальным в крайних положениях и максимальным при прохождении тела через точку O. Поэтому, чтобы колебания маятника были гармоническими, необходимо кроме малости углов отклонения дополнительно еще и условие нерастяжимости нити.

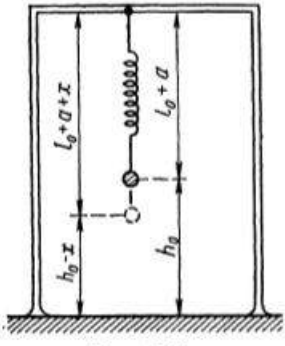


Рис. 1.42

Из этих примеров видно, что при малых амплитудах частота (или период) колебаний определяется только свойствами системы. Однако при больших отклонениях от положения равновесия линейная зависимость возвращающей силы от смещения $F = -kx$, а также возрастающего момента от угла поворота $M = -D\alpha$ строго не соблюдается и частота колебаний зависит в некоторой степени также и от амплитуды колебаний (x_0 или α_0).

Колебательные движения в механических системах сопровождаются периодическими превращениями кинетической энергии колеблющихся тел в потенциальную энергию взаимодействия частей системы и обратно. При этом *энергией колебаний* называют ту часть полной энергии системы, которая участвует в этих превращениях.

Например, энергия пружинного маятника, колеблющегося в поле тяготения Земли, состоит из потенциальной энергии деформированной пружины, потенциальной энергии положения груза и его кинетической энергии (рис. 1.42):

$$W = \frac{k(a+x)^2}{2} + mg(h_0 - x) + \frac{mv^2}{2},$$

где a — постоянное удлинение пружины, вызванное силой тяжести; $a = P/k = mg/k$; h_0 — высота груза в равновесном состоянии. Сокращая, получим

$$W = \left(\frac{ka^2}{2} + mgh_0\right) + \left(\frac{kx^2}{2} + \frac{mv^2}{2}\right) = W_0 + W_\tau. \tag{4.17}$$

Переменная часть этого выражения есть энергия колебаний в системе:

$$W_\tau = \frac{kx^2}{2} + \frac{mv^2}{2}.$$

Энергию колебательного движения можно представить в зависимости от амплитудных значений смещения и скорости: при $x = 0$ $v = v_0$; при $x = x_0$ $v = 0$. Следовательно,

$$W_\tau = \frac{mv_0^2}{2} = \frac{kx_0^2}{2} = \frac{m\omega^2 x_0^2}{2}. \quad (4.18)$$

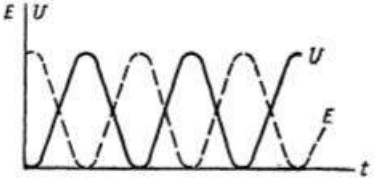


Рис. I.43

Таким образом, энергия колебаний W_τ периодически переходит из кинетической формы в потенциальную; период этих превращений вдвое меньше периода самих колебаний, так как амплитудные значения смещения $\pm x_0$ или скорости $\pm v_0$ появляются два раза за период, а энергия W_τ не зависит от знака этих величин. На рис. I.43 показаны изменения со временем составных частей этой энергии: потенциальной

$$W_\text{п} = \frac{kx^2}{2} = \frac{kx_0^2}{2} \sin^2 \omega t$$

и кинетической

$$W_\text{к} = \frac{mv^2}{2} = \frac{mv_0^2}{2} \cos^2 \omega t.$$

§ 18. СОБСТВЕННЫЕ, СВОБОДНЫЕ И ВЫНУЖДЕННЫЕ КОЛЕБАНИЯ. ПАРАМЕТРИЧЕСКОЕ ВОЗБУЖДЕНИЕ КОЛЕБАНИЙ

Если система изолирована от внешних воздействий и колебательное движение происходит без трения, то ее полная механическая энергия W с течением времени не изменяется. Тогда из постоянства W и W_0 следует, что и энергия колебаний W_τ будет также оставаться с течением времени постоянной. В реальных системах эта энергия может уменьшаться не только вследствие трения, но и при наличии излучения, когда колеблющиеся тела приводят в периодическое движение соприкасающиеся с ними частицы окружающей среды и в этой среде образуются упругие волны. Например, камертон или струна вызывает в воздухе звуковые волны, постепенно расходуя на это свою энергию колебаний.

Потери энергии в системе вследствие трения или излучения можно компенсировать при помощи внешнего воздействия на систему, тогда колебания могут стать незатухающими. Сообщение системе энергии извне можно осуществить, например, путем прерывистых, периодически повторяющихся воздействий, при которых система сразу получает значительное количество энергии и затем постепенно расходует их на трение или излучение. Так поступают, ударяя молоточком по маятнику, камертону или струне. Колебания при таких воздействиях носят сложный характер; после получения энергии амплитуда коле-

баний резко возрастает и затем вследствие потерь постепенно убывает до следующего удара. Заметим, что при этих ударах внешняя сила, приложенная к колеблющемуся телу, увеличивает его энергию только в том случае, если она совершает положительную работу, т. е. действует в направлении движения этого тела.

Рассмотрим другой способ компенсации потерь, когда энергия сообщается системе так же непрерывно, как она расходуется на трение или излучение. Это можно осуществить, если к колеблющемуся телу приложить непрерывно действующую силу. Очевидно, для этой цели постоянная по величине и направлению сила непригодна, так как она при движении тела в одном направлении будет совершать положительную работу, а в противоположном направлении — равную отрицательную работу; действие такой силы выразится только в смещении положения равновесия, вокруг которого происходят колебания. Например, сила тяжести, действующая на пружинный маятник (см. рис. I.42), не может предотвратить затухания колебания, а только опускает точку, соответствующую положению равновесия колеблющегося тела. Поэтому для компенсации потерь на трение или излучение непрерывно действующая внешняя сила должна быть переменной; только в этом случае в течение каждого периода колебаний положительная работа внешней силы может быть больше отрицательной.

Рассмотрим простейшую колебательную систему с одной степенью свободы, в которой колебания происходят вдоль некоторой линии, например оси OX. Согласно второму закону механики, движение тела должно определяться уравнением

$$m\frac{d^2x}{dt^2} = \sum F_i. \tag{4.19}$$

Правая часть этого уравнения есть сумма всех сил, действующих на тело; допустим, что в этой сумме содержатся только три силы:

1) *сила упругости*, пропорциональная смещению тела от положения равновесия (см. формулу (4.9)),

$$F_1 = -kx;$$

2) *сила трения*, пропорциональная первой степени скорости тела,

$$F_2 = -rv = -r\frac{dx}{dt} \tag{4.20}$$

(знак минус указывает, что направление силы трения всегда противоположно направлению скорости движения);

3) *внешняя сила*, непрерывно изменяющаяся со временем по какому-нибудь закону; в частности, предположим, что внешняя сила является синусоидальной:

$$F_3 = F_0 \sin \Omega t. \tag{4.21}$$

Если силы трения и внешняя сила отсутствуют, формула (4.19) дает дифференциальное уравнение

$$m\frac{d^2x}{dt^2} + kx = 0, \tag{4.22}$$

которое имеет решение в виде

$$x = x_0 \sin(\omega_0 t + \varphi),$$

где $\omega_0 = \sqrt{k/m}$. Амплитуда колебаний x_0 и начальная фаза φ должны быть определены из дополнительно заданного состояния системы в какой-нибудь момент времени, например при $t = 0$; частота колебаний ω_0 определяется только свойствами колебательной системы: массой колеблющегося тела m и коэффициентом упругой силы k. Амплитуда колебаний может быть сделана различной в зависимости от той энергии, которая была сообщена системе при выводе ее из состояния покоя. Начальная фаза φ определяется выбором того момента t, с которого начинается отсчет времени.

Колебания, происходящие при отсутствии внешних сил и трения, называются собственными; частота собственных колебаний зависит только от свойств системы.

Допустим теперь, что в системе действуют две силы: F_1 и F_2; уравнение движения тела будет иметь вид

$$m \frac{d^2x}{dt^2} = -kx - r \frac{dx}{dt}.$$

Разделим это уравнение на массу тела и обозначим:

$$\frac{k}{m} = \omega_0^2; \quad \frac{r}{m} = 2\delta. \tag{4.23}$$

Тогда получим дифференциальное уравнение

$$\frac{d^2x}{dt^2} + 2\delta \frac{dx}{dt} + \omega_0^2 x = 0. \tag{4.24}$$

Этому уравнению удовлетворяет функция

$$x = x_0 e^{-\delta t} \cos \omega t, \tag{4.25}$$

где

$$\omega = \sqrt{\omega_0^2 - \delta^2}.$$

График этой функции изображен на рис. I.44. Начальная амплитуда колебаний x_0 должна быть задана дополнительно.

Таким образом, если на тело кроме силы упругости действует сила трения, пропорциональная первой степени скорости движения, то тело будет совершать колебательное (но не гармоническое!) движение с частотой, зависящей от m, k и δ. Амплитуда колебаний (т. е. значение x при $\cos \omega t = 1$) будет с течением времени изменяться по экспоненциальному закону $x_0 e^{-\delta t}$.

Величина

$$\delta = r/2m,$$

определяющая быстроту убывания амплитуды колебаний с течением времени, называется *коэффициентом затухания* и имеет размерность 1/с. Произведение коэффициента затухания на период колебания T

$$d = \delta T, \tag{4.26}$$

равное логарифму отношения двух соседних амплитуд:

$$x_1 = x_0 e^{-\delta t}; \quad x_2 = x_0 e^{-\delta (t+T)}; \quad \ln (x_1/x_2) = \delta T = d,$$

есть безразмерная величина и называется *логарифмическим декрементом затухания*.

Заметим, что при нелинейной зависимости силы трения от скорости колебание тела будет иметь более сложный вид.

Колебания, происходящие в системе при отсутствии внешних сил (но при наличии потерь на трение или излучение), *называются* с в о б о д н ы м и. Частота свободных колебаний зависит от свойств системы и интенсивности потерь; с увеличением δ эта частота уменьшается.

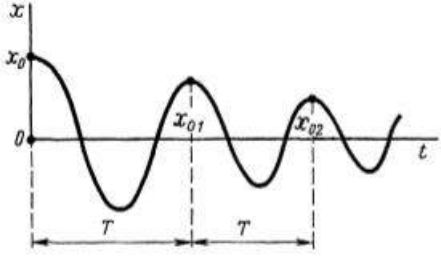


Рис. I.44

Рассмотрим более общую задачу — определение движения тела, на которое действуют все три упомянутые выше силы:

$$m \frac{d^2x}{dt^2} = -kx - rv + F_0 \sin \Omega t. \qquad (4.27)$$

Разделим это уравнение на массу тела и к обозначениям (4.23) добавим

$$f_0 = \frac{F_0}{m}.$$

Тогда уравнение примет вид

$$\frac{d^2x}{dt^2} + 2\delta \frac{dx}{dt} + \omega_0^2 x = f_0 \sin \Omega t. \qquad (4.28)$$

Внешняя сила будет совершать работу, знак которой зависит от разности фаз между силой и скоростью движения тела. Если направление внешней силы противоположно направлению движения (т. е. скорости колеблющегося тела), то она совершает отрицательную работу и поэтому тормозит движение колеблющегося тела; если же направление силы совпадает с направлением движения тела, то она совершает положительную работу и, следовательно, ускоряет движение тела. Со временем это приведет к тому, что тело будет вынуждено колебаться с той же частотой, с какой изменяется внешняя сила.

К такому же выводу можно прийти, если проанализировать уравнение (4.27). Оно должно соблюдаться для каждого момента времени, поэтому если сила F со временем изменяется по какому-нибудь закону, то вслед за ней и одновременно должны изменяться смещение, скорость и ускорение колеблющегося тела; следовательно, частоты изменения этих величин должны совпадать с частотой изменения внешней силы.

Однако фазы колебаний этих величин могут отличаться от фазы внешней силы; если, например, внешняя сила достигла нуля, то это вовсе не означает, что одновременно должны равняться нулю каждое

из выражений m ($\mathrm{d}^2x/\mathrm{d}t^2$), kx и rv, так как их знаки могут быть различными. В частности, для компенсации потерь на трение разность фаз между внешней силой и скоростью движения должна быть такой, чтобы результирующая работа внешней силы была положительной и равнялась работе силы трения.

Допустим, что разность фаз между внешней силой и смещением колеблющегося тела равна φ; так как тело вынуждено совершать колебания с частотой внешней силы, то

$$x = x_0 \sin(\Omega t + \varphi) \tag{4.29}$$

(эту функцию можно получить аналитически, решая уравнение (4.28)). Тогда скорость и ускорение будут иметь фазы, отличающиеся дополнительно на $\pi/2$ и π:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = x_0\Omega \sin\left(\Omega t + \varphi + \frac{\pi}{2}\right); \quad \frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -x_0\Omega^2 \sin(\Omega t + \varphi + \pi). \tag{4.30}$$

Нам необходимо найти x_0 и φ. Подставим выражения (4.29) и (4.30) в уравнение (4.28), произведем упрощения и приравняем нулю отдельно коэффициенты перед $\sin \Omega t$ и $\cos \Omega t$. В результате вычислений получим:

$$x_0 = \frac{f_0}{\sqrt{(\omega_0^2 - \Omega^2)^2 + 4\Omega^2\delta^2}}; \quad \operatorname{tg}\varphi = -\frac{2\delta\Omega}{\omega_0^2 - \Omega^2}. \tag{4.31}$$

Таким образом, если на колеблющееся тело действует периодическая синусоидальная сила с частотой Ω, то тело совершает колебания с той же частотой, причем амплитуда колебаний будет зависеть от амплитуды и частоты внешней силы, от коэффициента затухания, от упругих свойств системы и массы колеблющегося тела; такие колебания называются *вынужденными*.

При некоторой частоте внешней силы знаменатель в выражении (4.31) для x_0 будет иметь минимальное, а амплитуда вынужденных колебаний — максимальное значение. Для нахождения этой частоты (называемой *резонансной*) приравняем нулю производную:

$$\frac{\mathrm{d}}{\mathrm{d}\Omega}\left[(\omega_0^2 - \Omega^2)^2 + 4\Omega^2\delta^2\right] = 0,$$

откуда следует

$$\Omega_{\text{рез}} = \sqrt{\omega_0^2 - 2\delta^2}. \tag{4.32}$$

Таким образом, для данной колебательной системы, имеющей собственную частоту колебаний ω_0, резонансная частота внешней силы (при которой амплитуда вынужденных колебаний достигает максимума) зависит от коэффициента затухания. Наименьшая частота $\Omega_{\text{рез}} \approx 0$ соответствует предельному соотношению $2\delta^2 \approx \omega_0^2$, $\delta = \omega_0/\sqrt{2}$; при этом период внешней силы $T = 2\pi/\Omega$ равен бесконечности, а $x_0 = f_0/\omega_0^2$. По мере уменьшения коэффициента затухания δ резонансная частота внешней силы увеличивается, а при $\delta = 0$ $\Omega_{\text{рез}} = \omega_0$, $x \to \infty$.

На рис. 1.45 показан характер изменения амплитуды вынужденных колебаний в зависимости от Ω и δ при данном ω_0. Чем меньше

коэффициент затухания, тем резче изменяется амплитуда вынужденных колебаний при приближении частоты внешней силы к резонансному значению, т. е. тем более острым является максимум x_0. При $\Omega \approx 0$ амплитуда вынужденных колебаний одинакова для всех δ и равна

$$\frac{f_0}{\omega_0^2} = \frac{mf_0}{k} = \frac{F_0}{k},$$

т. е. соответствует тому отклонению, которое может вызвать в системе статическая сила F_0. Максимальную амплитуду для какого-нибудь заданного значения δ можно найти, подставив в (4.31) значение $\Omega = \Omega_{рез}$:

$$x_0 = \frac{f_0}{2\delta\sqrt{\omega_0^2 - \delta^2}} \approx \frac{f_0}{2\delta\omega_0} = \frac{f_0}{\omega_0^2}\frac{\omega_0}{2\delta} = \frac{F_0}{k}\frac{\omega_0}{2\delta}.$$

Множитель

$$D = \frac{\omega_0}{2\delta} = \frac{\omega_0 T}{2d} = \frac{\pi}{d}, \qquad (4.33)$$

показывающий, насколько максимальная амплитуда превышает величину $f_0/\omega_0^2 = F_0/k$, соответствующую $\Omega = 0$, называется *добротностью* данной колебательной системы. Чем меньше коэффициент затухания δ или логарифмический декремент затухания $d = \delta T = 2\pi\delta/\omega_0$, тем больше добротность системы и тем более острым будет максимум амплитуды.

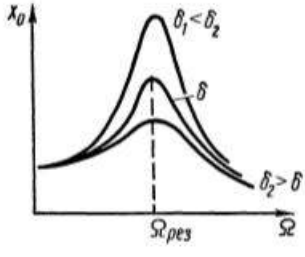


Рис. I.45

Согласно формуле для фазы (4.31), если δ очень мало, то $\omega_{рез} \approx \omega_0$ и $\varphi = -\pi/2$; тогда фаза скорости колеблющегося тела $\Omega t + \varphi + \pi/2 = \Omega t$ совпадает с фазой внешней силы. Это означает, что скорость тела и внешняя сила растут и убывают одновременно, имея всегда одинаковый знак, т. е. одинаковое направление. В этом случае внешняя сила в течение всего периода колебаний совершает только положительную работу. Если же $\Omega \neq \omega_0$ и $\varphi \neq -\pi/2$, то внешняя сила совершает в течение одной части периода положительную, а в течение другой части периода — несколько меньшую отрицательную работу. Результирующая положительная работа будет тем меньше, чем больше различие между Ω и ω_0.

Допустим, что система совершает свободные (затухающие) колебания и в некоторый момент на нее начинает действовать периодическая сила (см. выражение (4.21)). Обозначим работу сил трения через $w_{тр}$, а положительную работу внешней силы через w. Если $w_{тр} > w$, то и при наличии внешней силы затухание сохранится, но будет несколько ослаблено. Однако с уменьшением амплитуды и скоростей колеблющегося тела уменьшается и работа сил трения $w_{тр}$. Когда $w_{тр}$ становится равной w, изменение амплитуды колебаний прекращается и работа сил трения полностью компенсируется работой внешней силы. Если же в начальный период $w > w_{тр}$, то избыток $w - w_{тр}$ пройдет на увеличение энергии колебаний; амплитуды смещения x_0 и скорости v_0

будут расти, следовательно, будет увеличиваться также и $w_{тр}$. Как только $w_{тр}$ сделается равной w, дальнейшее увеличение амплитуды колебаний прекратится. Однако в некоторых случаях возможно, что непрерывно «подкачиваемая извне энергия» настолько увеличит размах колебаний в системе, что она разрушится раньше, чем наступит состояние $w_{тр} = w$.

Указанное выше возрастание амплитуды колебаний под действием внешней силы происходит особенно быстро, когда коэффициент затухания мал, а частота внешней силы равна или близка к частоте собственных колебаний ω_0; при этих условиях положительная работа внешней силы достигает (при данных F_0, k, m) наибольшего значения. Поэтому опасность разрушения колебательной системы периодическими внешними силами особенно велика при совпадении частот Ω и ω_0.

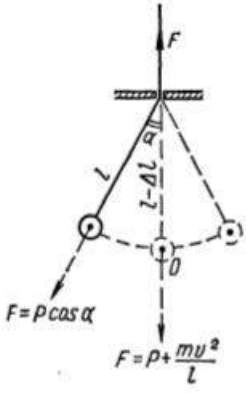


Рис. I.46

Колебания, происходящие в системе при условии $\Omega \approx \omega_0$, называются резонансными. Резонансные колебания могут привести или к установлению в системе определенного колебательного режима с большими амплитудами колебаний, или же, если потери на трение и излучение не смогут приостановить увеличения амплитуды колебаний, к разрушению системы.

Если на покоящуюся колебательную систему в некоторый момент времени начинает действовать внешняя периодическая сила, то вынужденные колебания наступают не сразу. В течение некоторого времени колебание тела носит сложный характер; происходит постепенное затухание («подавление») возникающих сначала собственных колебаний и также постепенное установление вынужденных колебаний с частотой внешней силы Ω и амплитудой x_0 (см. уравнение (4.31)). Поэтому функция (4.29) является частным решением дифференциального уравнения (4.28), соответствующим установившемуся состоянию колебательного движения в системе.

На величину амплитуды колебаний можно также воздействовать, изменяя *параметры системы,* т. е. величины, от которых зависит частота колебаний (длина маятника, коэффициенты возвращающих сил и моментов, коэффициенты трения и т. д.). Рассмотрим, например, простой маятник, длину которого можно изменять в процессе колебаний (рис. I.46). Если при прохождении тела через точку O (со скоростью v) уменьшить длину нити на Δl, то внешняя сила F, равная натяжению нити, совершит положительную работу $\Delta A_1 \approx (P + mv^2/l)\,\Delta l$. Обратное удлинение нити на Δl можно произвести в крайнем положении; тогда внешняя сила совершит меньшую отрицательную работу $\Delta A_2 = P\Delta l \cos\alpha$. Маятник получит извне энергию, равную $\Delta A_1 - \Delta A_2$. Повторяя такой процесс, можно увеличивать амплитуду колебаний тела.

Можно вызвать нарастание амплитуды колебаний также и воздействием внутренних сил, изменяющих один из параметров системы.

Допустим, например, что в системе, изображенной на рис. I.46, форма колеблющегося тела может изменяться таким образом, чтобы расстояние l от точки подвеса до центра тяжести тела увеличивалось в крайних положениях и уменьшалось при прохождении через положение равновесия (человек на качелях). При этом внутренние силы, изменяющие форму тела, будут совершать указанные выше работы ΔA_1 и ΔA_2, а результирующая положительная работа вызовет постепенное увеличение амплитуды колебаний. Заметим, что в обоих случаях начальное значение амплитуды колебаний может быть очень малым.

Подобное нарастание колебаний при изменении параметров системы (так называемое «параметрическое возбуждение») можно осуществить и в других колебательных системах; необходимо оказывать такое периодическое воздействие на тот или иной параметр системы, чтобы система могла получить энергию извне.

§ 19. ПОНЯТИЕ О НЕЛИНЕЙНЫХ КОЛЕБАНИЯХ

Рассмотренные выше колебательные движения в механических системах определяются свойствами системы и внешними воздействиями. Если в процессе колебаний параметры системы остаются постоянными, а силы, действующие на колеблющиеся тела, выражаются линейными функциями от этих параметров, то колебательные системы и сами колебания называются *линейными*. Например, системы, рассмотренные в § 17, будут линейными, если отклонения от положения равновесия настолько малы, что возвращающие силы или моменты выражаются линейными функциями $F = -kx$ и $M = -D\alpha$, а параметры системы m, k, J, D остаются в процессе колебаний постоянными. Линейными могут быть не только гармонические, но и затухающие колебания, если, например, возвращающая сила выражается линейной функцией от смещения ($F = -kx$), а сила трения — линейной функцией от скорости ($F_{тр} = -rv$); при этом опять-таки важно, чтобы коэффициент k и коэффициент трения r в течение всего процесса колебаний не изменялись. В частности, если сила трения зависит не от первой, а от второй степени скорости, как это имеет место при больших скоростях, то колебания не будут линейными.

Отметим важное свойство линейных колебательных систем: если на такую систему одновременно действуют несколько внешних периодических сил, то ввиду постоянства параметров системы каждая из этих сил оказывает свое действие независимо от существования или отсутствия других сил. Например, если одна сила при отсутствии других сил может вызвать смещение x_1, другая — x_2 и т. д., то при одновременном действии этих сил результирующее смещение равно $x_1 + x_2 + \ldots$ Этот так называемый «принцип суперпозиции» не имеет места для нелинейных систем. Действительно, допустим, одна сила вызвала смещение x_1 и при этом параметры системы несколько изменились; тогда вторая сила, которая при старых значениях параметров могла бы вызвать смещение x_2, при новых значениях параметров сможет вызвать только смещение x_2', и это смещение будет зависеть от того, насколько изменились параметры системы от действия первой силы.

Отличие линейных колебательных систем от нелинейных проявляется и в характере вынужденных колебаний: в линейных системах синусоидальная внешняя сила вызывает гармонические колебания, а в нелинейных системах эти колебания отличаются от гармонических тем сильнее, чем больше амплитуда вынужденных колебаний.

§ 20. СЛОЖЕНИЕ И РАЗЛОЖЕНИЕ КОЛЕБАНИЙ

В некоторых системах колеблющееся тело может одновременно участвовать в двух независимых колебательных движениях. Такая система показана, например, на рис. I.47, *а*, где точечная масса может колебаться в плоскости чертежа, как обычный маятник с длиной l_1, а в перпендикулярном направлении — как маятник с длиной l_2. Другая система изображена на рис. I.47, *б*, где колебания в направлениях

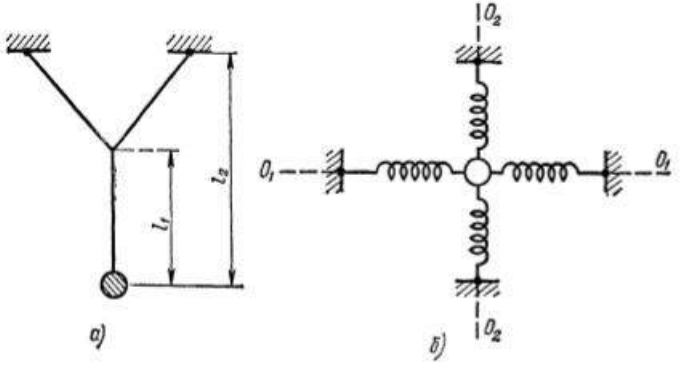


Рис. I.47

$O_1 - O_1$ и $O_2 - O_2$ происходят с различными частотами, если пружины имеют неодинаковые коэффициенты упругости. Кроме того, можно заставить тело участвовать одновременно в нескольких вынужденных колебаниях, если к этому телу приложить одновременно ряд периодических сил с различными направлениями, амплитудами и частотами.

В связи с этим возникает необходимость найти результирующее движение тела, участвующего в нескольких колебательных движениях. Ниже будут рассмотрены некоторые простые случаи сложения гармонических колебаний.

1. Сложение двух колебаний одного направления.

а) Частоты и фазы одинаковые, амплитуды различны:

$$\left.\begin{aligned} x_1 &= a\sin\omega t;\\ x_2 &= b\sin\omega t; \end{aligned}\right| \quad x = x_1 + x_2 = (a+b)\sin\omega t = x_0\sin\omega t.$$

Амплитуда результирующего колебания $x_0 = a + b$ равна сумме амплитуд складываемых колебаний.

б) Частоты и амплитуды одинаковые, фазы отличаются на φ:

$$\left.\begin{aligned} x_1 &= a\sin\omega t;\\ x_2 &= a\sin(\omega t+\varphi); \end{aligned}\right| \quad x = x_1 + x_2 = 2a\cos\frac{\varphi}{2}\sin\left(\omega t+\frac{\varphi}{2}\right) = x_0\sin\left(\omega t+\frac{\varphi}{2}\right).$$

Амплитуда результирующего колебания $x_0 = 2a\cos(\varphi/2)$ меньше суммы амплитуд складываемых колебаний; в частности, если $\varphi = \pi$, то $x_0 = 0$, т. е. два колебания с противоположными фазами взаимно друг друга «гасят» и тело будет покоиться.

в) Амплитуды одинаковые, начальные фазы равны нулю, частоты различные:

$$\left.\begin{aligned} x_1 &= a\sin\omega_1 t;\\ x_2 &= a\sin\omega_2 t; \end{aligned}\right| \quad x = x_1 + x_2 = 2a\cos\frac{\omega_1-\omega_2}{2}t\sin\frac{\omega_1+\omega_2}{2}t = 2a\cos\Omega t\sin\omega t.$$

Так как $\Omega = \frac{\omega_1-\omega_2}{2}$ меньше, чем $\omega = \frac{\omega_1+\omega_2}{2}$, то можно считать, что результирующее колебание происходит с частотой, равной полусумме частот складываемых колебаний, но амплитуда $x_0 = 2a\cos\Omega t$

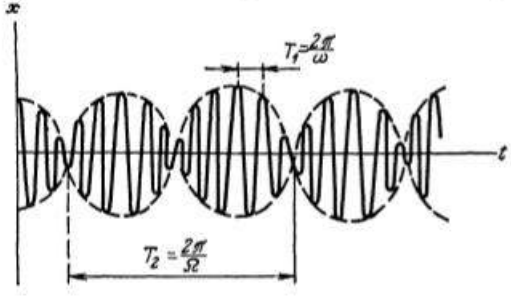


Рис. I.48

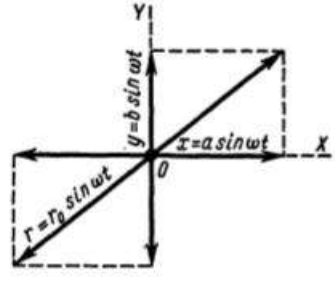


Рис. I.49

меняется со временем с частотой, равной их полуразности; график таких движений показан на рис. I.48; их называют *биениями*.

2. Сложение колебаний, происходящих в двух перпендикулярных направлениях.

Обозначим смещение тела в одном направлении через x, а в перпендикулярном направлении — через y. Уравнение траектории результирующего движения может быть представлено в виде уравнения линии на плоскости XOY (рис. I.49). Рассмотрим результирующее движение при различных соотношениях между смещениями, фазами и частотами.

а) Частоты и фазы одинаковые, амплитуды отличаются:

$$\left.\begin{aligned} x &= a\sin\omega t;\\ y &= b\sin\omega t; \end{aligned}\right| \quad y = \frac{b}{a}x.$$

Расстояние r от колеблющегося тела до точки O будет изменяться со временем по формуле

$$r = \sqrt{x^2+y^2} = \sqrt{a^2+b^2}\sin\omega t = r_0\sin\omega t.$$

Результирующее колебание происходит вдоль прямой, составляющей с осью OX некоторый угол φ, тангенс которого равен отношению b/a. Такое сложение можно наблюдать, если по телу, колеблющемуся в одном направлении, нанести удар в перпендикулярном направлении в момент, когда тело проходит точку O равновесного состояния.

б) Частоты одинаковые, фазы отличаются на $\pi/2$, амплитуды могут быть различными:

$$x = a \sin \omega t;$$
$$y = b \sin\left(\omega t + \frac{\pi}{2}\right) = b \cos \omega t; \quad \left| \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \right.$$

т. е. в результате сложения таких колебаний тело в плоскости XOY описывает эллипс, полуоси которого равны соответственно амплитудам складываемых колебаний. Если $a = b$, то тело описывает окружность. Такое сложение можно осуществить, если, например, по колеблющемуся телу маятника нанести удар в направлении, перпендикулярном его движению, в момент, когда тело находится в одном из крайних положений.

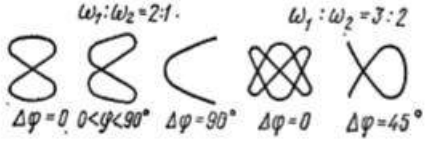


Рис. I.50

в) Частоты разные, но кратные между собой, например $\omega_1 : \omega_2 = 1 : 2;\ 2 : 3$ и т. д. В этом случае колеблющееся тело описывает сложные кривые *(фигуры Лиссажу)*, форма которых определяется отношением частот складывающихся колебаний, их амплитудами и разностью фаз между ними $\Delta\varphi = \varphi_1 - \varphi_2$. Некоторые из них приведены на рис. I.50.

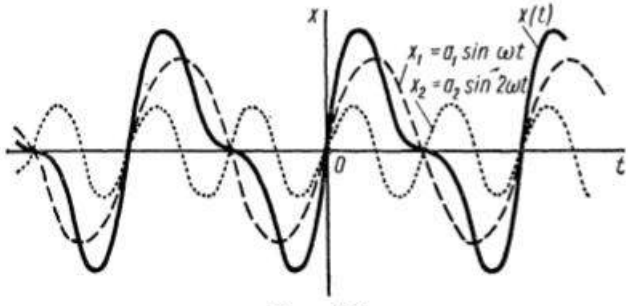


Рис. I.51

Важное значение в физике имеет сложение колебаний одинакового направления с кратными частотами ω, 2ω, 3ω и т. д. На рис. I.51 приведено сложение двух колебаний: $x_1 = a_1 \sin \omega t$ и $x_2 = a_2 \sin 2\omega t$. Период результирующего колебания (изображенного сплошной линией) оказывается равным $T = 2\pi/\omega$. К этим двум колебаниям можно добавить другие колебания с частотами, кратными ω: $x_3 = a_3 \sin 3\omega t$,

$x_4 = a_4 \sin 4\omega t$ и т. д. Форма кривой, т. е. вид функции $x = x(t)$, изображающей результирующее колебание, будет зависеть от амплитуд складываемых гармонических колебаний. Следовательно, подбирая эти амплитуды, можно получить результирующее периодическое движение с периодом $T = 2\pi/\omega$ и с самым различным характером изменения x со временем.

На рис. I.51 показано сложение колебаний, имеющих нулевые начальные фазы. В общем случае складываемые колебания с кратными частотами могут иметь не только различные амплитуды, но и различные начальные фазы, например

$$x_1 = a_1 \sin(\omega t + \varphi_1); \quad x_2 = a_2 \sin(2\omega t + \varphi_2) \text{ и т. д.}$$

Следовательно, они могут представлять собою комбинацию синусоидальных и косинусоидальных колебаний:

$$x_1 = a_1 \sin \omega t + b_1 \cos \omega t; \quad x_2 = a_2 \sin 2\omega t + b_2 \cos 2\omega t \text{ и т. д.}$$

Подбирая такие колебания с различными амплитудами и фазами, можно получить результирующее колебание с любым характером изменения x со временем. Очевидно, что будет верным и обратное утверждение: *если нам известно, что колеблющееся тело совершает сложное периодическое движение* (с периодом $T = 2\pi/\omega$), изображаемое некоторой функцией $x(t)$, *то это движение можно представить в виде суммы гармонических колебаний* с частотами ω, 2ω, 3ω, ..., амплитудами a_1, a_2, a_3, ... и фазами φ_1, φ_2, φ_3, ...:

$$x(t) = a_1 \sin(\omega t + \varphi_1) + a_2 \sin(2\omega t + \varphi_2) + \ldots = \sum a_n \sin(n\omega t + \varphi_n),$$

где $n = 1, 2, 3, \ldots$

Это утверждение является физическим применением математической теоремы Фурье о возможности разложения любой периодической функции x от некоторого параметра t в тригонометрический ряд и о способах вычисления постоянных a_n и φ_n для каждого члена этого ряда. Число членов ряда Фурье определяется видом периодической функции $x(t)$; в частности, для функции $x(t)$, изображенной на рис. I.51, ряд Фурье содержит только два члена. В некоторых случаях аналитический вид функции $x(t)$ может быть столь сложным, что заменяющий ее ряд Фурье должен содержать очень большое число членов. Если этот ряд сходится очень быстро, то в расчетах можно ограничиться только несколькими первыми членами, отбрасывая остальные, как относительно малые по величине.

Представление сложного периодического процесса в виде суммы отдельных гармонических колебаний называется *разложением этого процесса на гармонические составляющие*. Совокупность составных частей сложного периодического процесса с их частотами и амплитудами колебаний называется *спектром* этого процесса. Физически такое разложение в спектр можно осуществить и изучить при помощи, например, измерительной аппаратуры, содержащей систему, собственную частоту которой ω_0 можно плавно изменять. Если изучаемое сложное колебание $x(t)$ является внешним воздействием на эту систему, то при $\omega_0 = \omega$, 2ω, 3ω и т. д. в системе будут наблюдаться резонансные

максимумы вынужденных колебаний. Таким образом, определение спектра изучаемого колебания $x(t)$ сводится к измерению частот и амплитуд резонансных колебаний в измерительном аппарате. Наименьшая частота ω в спектре данного сложного колебания $x(t)$ называется *основной частотой*; частоты 2ω, 3ω и т. д. называются *обертонами*.

Так как энергия каждого отдельного гармонического колебания, согласно формуле (4.18), пропорциональна квадрату амплитуды этого колебания, то спектр сложного колебания $x(t)$ следует характеризовать не только набором частот составных гармонических колебаний, но и энергией, приходящейся на каждую из них.

Используя аналитически более удобный комплексный способ описания гармонических колебаний (см. формулу (4.8)), можно разложение сложного колебательного процесса в спектр представить также и в виде ряда:

$$x(t) = \sum_{-\infty}^{+\infty} A_n e^{in\omega t},$$

где n в общем случае пробегает все целочисленные значения от $-\infty$ до $+\infty$. Для некоторых функций $x(t)$ этот ряд может содержать небольшое число членов.

Оказывая воздействие на колеблющееся тело, можно вызвать изменение со временем амплитуды $x_0 = x(t)$ или частоты $\omega = \omega(t)$ колебаний по какому-нибудь закону (в частности, периодическому):

$$x = x_0(t)\sin(\omega t + \varphi); \quad x = x_0 \sin[\omega(t)\, t + \varphi].$$

Такие колебания называются *модулированными*; модуляция называется *амплитудной*, если изменяется x_0 при постоянной частоте колебаний ω, и *частотной*, если изменяется частота колебаний при постоянной амплитуде. Модуляция колебаний необходима, например, при радиопередачах; колебания, вызванные в микрофоне голосом диктора или звучанием музыкальных инструментов, изменяют амплитуду или частоту колебаний тока в антенне радиопередатчика. В приемниках производятся *демодуляция* колебаний (т. е. выделение тех воздействий, которые производили модуляцию), усиление их и обратное превращение в звуковые колебания.

Модулированные колебания также могут быть разложены в спектр, содержащий основную частоту ω и обертоны. Состав этого спектра — набор частот и распределение энергии между ними — является характеристикой данного способа модуляции.

§ 21. АВТОКОЛЕБАНИЯ

В вынужденных колебаниях подача энергии извне (для компенсации потерь на трение) осуществляется и регулируется внешней периодической силой, поэтому внешняя сила навязывает системе свою частоту колебаний и определяет амплитуду этих колебаний.

Однако можно вызывать незатухающие колебания также и постоянной силой, если только сама колеблющаяся система будет регулировать подачу энергии извне; для этого необходимо периодически прерывать или заменять действие внешней силы таким образом, чтобы совершаемая этой силой результирующая работа всегда была положительной.

Постоянная по величине и направлению внешняя сила может совершать положительную работу тогда, когда она действует в направлении движения тела. Как только колеблющееся тело изменяет направление своего движения, следует произвести «отключение» внешней силы от системы, чтобы предотвратить совершение ненужной отрицательной работы, т. е. исключить обратную передачу энергии от колеблющегося тела к внешнему источнику энергии. В некоторых системах вместо этого удается несколько уменьшить величину отрицательной работы по сравнению с положительной. В других системах вслед за изменением направления движения колеблющегося тела автоматически изменяется направление толчка, которое получает тело от внешней силы, так что за оба полупериода колебаний внешняя сила совершает только положительную работу. При помощи перечисленных выше приемов удается не только обеспечить подачу энергии извне от постоянной силы, но и отрегулировать количества получаемой энергии таким образом, чтобы в точности компенсировать потери энергии на трение и получить незатухающие колебания со стабильной амплитудой.

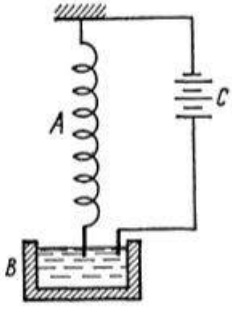


Рис. 1.52

Системы, автоматически регулирующие подачу энергии от внешнего источника, называются *автоколебательными,* а происходящие в них незатухающие периодические процессы — *автоколебаниями*. Такими системами являются часы, электрический звонок, ламповый генератор электромагнитных колебаний и т. д.

Рассмотрим колебательную систему, изображенную на рис. 1.52. Свободный конец подвешенной металлической пружины *A* опущен в сосуд с проводящей жидкостью *B*. К верхнему концу пружины и к жидкости присоединен источник тока *C*. При прохождении электрического тока через пружину между его витками действуют силы притяжения и пружина несколько сжимается. Если при этом свободный конец пружины остается погруженным в жидкость, то сжатое состояние пружины сохранится; возникшие вначале колебания пружины быстро затухнут вследствие наличия различных потерь (в частности, вследствие внутреннего трения между частицами пружины). Если же при сжатии пружины ее свободный конец выйдет из проводящей жидкости, то ток разомкнется и силы, сжимающие пружину, исчезнут. Тогда вследствие упругости пружина будет удлиняться и ее свободный конец вновь погрузится в жидкость. Таким образом, периодическая деформация пружины будет многократно повторяться. При этом потери на трение будут компенсированы источником тока, расходую-

щим энергию для возбуждения в контуре тока необходимой силы; для того чтобы возникли колебания и происходила компенсация потерь на трение, система автоматически отключает и включает внешнее воздействие. При включении тока пружина получает очередную порцию энергии для компенсации потерь.

Рассмотрим автоколебания, которые возникают при движении смычка по натянутой струне. Допустим, что смычок перемещается в одном направлении. В течение некоторого времени Δt_1 смычок и струна движутся вместе; сила f_1, действующая на струну, совершает положительную работу, деформирует струну и увеличивает ее потенциальную энергию. Совместное движение струны и смычка будет происходить до тех пор, пока упругая сила f_2, возбуждаемая в струне, не оторвет ее от смычка, после чего струна начнет двигаться в направлении, обратном движению смычка. Заметим, что это обратное движение вначале будет ускоренным, так как сила трения f_1 между смычком и струной при их относительном движении меньше, чем при их относительном покое (когда они двигались вместе вперед). Вследствие этого, как только произошел отрыв струны от смычка, упругая сила f_2 оказывается больше, чем сила трения f_1; разность $f_2 - f_1$ сообщает элементам струны ускоренное движение в обратном направлении. По мере выпрямления струны сила f_2 уменьшается и движение струны приближается к равномерному. Вследствие трения между смычком и струной (при обратном движении струны) энергия, приобретенная струной при прямом движении, постепенно превращается в тепло; в некоторый момент времени движение струны относительно смычка прекратится. Таким образом, обратное движение струны будет продолжаться некоторое время Δt_2. В дальнейшем смычок вновь увлекает струну, и процесс повторяется. Таким образом, за первую половину периода колебаний T струна получает от смычка некоторое количество энергии; за вторую половину периода эта энергия превращается в тепло и частично излучается в окружающую среду в виде звука.

В различного типа часах колеблющийся маятник приводит в движение определенные детали механизма, благодаря которым направление действия внешней силы оказывается всегда совпадающим с направлением движения маятника; внешняя сила (от сжатой пружины или опускающейся гири) совершает положительную работу за оба полупериода колебаний маятника часов.

Допустим, что за один период автоколебательная система получает от внешней силы некоторую энергию W. Если амплитуда колебаний велика, то потери на трение $W_{тр}$ (пропорциональные этой амплитуде или ее квадрату) могут быть больше, чем W, и колебания будут постепенно затухать до наступления равенства $W = W_{тр}$. Если же поступление энергии извне окажется больше, чем потери на трение, то колебания будут возрастать опять-таки до наступления равенства $W = W_{тр}$. Таким образом, автоколебательная система характеризуется не только определенной частотой колебаний, но и определенным значением их амплитуды.

Глава 5

ВОЛНОВЫЕ ПРОЦЕССЫ; ОСНОВЫ АКУСТИКИ

§ 22. ОБРАЗОВАНИЕ И РАСПРОСТРАНЕНИЕ ВОЛН В УПРУГОЙ СРЕДЕ

Если колеблющееся тело (камертон, струна, мембрана и т. д.) находится в упругой среде, то оно приводит в колебательное движение соприкасающиеся с ним частицы среды. Вследствие этого в прилегающих к телу элементах среды возникают периодические деформации (например, сжатия и растяжения). При этих деформациях в среде появляются упругие силы, стремящиеся вернуть элементы среды к первоначальным состояниям равновесия; благодаря взаимодействию соседних элементов среды упругие деформации будут передаваться от одних участков среды к другим, более удаленным от колеблющегося тела.

Таким образом, периодические деформации, вызванные в какомнибудь месте упругой среды, будут распространяться в среде с некоторой скоростью, зависящей от ее физических свойств. При этом частицы среды совершают колебательные движения около положений равновесия; от одних участков среды к другим передается только состояние деформации.

Процесс распространения колебательного движения в среде называется *волновым процессом* или просто *волной*. В зависимости от характера возникающих при этом упругих деформаций различают волны *продольные* и *поперечные*. В продольных волнах частицы среды колеблются вдоль линии, совпадающей с направлением распространения колебаний. В поперечных волнах частицы среды колеблются перпендикулярно направлению распространения волны. На рис. I.53 показано расположение частиц среды (условно изображенных в виде черточек) в продольных (*а*) и поперечных (*б*) волнах.

Жидкие и газообразные среды не имеют упругости сдвига, и поэтому в них возбуждаются только продольные волны, распространяющиеся в виде чередующихся сжатий и разрежений среды. Волны, возбуждаемые на поверхности воды, являются поперечными; они обязаны своим существованием земному тяготению. В твердых телах могут быть вызваны и продольные и поперечные волны; частным видом поперечных волн являются крутильные, возбуждаемые в упругих стержнях, к которым приложены крутильные колебания.

Предположим, что точечный источник волны начал возбуждать в среде колебания в момент времени $t = 0$; по истечении времени t это колебание распространится по различным направлениям на расстояния $r_i = v_i t$, где v_i — скорость волны в данном направлении. Поверхность, до которой доходит колебание в некоторый момент времени, называется *фронтом волны*. Форма фронта волны определяется конфигурацией источника колебаний и свойствами среды.

Среда называется *изотропной*, если ее физические свойства, в частности скорость распространения данной волны, одинаковы по всем направлениям. Среда называется *однородной*, если

ее физические свойства одинаковы в любом месте этой среды. Среда может быть однородной, но неизотропной, или неоднородной, но изотропной. Фронт волны от точечного источника колебаний в однородной и изотропной среде имеет вид сферы; такие волны называются *сферическими*.

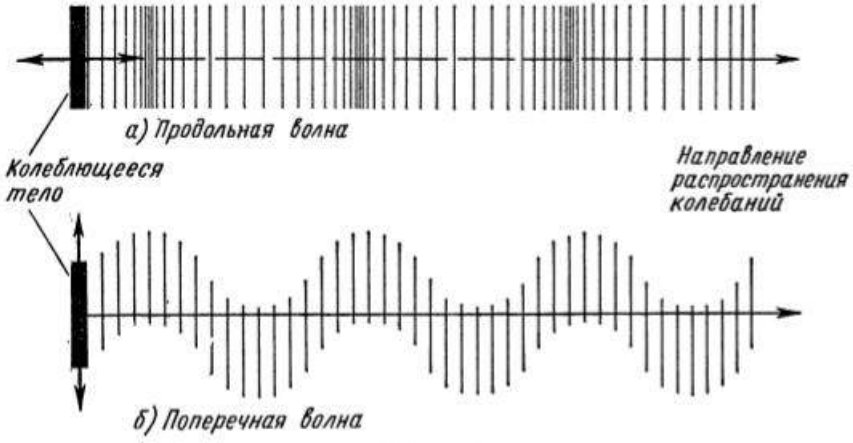


Рис. I.53

В неоднородной и неизотропной (*анизотропной*) среде, а также от неточечных источников колебаний фронт волны имеет сложную форму. Если фронт волны представляет собой плоскость и эта форма сохраняется по мере распространения колебаний в среде, то волну называют *плоской*. Малые участки фронта волны сложной формы

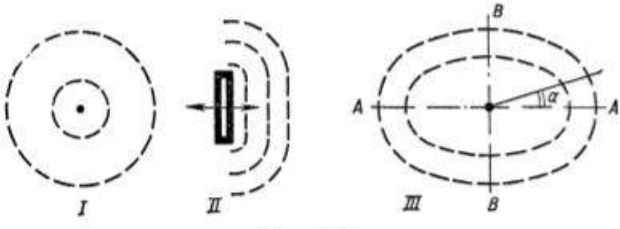


Рис. I.54

можно считать плоской волной (если только рассматривать небольшие расстояния, проходимые этой волной).

При описании волновых процессов выделяют поверхности, в которых все частицы колеблются в одинаковой фазе; эти «поверхности одинаковой фазы» называются *волновыми* или *фазовыми*. Они также могут быть сферическими, плоскими или иметь сложную форму в зависимости от конфигурации источника колебаний и свойств среды. На рис. I.54 условно показаны: *I* — шаровая волна от точечного источника, *II* — волна от колеблющейся пластинки, *III* — эллиптическая волна

от точечного источника в анизотропной среде, в которой скорость распространения волны v_l плавно изменяется по мере возрастания угла α, достигая максимума вдоль направления $A—A$ и минимума — вдоль $B—B$.

§ 23. ФОРМУЛА ГАРМОНИЧЕСКОЙ ВОЛНЫ

При описании волнового процесса требуется найти амплитуды и фазы колебательного движения в различных точках среды и изменение этих величин с течением времени. Эта задача может быть решена, если известно, по какому закону колеблется и как взаимодействует со средой тело, вызвавшее волновой процесс. Однако во многих случаях не существенно, каким телом возбуждена данная волна; решается более простая задача: дано состояние колебательного движения в одних точках среды в определенный момент времени, например известно расположение фронта волны или волновой поверхности; требуется определить состояние колебательного движения в других точках среды. Эта задача будет рассматриваться в § 25. Здесь же мы найдем связь между состояниями колебательного движения в различных точках среды в простейшем случае, когда в этой среде распространяется плоская или сферическая синусоидальная волна.

ВЫВОД ФОРМУЛЫ ВОЛНЫ

Допустим, что волновой процесс распространяется в положительном направлении оси OX, т. е. в сторону возрастания координаты x (рис. I.55). Обозначим через y колеблющуюся величину; этой величиной могут быть: смещение частиц среды относительно их положения

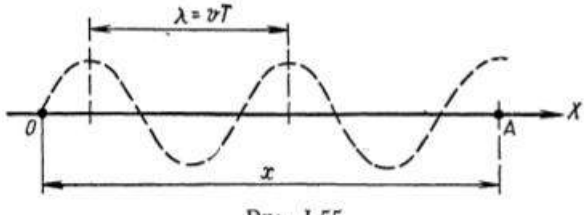


Рис. I.55

равновесия; отклонение давления или плотности в данном месте среды от равновесного значения и т. д. Для простоты рассуждений предположим, что распространяющаяся волна — синусоидальная, т. е. в каждой точке среды величина y изменяется со временем по гармоническому закону. Допустим, что начало отсчета времени выбрано так, что в точке O при $t = 0$ $y = 0$, т. е. $\varphi = 0$; тогда

$$y = y_0 \sin \omega t, \tag{5.1}$$

где $\omega = 2\pi/T$ — угловая частота; T — период; y_0 — амплитуда колебаний; ωt — аргумент синуса (определяющий значение колеблющейся величины в каждый заданный момент времени) — фаза колеба-

ний в точке O. Требуется найти фазу колебаний в любой другой точке A, отстоящей от O на расстоянии x.

Так как точка A расположена относительно O в направлении распространения волны, то в данный момент времени t в этой точке будет такое состояние колебательного движения, какое было в точке O на x/v секунд раньше; здесь v есть скорость распространения фазы колебаний в направлении OX. Таким образом, фаза колебаний в точке A в момент t равна фазе колебаний в точке O в более ранний момент $t - x/v$, т. е. равна $\omega(t - x/v)$.

Следовательно, значение колеблющейся величины в точке A в момент времени t

$$y = y_0 \sin \omega (t - x/v). \tag{5.2}$$

Мы получили формулу (уравнение) *синусоидальной волны*; величина v называется *фазовой скоростью* волны.

Допустим теперь, что волна распространяется в обратном направлении, т. е. от A к O, в сторону убывания координаты x. Тогда определенное состояние колебания, т. е. определенная фаза волны, достигает точки A на $\tau = x/v$ секунд раньше, чем точки O, следовательно, фаза в точке A в данный момент времени больше фазы в точке O на $\omega\tau = \omega x/v$. Если по-прежнему принять фазу в точке O в момент t равной ωt, то в точке A в этот же момент времени фаза будет равна $\omega t + \omega\tau = \omega(t + x/v)$. Таким образом, формулу синусоидальной волны можно написать в общем виде:

$$y = y_0 \sin \omega (t \pm x/v), \tag{5.3}$$

где знак минус берется для волны, распространяющейся в направлении возрастания x, а плюс — в обратном направлении.

При выводе формулы (5.3) предполагалось, что амплитуда колебаний y_0 по мере распространения волны не изменяется и среда однородна (т. е. скорость распространения фазы колебаний везде одинаковая). Эти два предположения означают, что мы рассматривали *плоскую* волну. У сферической волны амплитуда колебаний уменьшается обратно пропорционально расстоянию и формула волны имеет вид

$$y = \frac{y_0}{x} \sin \omega (t - x/v). \tag{5.4}$$

Расстояние λ, пройденное волной (т. е. определенной фазой колебаний) за один период колебаний, называется *длиной волны*; очевидно,

$$\lambda = vT = v(2\pi/\omega) = v/\nu; \quad v = \lambda/T = \lambda\nu.$$

В формуле волны (5.3) колеблющаяся величина зависит от двух переменных: x и t. Если найти производную от y по времени, полагая x постоянной, то эта *частная производная*

$$\lim_{\Delta t \to 0} \left(\frac{\Delta y}{\Delta t}\right)_{x = \text{const}} = \frac{\partial y}{\partial t}$$

показывает скорость изменения колеблющейся величины в данной точке среды. Производная же от y по x при постоянном t

$$\lim_{\Delta x \to 0} \left(\frac{\Delta y}{\Delta x}\right)_{t=\text{const}} = \frac{\partial y}{\partial x}$$

есть разность значений колеблющейся величины, рассчитанная на единицу расстояния между точками среды ($\Delta x = x_2 - x_1$), т. е. показывает, как увеличивается или уменьшается y вдоль оси OX (в данный момент времени t).

Найдем частные производные от колеблющейся величины y по времени при постоянном x:

$$u = \frac{\partial y}{\partial t} = y_0 \omega \cos \omega (t - x/v), \tag{5.5}$$

$$a = \frac{\partial^2 y}{\partial t^2} = -\, y_0 \omega^2 \sin \omega (t - x/v) = -\, \omega^2 y.$$

Если y есть смещение частиц среды при колебаниях, то u и a будут скоростью и ускорением этих частиц при их колебательном движении в точке с координатой x. Амплитудные значения этих величин связаны между собой:

$$u_0 = y_0 \omega; \quad a_0 = y_0 \omega^2 = u_0 \omega.$$

Частные производные от y по x при постоянном t будут равны:

$$\frac{\partial y}{\partial x} = y_0 \left(-\frac{\omega}{v}\right) \cos \omega \left(t - \frac{x}{v}\right),$$

$$\frac{\partial^2 y}{\partial x^2} = -\, y_0 \frac{\omega^2}{v^2} \sin \omega \left(t - \frac{x}{v}\right) = -\, \frac{\omega^2 y}{v^2}.$$

Следовательно,

$$\frac{\partial^2 y}{\partial t^2} = v^2 \frac{\partial^2 y}{\partial x^2}. \tag{5.6}$$

Это есть *дифференциальное уравнение плоской гармонической волны*, распространяющейся по оси OX. Оно получено нами из формулы волны (5.3). Однако можно сделать и обратное заключение: если какая-нибудь физическая величина $y = y(x, t)$ зависит от времени и координат так, что ее частные производные удовлетворяют уравнению (5.6), то величина y распространяется в среде в виде плоской волны (см. уравнение (5.3)) со скоростью

$$v = \sqrt{\frac{\partial^2 y}{\partial t^2} \Big/ \frac{\partial^2 y}{\partial x^2}}$$

и частотой колебаний

$$\omega = \sqrt{-\frac{1}{y}\frac{\partial^2 y}{\partial t^2}} = v \sqrt{-\frac{1}{y}\frac{\partial^2 y}{\partial x^2}}.$$

Дифференциальное уравнение плоской гармонической волны, распространяющейся вдоль оси OX, можно записать не только в виде (5.6), содержащем вторые производные по координате и времени, но и в виде

связи между этими производными и самой колеблющейся величиной:

$$\frac{\partial^2 y}{\partial t^2} + \omega^2 y = 0; \quad \frac{\partial^2 y}{\partial x^2} + \frac{\omega^2}{v^2} y = 0. \tag{5.7}$$

Второе уравнение явно не содержит времени; если произвести замену $\omega = 2\pi/T$, $vT = \lambda$ (длина волны), то

$$\frac{\partial^2 y}{\partial x^2} + \frac{4\pi^2}{\lambda^2} y = 0. \tag{5.8}$$

ВОЛНА В УПРУГОЙ СРЕДЕ

При распространении в среде упругих (в частности, звуковых) колебаний частицы среды совершают колебательное движение относительно своих положений равновесия. Можно было бы описывать волновое движение, учитывая только смещения и скорости частиц среды. Однако при наличии беспорядочного теплового движения частиц пользоваться таким описанием неудобно. Поэтому принято упругую (в частности, звуковую) *волну характеризовать периодическими изменениями давления и плотности*, которые происходят при последовательных сжатиях и растяжениях (расширениях) среды. Обозначим, например, давление и плотность воздуха в равновесном состоянии через p_0 и ρ_0, а их мгновенные значения в данном месте — через p и ρ. Тогда для описания звуковой волны в воздухе можно интересоваться периодическими изменениями избыточного давления $\Delta p = p - p_0$ или избыточной плотности $\Delta\rho = \rho - \rho_0$.

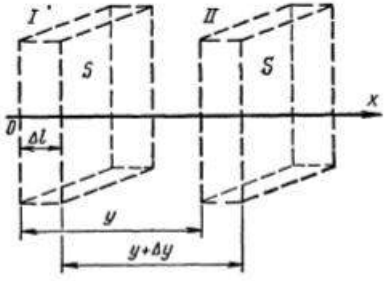


Рис. 1.56

Выясним, при каких условиях в упругих средах возможны гармонические волны вида (5.3). Выделим перпендикулярно OX некоторую площадку S (рис. 1.56) и слой малой толщины Δl. Допустим, что в положении I избыточное давление слева равно Δp_1, а справа $\Delta p_2 = \Delta p_1 + \frac{\partial(\Delta p)}{\partial x}\Delta l$, следовательно, на выделенный элемент среды будет действовать результирующая сила $\Delta F = S(\Delta p_1 - \Delta p_2) = -S\frac{\partial(\Delta p)}{\partial x}\Delta l$. Масса этого элемента $\Delta m = \rho S \Delta l$, где ρ — средняя плотность среды в объеме элемента. Тогда, согласно второму закону механики, рассматриваемый элемент среды будет иметь ускорение

$$a = \frac{\Delta F}{\Delta m} = -\frac{1}{\rho}\frac{\partial(\Delta p)}{\partial x}$$

(знак минус означает, что если избыточное давление Δp в положительном направлении x возрастает, то сила ΔF и ускорение a будут направлены в обратную сторону).

Обозначим смещение частиц среды в направлении распространения упругой волны через y. Так как смещение y зависит от двух переменных: времени и координаты, то ускорение элемента запишем в виде частной производной $a = \frac{\partial^2 y}{\partial t^2}$; тогда

$$\frac{\partial^2 y}{\partial t^2} = -\frac{1}{\rho}\frac{\partial (\Delta p)}{\partial x}. \tag{5.9}$$

Исследуем правую часть этой формулы. Если бы все частицы среды, находящиеся в рассматриваемом элементе, имели одинаковое смещение y, то объем элемента, следовательно, и давление p и плотность ρ внутри него оставались бы постоянными. В этом случае правая часть уравнения (5.9) будет равна нулю и упругой волны в среде не будет. Поэтому необходимо допустить, что при переходе из положения *I* в *II* одна грань рассматриваемого элемента среды смещается на y, а другая — на $y + \Delta y$. При таком перемещении объем элемента изменится, вследствие чего давление p станет функцией от координаты x и правая часть уравнения (5.9) будет отлична от нуля.

Однако в формуле (5.9) имеются две переменные величины y и p; если исключить одну из них, например давление p, то получим дифференциальное уравнение для смещения элементов среды от положения равновесия. Для этой цели сначала учтем, что величину Δy следует полагать пропорциональной толщине элемента среды Δl:

$$\Delta y = \frac{\partial y}{\partial x}\Delta l;$$

где $\partial y/\partial x$ показывает, какое изменение смещения y приходится на единицу длины вдоль оси OX.

Тогда относительное изменение объема элемента будет равно

$$\varepsilon = \frac{S(y+\Delta y) - Sy}{S\,\Delta l} = \frac{\partial y}{\partial x}.$$

Масса среды в объеме элемента не изменяется, поэтому относительное увеличение плотности будет равно относительному уменьшению объема элемента, т. е.

$$\frac{\Delta\rho}{\rho} = -\varepsilon = -\frac{\partial y}{\partial x}.$$

Теперь для того чтобы рассчитать изменение избыточного давления Δp внутри элемента, необходимо знать зависимость Δp от ρ или ε.

Если среда — твердое тело, то при малых деформациях можно воспользоваться законом Гука $p = \varepsilon E$. Для плоской волны ($S = \text{const}$) относительное удлинение или сжатие элемента объема будет совпадать с относительным изменением его объема; напряжение сжатия или растяжения можно полагать равным среднему значению Δp внутри элемента, причем увеличение Δp сопровождается уменьшением объема элемента, поэтому

$$\Delta p = -\varepsilon E = -E\frac{\partial y}{\partial x};\quad \frac{\partial (\Delta p)}{\partial x} = -E\frac{\partial^2 y}{\partial x^2}.$$

Подставив последнее соотношение в формулу (5.9) получим дифференциальное уравнение плоской волны, распространяющейся в *твердых телах*:

$$\frac{\partial^2 y}{\partial t^2} = \frac{E}{\rho}\frac{\partial^2 y}{\partial x^2}. \tag{5.10}$$

Сравнивая уравнения (5.10) и (5.6), замечаем, что величину E/ρ следует отождествить с квадратом скорости распространения волны, поэтому сама скорость волны оказывается равной

$$v = \sqrt{E/\rho}. \tag{5.11}$$

Для железа $E = 2 \cdot 10^{11}$ Н/м², $\rho = 7800$ кг/м³ и $v \approx 5100$ м/с.

В *газах* процессы сжатия и расширения описываются уравнением

$$pV^{\gamma} = \text{const}$$

(см. ч. II, § 10), где p — давление; V — удельный объем газа, а γ — некоторая постоянная величина, зависящая от того, как происходят процессы сжатия и расширения. Из этого уравнения следует

$$\frac{dp}{p} = -\gamma\frac{dV}{V} = -\gamma\varepsilon.$$

Если избыточное давление Δp мало по сравнению с давлением газа p_0, то

$$\Delta p \approx -p_0\gamma\varepsilon; \quad \frac{\partial(\Delta p)}{\partial x} = -p_0\gamma\frac{\partial \varepsilon}{\partial x} = -p_0\gamma\frac{\partial^2 y}{\partial x^2}.$$

Подставив это выражение для $\frac{\partial(\Delta p)}{\partial x}$ в формулу (5.9), вновь получим дифференциальное уравнение (5.6) плоской волны, причем скорость распространения оказывается равной (полагая $\rho \approx \rho_0$)

$$v = \sqrt{\gamma\frac{p_0}{\rho_0}}. \tag{5.12}$$

Дифференциальное уравнение плоской волны и формулы (5.11) и (5.12) для скоростей распространения получены при предположении, что избыточные давления Δp и плотности $\Delta\rho$ малы. Найдем изменение этих величин со временем; для любой среды, полагая $\Delta\rho \approx -\rho_0\varepsilon$, получим для плотности

$$\Delta\rho \approx -\rho_0\frac{\partial y}{\partial x} = \rho_0 y_0\frac{\omega}{v}\cos\omega\left(t - \frac{x}{v}\right) = \Delta\rho_0\cos\omega\left(t - \frac{x}{v}\right),$$

где через $\Delta\rho_0$ обозначена амплитуда колебаний плотности среды в волне:

$$\Delta\rho_0 = \rho_0 y_0\frac{\omega}{v} = \frac{\rho_0 u_0}{v}.$$

Для колебаний давления Δp также получаются формулы, одинаковые для всех сред:

$$\Delta p \approx \Delta p_0\cos\omega(t - x/v); \quad \Delta p_0 = \rho_0 u_0 v. \tag{5.13}$$

Таким образом, Δp и $\Delta \rho$ пропорциональны не смещениям частиц среды y, а их скоростям u. Из этих уравнений можно получить общее выражение для скорости распространения плоской волны в упругой среде:

$$v = \sqrt{\frac{\Delta p}{\Delta \rho}}, \text{ или } v = \sqrt{\frac{dp}{d\rho}}.$$

§ 24. ПОТОК ЭНЕРГИИ В ВОЛНОВЫХ ПРОЦЕССАХ

Процесс распространения волны в каком-нибудь направлении в среде сопровождается переносом энергии колебаний в этом направлении. Допустим, что S есть часть фронта плоской волны в некоторый момент t (рис. I.57). По истечении времени Δt фронт волны переместится на расстояние $\Delta l = v\Delta t$, вследствие чего частицы среды в объеме $S\Delta l$ приводятся в колебательное движение. Обозначим через w энергию колебаний частиц среды, содержащихся в единице объема; допустим, что объем $S\Delta l$ очень мал и поэтому в пределах этого объема энергия w везде одинакова. Можно утверждать, что за время Δt среда, заключенная в объеме $S\Delta l$, получила через площадку S энергию $wS\Delta l$. Таким образом, за единицу времени через площадку S прошла энергия

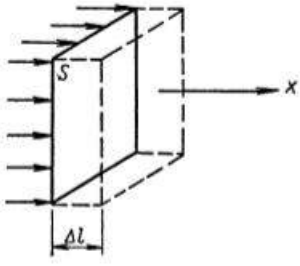


Рис. I.57

$$\Phi = \frac{wS\,\Delta l}{\Delta t} = wSv.$$

Величина Φ есть *поток энергии волны* (в единицу времени) через площадку S (S ориентируют перпендикулярно направлению распространения волны). *Плотностью потока энергии* (*интенсивностью волны*) называют энергию, проходящую в единицу времени через единицу площадки, перпендикулярной направлению распространения волны:

$$I = \Phi/S = wv. \qquad (5.14)$$

Так как скорость распространения волны v есть вектор, то и плотность потока энергии есть векторная величина (*вектор Умова*). Поток энергии выражается в Вт, плотность потока — в Вт/м².

В сферической волне, вызванной точечным источником колебаний, плотность потока энергии убывает обратно пропорционально квадрату расстояния от источника колебаний. Для доказательства допустим, что источник колебаний ежесекундно отдает в окружающую среду одну и ту же энергию, равную W. Эта энергия равномерно распределяется по шаровой поверхности фронта волны $S = 4\pi R^2$, поэтому через единицу площади этой поверхности в единицу времени проходит энергия $I = W/4\pi R^2$, т. е. $I \sim 1/R^2$. Это соотношение, так же как и формула (5.14), применимо, если только в среде не происходит превращения энергии колебательного движения в другие виды энергии, например в теплоту.

Найдем формулу для расчета w. Рассмотрим некоторый объем среды, размеры которого настолько малы по сравнению с длиной волны λ, что в пределах этого объема фазы и амплитуды колебаний для всех частиц можно считать одинаковыми. Обозначим массу одной частицы через m, число частиц в единице объема — через n. Тогда полная энергия колебаний (сумма кинетической и потенциальной энергии) частиц, содержащихся в единице объема, согласно формуле (4.18), будет равна:

$$w = n\frac{mu_0^2}{2} = \frac{\rho u_0^2}{2} = \frac{\rho y_0^2\omega^2}{2},$$
$$I = \frac{\rho v y_0^2\omega^2}{2} = \frac{\rho v u_0^2}{2}, \qquad (5.15)$$

где $\rho = nm$ — плотность среды; y_0 — амплитуда смещения; $u_0 = y_0\omega$ — амплитуда скорости; ω — угловая частота колебаний частиц среды. Для сферической волны плотность потока энергии убывает обратно пропорционально квадрату расстояния от центра колебаний, поэтому амплитуда колебаний y_0 убывает обратно пропорционально первой степени этого расстояния.

Заметим, что максимумы избыточных давлений и плотностей в волне совпадают не с максимумами смещений y, а с максимумами скоростей частиц среды $u = dy/dt$. Поэтому потенциальная энергия сжатых и растянутых участков среды изменяется в фазе с кинетической энергией колеблющихся частиц среды, т. е. вдоль волны оба эти вида энергии (потенциальная и кинетическая) одновременно приобретают наибольшие и наименьшие значения.

При распространении волны в среде происходит *поглощение энергии*, т. е. превращение механической энергии колебаний в беспорядочное (тепловое) движение частиц среды.

Допустим, что интенсивность плоской волны, распространяющейся по оси OX, при $x = 0$ равна I_0. Полагая, что в слое толщиной dx количество поглощенной энергии dI на единице площади пропорционально величине I и толщине слоя, получим:

$$dI = -\beta I\,dx; \quad I = I_0 e^{-\beta x}.$$

Величина β называется *коэффициентом поглощения* энергии волны в данной среде. Так как энергия волны пропорциональна квадрату амплитуды (смещения, скорости, избыточного давления или плотности), то при прохождении слоя среды толщиной x эти амплитуды будут убывать по закону

$$y_{0x} = y_0 e^{-\alpha x}.$$

Для избыточного давления в среде

$$\Delta p = \Delta p_0 e^{-\alpha x}.$$

Величина α (равная $\beta/2$) называется *коэффициентом затухания* колебаний в среде. Для воздуха и некоторых других газов он пропорционален квадрату частоты колебаний; $\alpha \approx 2\cdot10^{-11}\nu^2$ с^2/м; для воды (до $\nu = 10^6$ Гц) $\alpha \approx 3\cdot10^{-14}\nu^2$ с^2/м, т. е. в 700 раз меньше, чем в воз-

духе. Например, амплитуда ультразвуковых колебаний с частотой $\nu = 10^6$ Гц уменьшается в $e = 2{,}71$ раза в воздухе на расстоянии $1/\alpha = 0{,}05$ м, а воде — на расстоянии 33,4 м. Этим объясняется широкое применение звуковых и ультразвуковых волн при исследовании морей, при гидролокации и т. д.

§ 25. ИНТЕРФЕРЕНЦИЯ И ДИФРАКЦИЯ ВОЛН. РАСПРОСТРАНЕНИЕ ВОЛН В СРЕДАХ С ДИСПЕРСИЕЙ

ПРИНЦИП ГЮЙГЕНСА. ПРИНЦИП СУПЕРПОЗИЦИИ

При изучении волновых процессов в самом общем случае ставится следующая задача: дано тело определенной формы и размеров, совершающее колебания в некоторой среде, свойства которой известны; даны амплитуды, фазы и направления колебаний всех точек тела; требуется определить форму и расположение фронта волны или волновой поверхности в среде в каждый интересующий нас момент времени, а также амплитуды колебаний в различных точках этих поверхностей.

Для этой цели сначала рассматривается взаимодействие колеблющегося тела со средой и определяется фронт волны в непосредственной близости от этого тела. Дальнейшее распространение волны в среде определяется при помощи принципа Гюйгенса и принципа суперпозиции.

В основе принципа (или правила) Гюйгенса лежит утверждение: *все точки волнового фронта, заданного в некоторый момент времени* t_0, *можно рассматривать как самостоятельные источники волны, начавшие излучать в момент* t_0 (*эти волны называются элементарными или вторичными*).

Для нахождения фронта волны в последующий момент времени $t = t_0 + \Delta t$ *необходимо найти фронты всех элементарных волн и затем провести их огибающую.*

Элементарные волны могут быть сферическими (если среда изотропная), эллипсоидальными или иными в зависимости от свойств среды. Таким образом, при помощи принципа Гюйгенса можно чисто геометрическим построением найти расположение фронта волны в последующие моменты времени, если это расположение задано в момент t_0. Однако принцип Гюйгенса не определяет амплитуды колебаний в различных точках, куда доходит волна. Для этого дополнительно используется *принцип суперпозиции*, содержание которого заключается в следующем.

Допустим, что S — фронт волны (или волновая поверхность) в момент t_0, а S' — в момент $t = t_0 + \Delta t$ (рис. I.58, *а*). Элементарные волны, исходящие из различных точек M_1, M_2, ... поверхности S, имеют в этих точках одинаковые фазы (амплитуды их могут отличаться). В момент t эти волны дойдут до точки A с фазами и амплитудами, зависящими от расстояний r_1, r_2, ... Принцип суперпозиции утверждает, что *амплитуду колебаний в точке A можно найти, сложив все колеба-*

ния, вызванные в этой точке. При этом обычно предполагается, что колебания в среде *линейные* и поэтому в точке A суммарное отклонение от состояния равновесия y равно сумме отклонений $y_1 + y_2 + \ldots$, вызванных каждой волной в отдельности.

Допустим, что в данной среде имеются несколько точечных источников волны (одинаковой частоты). Рассмотрим две точки среды: A, в которой смещения, вызванные этими волнами, имеют одинаковое направление и поэтому складываются, и B, в которой смещения имеют противоположные направления и вычитаются. Так как энергия колебаний в единице объема среды пропорциональна квадрату амплитуды, то вблизи точки A плотность энергии колебаний значительно больше,

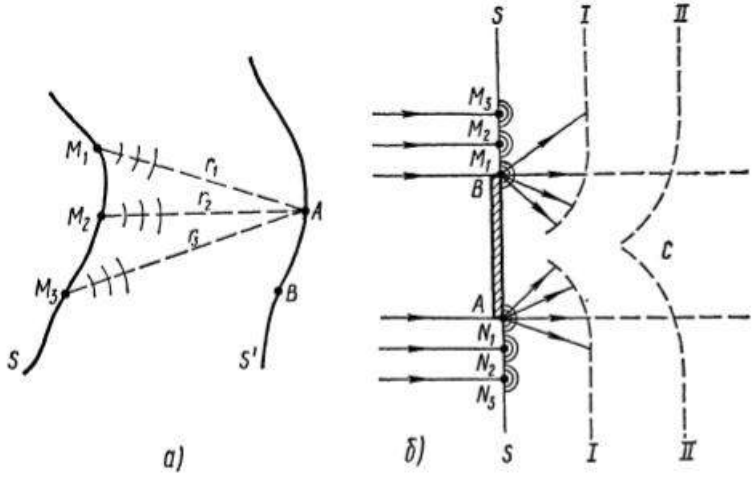


Рис. I.58

чем около точки B, т. е. распределение энергии в среде будет неравномерным. Это распределение, если оно со временем не изменяется и поэтому может быть обнаружено и изучено, называется *интерференционной картиной.* Образование интерференционной картины (*интерференция*) возможно, если волны имеют одинаковые частоты и приходят в данную точку среды с постоянной разностью фаз (не изменяющейся со временем); такие волны называются *когерентными.* Колеблющиеся тела, вызывающие в среде когерентные волны, называются *когерентными источниками.* Например (рис. I.58, *а*), точки M_1, M_2, ... поверхности S колеблются с одинаковыми частотами и фазами, а исходящие от них элементарные волны приходят в точку A с разностями фаз, которые зависят только от расстояний r_1, r_2, ... и не изменяются со временем; поэтому все точки волновой поверхности являются когерентными источниками, а испускаемые ими элементарные волны — когерентными волнами.

При помощи принципа Гюйгенса и принципа суперпозиции можно решить ряд важных задач по распространению волн в неоднородной и неизотропной среде. В частности, рассмотрим распространение

волны в среде, в которой имеются тела, не пропускающие этой волны, поглощающие или отражающие ее. Допустим, плоская волна S, имеющая во всех своих точках одинаковую амплитуду колебаний, встречает «непрозрачное» для этой волны тело P, не пропускающее участок волны AB (рис. I.58,*б*). Применим принцип Гюйгенса для момента времени, когда волновой фронт находится в S-S. Построив фронты элементарных волн, исходящих из точек M_1, M_2, ..., N_1, N_2, ..., и проведя огибающие I, II и т. д., заметим, что волновой фронт заходит в область C, которую называют *областью геометрической тени* (из геометрических соображений можно предположить, что волна, отрезаемая телом, не должна заходить в эту область).

Это проникновение волны в область геометрической тени называется *дифракцией волны*. Применяя принцип суперпозиции, можно установить, что вдоль нового фронта волны I или II амплитуда колебаний (или плотность потока энергии) уже не везде одинакова, как это было задано для волнового фронта S. Подобная задача рассматривается также в разделе оптики для световых волн; здесь же заметим, что при дифракции волны происходит некоторое перераспределение энергии в пределах фронта волны, причем это перераспределение особенно заметно вблизи границ непрозрачных тел.

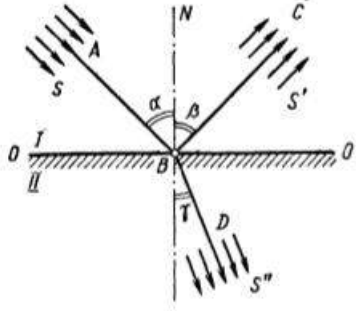


Рис. I.59

При помощи принципа Гюйгенса объясняются также законы отражения и преломления волн на границе раздела двух сред. Соответствующие построения для световых волн приводятся в разделе оптики; они применимы и для упругих волн в механических средах. Допустим, что на плоскую границу раздела 0—0 двух однородных и изотропных сред I и II, скорости распространения колебаний в которых равны v_1 и v_2, падает плоская волна S, причем нормаль AB к фронту этой волны составляет с нормалью N к поверхности раздела угол α (рис. I.59). Тогда построение Гюйгенса в согласии с опытными данными дает следующие результаты:

1) нормаль BC к фронту S' отраженной волны составляет с N угол β, равный α;

2) нормаль BD к фронту S'' преломленной волны составляет с N угол γ, удовлетворяющий соотношению

$$\frac{\sin\alpha}{\sin\gamma} = \frac{v_1}{v_2}.$$

Заметим, что при переходе волны из одной среды в другую частота колебаний ν сохраняется, а длина волны изменяется в зависимости от скорости распространения волны:

$$\lambda_1 = v_1/\nu; \quad \lambda_2 = v_2/\nu; \quad \lambda_1/\lambda_2 = v_1/v_2.$$

Выше предполагалось, что волны, распространяющиеся в данной среде, являются плоскими и гармоническими, т. е. в каждой точке среды колеблющаяся величина (которую теперь будем обозначать через ψ) изменяется со временем по синусоидальному (или косинусоидальному) закону, а в каждый определенный момент времени эта величина распределена вдоль линии распространения (например, оси OX) по тому же закону. Формула такой волны имеет вид

$$\psi = \psi_0 \sin \omega (t - x/v),$$

где v — скорость распространения волны в среде (фазовая): она показывает скорость перемещения определенной фазы колебаний в данной среде. Допустим, что нас интересует некоторое значение фазы $\alpha = \omega (t - x/v)$, например $\alpha = \pi/2$. Полагая $\alpha = \text{const}$ и дифференцируя по времени, получим $v = dx/dt$. Очевидно, что любое значение фазы волны распространяется с этой же скоростью.

ДИСПЕРСИЯ

Величина v определяется физическими свойствами среды; в некоторых средах она различна для различных частот колебаний (ω или ν). Зависимость скорости распространения волны от частоты колебаний называется *дисперсией*. Дисперсия имеет важное значение при распространении в среде сложного периодического процесса, содержащего гармонические волны различных частот колебаний.

Допустим, например, что по камертону, имеющему основную частоту колебаний ν_0, наносятся одинаковые удары молотком через строго определенные промежутки времени T. Тогда по воздуху вдоль направления оси OX будет распространяться периодический процесс, описываемый, например, некоторой функцией $\Psi(x, t)$. Эта функция должна зависеть от спектра камертона, т. е. от набора частот, которые он излучает, и от распределения энергии между этими частотами (т. е. между гармоническими составляющими спектра). Кроме того, эта функция должна содержать частоту $\nu = 1/T$, с которой наносятся удары, а также должна зависеть от характера самого удара (продолжительности удара, изменения силы удара со временем) и от быстроты затухания колебаний камертона вследствие потерь на трение и на излучение. На рис. 1.60 схематично показаны значения $\Psi(x, t)$, соответствующие двум последовательным ударам молотка. Условимся каждую фигуру, охваченную на этом рисунке пунктирной линией (содержащую только большие отклонения от положения равновесия), называть *волновым импульсом* или *формой волны* Ψ (ее называют также *волновым пакетом*). За время T волновой импульс пройдет расстояние l и, следовательно, скорость распространения этого импульса будет равна

$$u = l/T.$$

Для описания процесса распространения звука, возбужденного таким образом, с учетом физических свойств среды (и, в частности, дисперсии) целесообразно представить функцию $\Psi(x, t)$ в виде суммы

гармонических слагаемых $\psi_i(x, t)$, имеющих различные частоты колебаний ω_i:

$$\Psi(x, t) = \sum \psi_i(x, t).$$

Если среда не обладает дисперсией, то все гармонические волны $\psi_i(x, t)$ будут иметь одинаковые фазовые скорости: $v_i = v$. Вследствие этого весь волновой импульс (как «сумма» этих волн) будет перемещаться в среде с той же скоростью ($u = v$), поэтому «форма волны» по мере ее распространения изменяться не будет. Если же среда обладает дисперсией, т. е. фазовые скорости гармонических волн $\psi_i(x, t)$ не одинаковы и зависят от частоты ω или длины волны λ, то «форма» волнового импульса будет по мере распространения в среде изменяться.

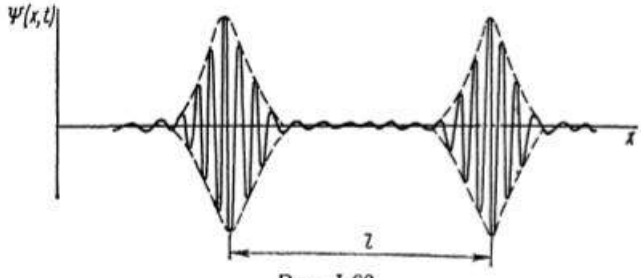


Рис. 1.60

В том случае, когда форма волны (т. е. форма кривой на рис. 1.60, изображающей волну Ψ) при распространении изменяется не очень сильно (хотя бы на малых расстояниях от источника), скорость волнового импульса (пакета) можно определить по перемещению l точки с наибольшим отклонением, показанному на рис. 1.60. Определенная таким образом скорость u называется групповой скоростью волнового импульса (пакета). Существует связь между фазовой скоростью $v(\lambda)$ какой-нибудь из гармонических составляющих импульса и групповой скоростью самого импульса

$$v - u = \frac{dv}{d\lambda},$$

где $dv/d\lambda$ — производная от фазовой скорости по длине волны — характеризует дисперсионные свойства данной среды. Фазовые скорости v гармонических составляющих любого волнового импульса (пакета) в диспергирующей среде могут в зависимости от знака производной $dv/d\lambda$ быть и больше и меньше групповой скорости u самого импульса (пакета).

§ 26. СТОЯЧИЕ ВОЛНЫ

Рассмотрим результат интерференции двух синусоидальных плоских волн одинаковой амплитуды и частоты, распространяющихся в противоположных направлениях. Для простоты рассуждений допустим, что уравнения этих волн имеют вид:

$$y_1 = y_0 \sin \omega (t - x/v); \quad y_2 = y_0 \sin \omega (t + x/v).$$

Это означает, что в начале координат O ($x = 0$) обе волны вызывают колебания в одинаковой фазе. В точке A с координатой x суммарное значение колеблющейся величины, согласно принципу суперпозиции (см. § 19), равно

$$y = y_1 + y_2 = 2y_0 \cos \omega \frac{x}{v} \sin \omega t. \qquad (5.16)$$

Данное уравнение показывает, что в результате интерференции прямой и обратной волн в каждой точке среды (с фиксированной координатой x) происходит гармоническое колебание с той же частотой ω, но с амплитудой

$$Y_0 = 2y_0 \cos \omega \frac{x}{v} = 2y_0 \cos 2\pi \frac{x}{\lambda},$$

зависящей от значения координаты x. В точках среды, в которых $\cos(2\pi x/\lambda) = 0$, колебания отсутствуют вовсе: $Y_0 = 0$; эти точки называются *узлами колебаний*. В точках, где $\cos(2\pi x/\lambda) = \pm 1$, амплитуда

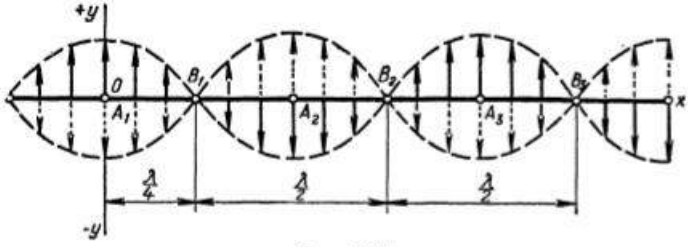


Рис. I.61

колебаний имеет наибольшее значение, равное $2y_0$. Эти точки называются *пучностями колебаний*. Легко показать, что расстояние между соседними узлами или соседними пучностями равно $\lambda/2$, расстояние между пучностью и ближайшим узлом равно $\lambda/4$. При изменении x на $\lambda/2$ косинус в формуле (5.16) меняет знак на обратный (его аргумент изменяется на π); поэтому если в пределах одной полуволны — от одного узла до другого — частицы среды отклонились в одну сторону, то в пределах соседней полуволны частицы среды будут отклонены в противоположную сторону.

Волновой процесс в среде, описываемый формулой (5.16), называется *стоячей волной*. Графически стоячая волна может быть изображена так, как это показано на рис. I.61. Допустим, что y есть смещение точек среды от состояния равновесия; тогда формула (5.16) описывает «стоячую волну смещения». В некоторый момент времени, когда $\sin \omega t = 1$, все точки среды имеют максимальные смещения, направление которых в зависимости от величины координаты x определяется знаком $\cos(2\pi x/\lambda)$. Эти смещения показаны на рис. I.61 сплошными стрелками. Спустя четверть периода, когда $\sin \omega t = 0$, смещения всех точек среды равны нулю; частицы среды проходят через линию Ox с различными скоростями. Спустя еще четверть периода, когда $\sin \omega t = -1$, частицы среды опять будут иметь максимальные смещения, но противоположного направления; эти смещения показаны на

рис. I.61 пунктирными стрелками. Точки A_1, A_2, ... суть пучности стоячей волны смещения; точки B_1, B_2, ... — узлы этой волны.

Характерные особенности стоячей волны в отличие от обычной распространяющейся, или бегущей, волны следующие (имеются в виду плоские волны при отсутствии затухания):

1) в стоячей волне амплитуды колебаний различны в различных местах системы; в системе имеются узлы и пучности колебаний. В «бегущей» волне эти амплитуды везде одинаковы;

2) в пределах участка системы от одного узла до соседнего все точки среды колеблются в одинаковой фазе; при переходе к соседнему участку фазы колебаний меняются на обратные. В бегущей волне фазы колебаний, согласно формуле (5.2), зависят от координат точек;

3) в стоячей волне нет одностороннего переноса энергии, как это имеет место в бегущей волне.

При описании колебательных процессов в упругих системах за колеблющуюся величину y можно принять не только смещение или скорости частиц системы, но и величину относительной деформации или величину напряжения на сжатие, растяжение или сдвиг и т. д. При этом в стоячей волне, в местах, где образуются пучности скоростей частиц, располагаются узлы деформаций и, наоборот, узлы скоростей совпадают с пучностями деформаций. Преобразование энергии из кинетической формы в потенциальную и обратно происходит в пределах участка системы от пучности до соседнего узла. Можно считать, что каждый такой участок не обменивается энергией с соседними участками. Заметим, что превращение кинетической энергии движущихся частиц в потенциальную энергию деформированных участков среды за один период происходит дважды.

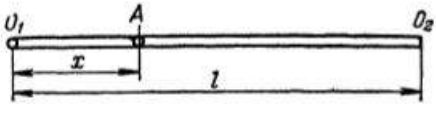


Рис. 1.62

Выше, рассматривая интерференцию прямой и обратной волн (см. выражения (5.16)), мы не интересовались происхождением этих волн. Допустим теперь, что среда, в которой происходит распространение колебаний, имеет ограниченные размеры, например колебания вызываются в каком-нибудь сплошном теле — в стержне или струне, в столбе жидкости или газа и т. д. Волна, распространяющаяся в такой среде (теле), отражается от границ, поэтому в пределах объема этого тела непрерывно происходит интерференция волн, вызванных внешним источником и отраженных от границ.

Рассмотрим простейший пример; допустим, в точке O_1 (рис. I.62) стержня или струны при помощи внешнего синусоидального источника возбуждается колебательное движение с частотой ω; начало отсчета времени выберем так, чтобы в этой точке смещение выражалось формулой

$$y = y_0 \sin \omega t,$$

где y_0 — амплитуда колебаний в точке O_1. Вызванная в стержне волна отразится от второго конца стержня O_2 и пойдет в обратном направ-

лении. Найдем результат интерференции прямой и отраженной волн в некоторой точке стержня A, имеющей координату x. Для простоты рассуждений предположим, что в стержне нет поглощения энергии колебаний и поэтому амплитуды прямой и отраженной волн равны.

В некоторый момент времени t, когда смещение колеблющихся частиц в точке O_1 равно y, в другой точке стержня A смещение y_1, вызванное прямой волной будет, согласно формуле волны, равно

$$y_1 = y_0 \sin \omega (t - x/v).$$

Через эту же точку A проходит также и отраженная волна. Чтобы найти смещение y_2, вызванное в точке A отраженной волной (в тот же самый момент времени t), необходимо рассчитать время Δt, в течение которого волна пройдет путь от O_1 до O_2 и обратно до точки A. Так как $O_1O_2 + O_2A = l + (l - x) = 2l - x$, а $\Delta t = (2l - x)/v$, то смещение, вызванное в точке A отраженной волной, будет равно

$$y_2 = y_0 \sin \omega \left(t - \frac{2l-x}{v}\right).$$

При этом предполагается, что на отражающем конце стержня в процессе отражения не происходит скачкообразного изменения фазы колебания; в некоторых случаях такое изменение фазы (называемое *потерей фазы*) имеет место и должно быть учтено.

Сложение колебаний, вызванных в различных точках стержня A (x) прямой и отраженной волнами, дает стоячую волну; действительно,

$$y = y_1 + y_2 = 2y_0 \cos \omega \frac{l-x}{v} \sin \omega \left(t - \frac{l}{v}\right) = y_{0x} \sin (\omega t - \varphi), \quad (5.17)$$

где $\varphi = \omega l/v = 2\pi l/\lambda$ — некоторая постоянная фаза, не зависящая от координаты x, а величина

$$y_{0x} = 2y_0 \cos \omega \frac{l-x}{v} = 2y_0 \cos 2\pi \frac{l-x}{\lambda} \quad (5.18)$$

является амплитудой колебаний в точке A; она зависит от координаты x, т. е. различна в различных местах стержня.

Найдем координаты тех точек стержня, в которых образуются узлы $\left(\cos 2\pi \frac{l-x}{\lambda} = 0\right)$ и пучности $\left(\cos 2\pi \frac{l-x}{\lambda} = 1\right)$ стоячей волны. Обращение косинуса в нуль или единицу происходит при значениях аргумента, кратных $\pi/2$:

$$2\pi \frac{l-x}{\lambda} = k\frac{\pi}{2}; \quad x = l - k\frac{\lambda}{4}, \quad (5.19)$$

где k — целое число. При нечетном значении этого числа косинус обращается в нуль и формула (5.19) дает координаты узлов стоячей волны; при четных k мы получим координаты пучностей.

Выше было произведено сложение только *двух волн*: прямой, идущей от O_1 к O_2, и отраженной, распространяющейся от O_2 к O_1. Однако следует учесть, что отраженная волна на границе стержня O_1 вновь отразится и пойдет в направлении прямой волны. Таких отражений

от концов стержня будет много, и поэтому необходимо найти результат интерференции не двух, а всех одновременно существующих в стержне волн.

Предположим, что внешний источник колебаний вызывал в стержне волны в течение некоторого времени τ, после чего поступление энергии колебаний извне прекратилось. За это время в стержне произошло $N = \tau/\tau'$ отражений, где $\tau' = l/v$ — время, в течение которого волна прошла от одного конца стержня к другому. Следовательно, в стержне будет одновременно существовать $N/2$ волн, идущих в прямом, и $N/2$ волн, идущих в обратном направлениях.

Допустим, что в результате интерференции одной пары волн (прямой и отраженной) смещение в точке A оказалось равным y. Найдем условие, при котором все смещения y, вызываемые каждой парой волн, имеют в точке A стержня одинаковые направления и поэтому складываются. Для этого фазы колебаний, вызванных каждой парой волн в точке A, должны отличаться на 2π от фазы колебаний, вызванных следующей парой волн. Но каждая волна вновь возвращается в точку A с тем же направлением распространения лишь спустя время $2l/v$, т. е. отстает по фазе на $\omega \cdot 2l/v = 4\pi l/\lambda$; приравнивая это отставание $2\pi n$, где n — целое число, получаем

$$4\pi \frac{l}{\lambda} = n \cdot 2\pi; \quad l = n \frac{\lambda}{2}, \tag{5.20}$$

т. е. вдоль длины стержня должно уместиться целое число полуволн. Заметим, что при этом условии фазы всех волн, идущих от O_1 в прямом направлении, отличаются друг от друга на $2\pi n$, где n — целое число; точно так же фазы всех волн, идущих от O_2 в обратном направлении, отличаются друг от друга на $2\pi n$. Поэтому, если одна пара волн (прямая и обратная) дает вдоль стержня распределение смещений, определяемое формулой (5.17), то при интерференции $N/2$ пар таких волн распределение смещений не изменится; увеличатся только амплитуды колебаний. Если максимальная амплитуда колебаний при интерференции двух волн, согласно формуле (5.18), равна $2y_0$, то при интерференции многих волн она будет больше. Обозначим ее через Y_0; тогда распределение амплитуды колебаний вдоль стержня вместо выражения (5.18) определится по формуле

$$y_{0x} = Y_0 \cos 2\pi \frac{l-x}{\lambda}. \tag{5.21}$$

Из выражений (5.19) и (5.20) определяются точки, в которых косинус имеет значения 0 или 1:

$$x = l - k\frac{\lambda}{4} = n\frac{\lambda}{2} - k\frac{\lambda}{4} = \frac{\lambda}{4}(2n-k), \tag{5.22}$$

где n — целое число ($n = 1, 2, 3, \ldots$). Координаты *узлов* стоячей волны получатся из этой формулы при нечетных значениях $k = 1, 3, 5, \ldots$; тогда в зависимости от длины стержня, т. е. величины n,

$$x = \frac{\lambda}{4}, \; 3\frac{\lambda}{4}, \; 5\frac{\lambda}{4} \text{ и т. д.;}$$

координаты *пучностей* получатся при четных значениях $k = 0, 2, 4, \ldots$:

$$x = 0;\ \frac{\lambda}{2};\ \lambda,\ \frac{3\lambda}{2},\ 2\lambda,\ \frac{5\lambda}{2}\ \text{и т. д.}$$

На рис. I.63 схематически показана стоячая волна в стержне, длина которого $l = 3\lambda/2$; точки A_1, A_2, A_3 суть пучности, точки B_1, B_2, B_3 — узлы этой стоячей волны.

В гл. 4 ч. I было показано, что при отсутствии периодических внешних воздействий характер колебательных движений в системе и прежде всего основная величина — частота колебаний — определяются размерами и физическими свойствами системы. Каждая колебательная система обладает собственным, ей присущим колебательным движением; это колебание можно наблюдать, если вывести систему из состояния равновесия и затем устранить внешние воздействия.

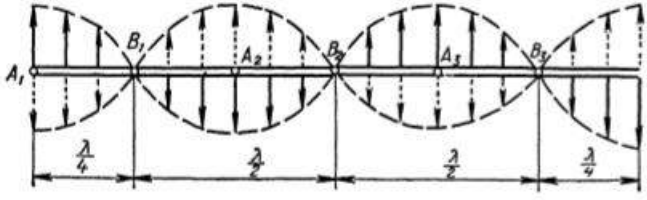


Рис. I.63

В гл. 4 ч. I рассматривались преимущественно колебательные *системы с сосредоточенными параметрами*, в которых инертной массой обладали одни тела (точечные), а упругими свойствами — другие тела (пружины). В отличие от них колебательные системы, в которых масса и упругость присущи каждому элементарному объему, называются *системами с распределенными параметрами*. К ним относятся рассмотренные выше стержни, струны, а также столбы жидкости или газа (в духовых музыкальных инструментах) и т. д. *Для таких систем собственными колебаниями являются стоячие волны*; основная характеристика этих волн — длина волны или распределение узлов и пучностей, а также частота колебаний — определяется только размерами и свойствами системы. Стоячие волны могут существовать и при отсутствии внешнего (периодического) воздействия на систему; это воздействие необходимо только для того, чтобы вызвать или поддержать в системе стоячие волны или же изменить амплитуды колебаний. В частности, если внешнее воздействие на систему с распределенными параметрами происходит с частотой, равной частоте ее собственных колебаний, т. е. частоте стоячей волны, то имеет место явление *резонанса*, рассмотренное в гл. 5.

Возвращаясь опять к стержню или струне, можно заметить, что при данной длине l в них могут одновременно возникать (существовать) стоячие волны различных частот. Действительно, условие $l = n\lambda/2$ может соблюдаться для некоторого множества длин волн:

$$l = n_1 \frac{\lambda_1}{2} = n_2 \frac{\lambda_2}{2} = n_3 \frac{\lambda_3}{2} = \ldots,$$

где n_1, n_2, ... — целые числа. Очевидно, эти характерные для данной системы длины волн λ_1, λ_2, ... должны быть кратны между собой: если $n_1 = 1$, $n_2 = 2$, $n_3 = 3$, ..., то

$$\lambda_2 = \lambda_1 \frac{n_1}{n_2} = \frac{\lambda_1}{2}; \quad \lambda_3 = \lambda_1 \frac{n_1}{n_3} = \frac{\lambda_1}{3} \text{ и т. д.}$$

Каждой из этих длин волн соответствует определенная частота колебаний:

$$\omega_1 = 2\pi \frac{v}{\lambda_1}; \quad \omega_2 = 2\pi \frac{v}{\lambda_2} = 2\omega_1;$$

$$\omega_3 = 2\pi \frac{v}{\lambda_3} = 3\omega_1 \text{ и т. д.}$$

Здесь предполагается, что вещество, из которого сделан стержень (т. е. колеблющаяся среда), не обладает дисперсией, т. е. скорость распространения волны v для различных частот одинакова.

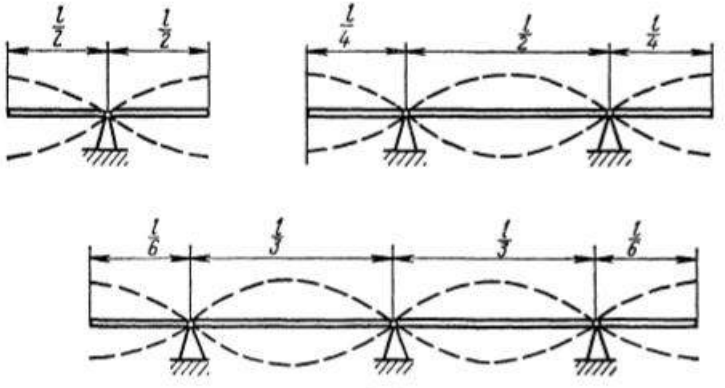


Рис. 1.64

Таким образом, *у систем с распределенными параметрами собственные колебания — стоячие волны — характеризуются целым спектром частот, кратных между собой*. Наименьшая из этих частот, соответствующая наибольшей длине волны ($n = 1$, $\lambda_1 = 2l$, $\omega_1 = \pi v/l$), называется *основной частотой*; остальные ($n = 2$, $\omega_2 = 2\omega_1$; $n = 3$, $\omega_3 = 3\omega_1$ и т. д.) — *обертонами* или *гармониками*.

Каждая система характеризуется не только наличием такого спектра колебаний, но и определенным *распределением энергии* между колебаниями различных частот. Для музыкальных инструментов это распределение придает звуку своеобразную особенность, так называемый *тембр звука*, различный для различных инструментов.

Изложенные выше расчеты относятся к свободному колеблющемуся стержню длиной l. Однако обычно мы имеем стержни, закрепленные на одном или обоих концах (например, колеблющиеся струны), или же вдоль стержня имеется одна или несколько точек закрепления. Места закрепления, где частицы системы не могут совершать колебательного движения, являются вынужденными узлами смещения. Например,

если в стержне необходимо получить стоячие волны при одной, двух, трех точках закрепления и т. д., то эти точки не могут быть выбраны произвольно, а должны располагаться вдоль стержня так, чтобы они оказались в узлах образовавшейся стоячей волны. Это показано, например, на рис. I.64. На этом же рисунке пунктиром показаны смещения точек стержня при колебаниях; на свободных концах всегда образуются пучности смещения, на закрепленных — узлы смещения. Для колеблющихся воздушных столбов в трубах узлы смещения (и скорости) получаются у отражающих твердых стенок; на открытых концах трубок образуются пучности смещений и скоростей.

§ 27. ЗВУКОВЫЕ КОЛЕБАНИЯ И ВОЛНЫ

Звуковые колебания, воспринимаемые человеческим ухом, имеют частоты, лежащие в пределах от 20 до 20 000 Гц; частоты, меньшие 20 Гц, называются *инфразвуковыми*, а бóльшие 20 000 Гц — *ультразвуковыми*.

При распространении в воздухе гармонической звуковой волны избыточное давление $\Delta p = p - p_0$ или избыточная плотность $\Delta\rho = \rho - \rho_0$ изменяются по формулам:

$$\Delta p = \Delta p_0 \cos\omega(t - x/v); \quad \Delta\rho = \Delta\rho_0 \cos\omega(t - x/v) \tag{5.23}$$

(если только Δp и $\Delta\rho$ малы по сравнению с p и ρ). Избыточное давление Δp иногда просто называют «звуковым давлением» (не смешивать с давлением звука на препятствия), а Δp_0 — амплитудой звуковой волны.

Интенсивностью или *силой звука* называют количество энергии, проносимое звуковой волной в единицу времени через единичную площадку, расположенную перпендикулярно направлению распространения; согласно формулам (5.15) и (5.13), интенсивность звука равна ($\rho_0 \approx \rho$)

$$I = \frac{1}{2}\rho v u_0^2 = \frac{\Delta p_0^2}{2\rho v}. \tag{5.24}$$

Амплитуды смещения, скорости и ускорения частиц среды при их колебательном движении в звуковой волне, а также $\Delta\rho_0$ можно выразить через интенсивность звука; так как $u_0 = y_0\omega$, $a_0 = u_0\omega$, то

$$y_0 = \frac{1}{\omega}\sqrt{\frac{2I}{\rho v}}; \quad u_0 = \sqrt{\frac{2I}{\rho v}}; \quad a_0 = \omega\sqrt{\frac{2I}{\rho v}}; \qquad \Delta\rho_0 = \frac{\Delta p_0}{v^2}. \tag{5.25}$$

Интенсивность звука выражается также через так называемое *эффективное звуковое давление* $\Delta p_{эф} = \Delta p/\sqrt{2}$:

$$I = \frac{\Delta p_{эф}^2}{\rho v}. \tag{5.26}$$

Произведение скорости звука в данной среде на плотность среды является важнейшей акустической характеристикой среды и назы-

вается *удельным акустическим сопротивлением*. Эта величина входит во все важнейшие расчетные формулы акустики. Например, при перпендикулярном падении звука на границу двух сред коэффициент k, показывающий, какая часть звуковой энергии отражается от этой границы, равен

$$k = \left(\frac{\rho_1 v_1 - \rho_2 v_2}{\rho_1 v_1 + \rho_2 v_2}\right)^2.$$

Для воздуха при нормальных условиях $\rho v \approx 430$ кг/(м^2·с), для воды $\rho v = 145 \cdot 10^4$ кг/(м^2·с), поэтому коэффициент отражения звука на границе между водой и воздухом равен $\approx$99,9%. В неоднородной среде, например в атмосфере, состоящей из множества слоев с различными удельными акустическими сопротивлениями, звук частично отражается от границ каждого слоя и поэтому быстро рассеивается в этой среде. В однородной атмосфере звук рассеивается меньше и поэтому слышен далеко. Для твердых тел удельные акустические сопротивления имеют более высокие значения [железо — $40 \cdot 10^6$, кварц — $15 \cdot 10^6$ кг/(м^2·с)].

Согласно формуле (5.24), интенсивность звука при данной амплитуде Δp_0 обратно пропорциональна удельному акустическому сопротивлению среды, а квадрат амплитуды звуковой волны Δp_0^2 при данной интенсивности звука прямо пропорционален этому сопротивлению.

При переходе звука из одной среды в другую частота звука ν остается неизменной, но длина волны изменяется пропорционально скорости распространения:

$$\frac{\lambda_1}{\lambda_2} = \frac{v_1}{v_2}. \qquad (5.27)$$

Например, если из воздуха в воду проходит звуковая волна частоты $\nu = 1000$ Гц, то длина волны в воздухе $\lambda_1 = 332$ (м/с) : 1000 (1/с) = = 0,332 м, а в воде $\lambda_2 = 1400$ (м/с) : 1000 (1/с) = 1,4 м.

ЭФФЕКТ ДОПЛЕРА

Если источник или приемник звука движется относительно среды, в которой происходит распространение звука, то частота звука ν, испускаемая источником, и частота ν' этого же звука, воспринимаемая приемником, будут отличаться друг от друга (эффект Доплера). Рассмотрим частный случай, когда вдоль O—O_1 (рис. I.65) распространяется плоская гармоническая звуковая волна. Допустим:

1) источник звука движется относительно среды со скоростью $\pm u$; приемник звука покоится. В этом случае за период колебаний T звуковая волна отойдет от источника (камертона) на расстояние vT, а сам источник сместится на uT. На рис. I.65 кривые a и b условно изображают звуковые волны от неподвижного (a) и движущегося ($б$) источника. Если источник звука удаляется от приемника, как это показано на рис. I.65, т. е. движется в направлении, обратном направлению распространения, то длина волны $\lambda' = vT + uT = (v + u)\,T$; если же источник звука движется к приемнику, т. е. в направлении распространения звука, то, очевидно, $\lambda' = vT - uT =$

$= (v - u)\,T$. Тогда частота звука, воспринимаемая приемником в обоих указанных случаях, может быть представлена в виде

$$\nu' = \frac{v}{\lambda'} = \frac{\nu}{(1 \pm u/v)}, \tag{5.28}$$

где $\nu = v/\lambda = 1/T$ — частота колебаний источника;

2) источник звука покоится, приемник движется относительно среды со скоростью $\pm w$. В этом случае длина волны в среде не изменяется и равна $\lambda = vT$; однако две последовательные амплитуды волны дойдут до приемника не через T секунд, а через большее или меньшее

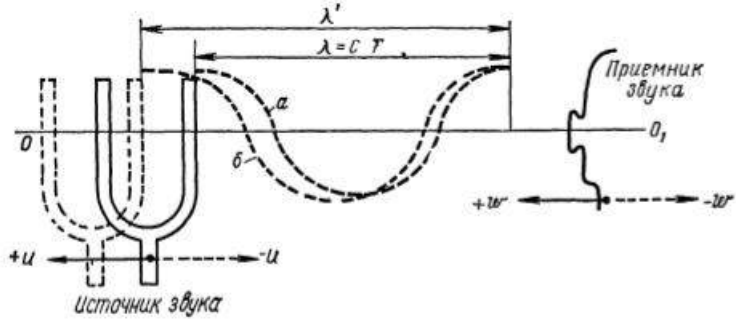


Рис. I.65

время T' в зависимости от того, удаляется или приближается приемник к источнику звука. За это время T' звук распространится на расстояние vT', а приемник сместится на $\pm wT'$. Сумма этих величин должна равняться длине волны, следовательно, $\lambda = vT' \pm wT' = (v \pm w)\,T'$. Так как частота звука, воспринимаемая приемником $\nu' = 1/T'$, а $v/\lambda = \nu = 1/T$, то

$$\nu' = \nu\,(1 \pm w/v); \tag{5.29}$$

3) источник и приемник звука движутся относительно среды. Соединяя результаты, полученные для двух предыдущих случаев, получим

$$\nu' = \nu \frac{1 \pm w/v}{1 \pm u/v}. \tag{5.30}$$

ХАРАКТЕРИСТИКИ ЗВУКОВОГО ОЩУЩЕНИЯ

Звуковое ощущение характеризуется высотой звука, тембром и громкостью. *Высота звука* определяется частотой колебаний. Однако источники звука испускают не одну, а целый спектр частот, причем энергия волны как-то распределяется между различными частотами. Высота звука определяется по одной — основной частоте, если на долю этой частоты приходится значительно большее количество энергии, чем на долю других частот. Если спектр состоит из отдельных частот, то он называется *линейчатым*, если же из непрерывного набора частот, то *сплошным*. На рис. I.66 показан акустический спектр рояля,

соответствующий основной частоте (тону) 256 Гц. Этот спектр является примером смешанного спектра; он состоит из набора отдельных частот и участка сплошного спектра (в области *а — б*). В этом спектре выделяется основной тон звука — частота 256 Гц; на звук этой частоты приходится относительно наибольшая интенсивность; на другие частоты приходится меньше энергии и они слабее слышны.

Акустический спектр звука в зависимости от своего характера (линейчатый, сплошной или смешанный) и от распределения энергии между частотами определяет своеобразие звукового ощущения, называемое *тембром звука*. Различные музыкальные инструменты, испускающие звук одного и того же тона, имеют различный акустический спектр, т. е. отличаются тембром звука.

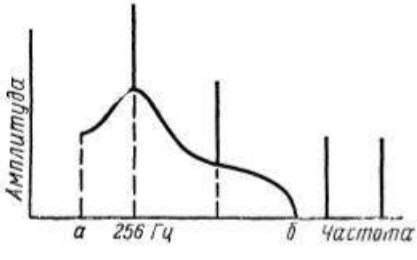


Рис. 1.66

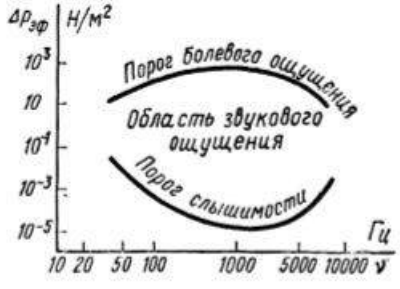


Рис. 1.67

При оценке звукового восприятия следует учесть, что чувствительность человеческого уха различна для различных частот. Интенсивность (сила) звука данной частоты должна быть достаточно большой, чтобы вызвать звуковое ощущение, однако если эта интенсивность превышает некоторый предел, то звук не слышен и вызывает только болевое ощущение. Таким образом, для каждой частоты колебаний существует наименьшая (*порог слышимости*) и наибольшая (*порог болевого ощущения*) интенсивность звука, которая способна вызвать звуковое восприятие. На рис. 1.67 показана область звукового восприятия человеческого уха. Наибольшую чувствительность ухо имеет к частотам от 800 до 6000 Гц. В этом интервале наименьшая интенсивность звука (порог слышимости) равна 10^{-12} Вт/м²; порог болевого ощущения приблизительно равен 1 Вт/м².

Установлено, что ощущение звука человеческим ухом, называемое *громкостью*, зависит от интенсивности этого звука I, но не пропорционально ей. Звуки, интенсивности которых лежат в некоторых пределах I и $I + \Delta I$, воспринимаются как имеющие одинаковую громкость; поэтому, чтобы ухо могло установить различие в громкости двух звуков, их интенсивности I_1 и I_2 должны отличаться больше чем на ΔI. Таким образом, ΔI есть минимальное изменение интенсивности звука, которое может быть отмечено ухом как изменение громкости этого звука.

Согласно общефизиологическому закону Вебера — Фехнера, можно полагать, что *величина* ΔI *пропорциональна* I, *т. е.*

способность уха различать громкости звуков ослабевает с увеличением интенсивности. Отношение $\Delta I/I$ оказывается почти постоянной величиной (это утверждение соблюдается хорошо для частот от 100 до 4000 Гц). Поэтому с увеличением интенсивности звука увеличивается ΔI, следовательно, чувствительность уха к восприятию изменения громкости уменьшается.

Допустим, что громкость звука оценивается некоторой величиной L. Изменение громкости dL было бы прямо пропорционально изменению интенсивности звука dI только в том случае, если бы чувствительность уха к интенсивности звука оставалась неизменной. Однако эта чувствительность ослабевает с возрастанием I, причем dL оказывается пропорциональной отношению dI/I. Интегрируя равенство $dL = kdI/I$ (где k — постоянный коэффициент пропорциональности) в пределах от I до наименьшего значения интенсивности I_0, соответствующего порогу слышимости, получим $L = k \ln (I/I_0)$. Таким образом, субъективное ощущение громкости звука оказывается пропорциональным логарифму отношения интенсивности звука к соответствующему порогу слышимости. На этом основании условились ввести *объективную оценку громкости звука* по измеренному значению его интенсивности I:

$$L = \lg (I/I_0) \text{ белов} \quad \text{или} \quad L = 10 \lg (I/I_0) \text{ децибелов}, \qquad (5.31)$$

где для всех звуков условились полагать $I_0 = 10^{-12}$ Вт/м². Величину L называют *уровнем интенсивности звука* (обычно в децибелах). Если измеряется не интенсивность звука, а эффективное звуковое давление (см. формулу (5.26)), то

$$L = 20 \lg \frac{\Delta p_{эф}}{\Delta p_{сл}}, \qquad (5.32)$$

где $\Delta p_{сл} = 2 \cdot 10^{-5}$ Н/м² есть звуковое давление нижнего порога восприятия звука. В формуле (5.32) L называют *уровнем звукового давления* (в децибелах).

Физиологической характеристикой звука является *уровень громкости звука*, который измеряется в *фонах* и определяется следующим образом: изменяя интенсивность стандартного звука частоты 1000 Гц, добиваются, чтобы измеряемый и стандартный звук имели одинаковую громкость (по звуковому восприятию). Громкость в фонах измеряемого звука приравнивается числу децибел уровня звукового давления стандартного звука.

Ниже приводятся уровни L звукового давления (дБ), интенсивности звука I (Вт/м²) и эффективные звуковые давления $\Delta p_{эф}$ (Па) для некоторых звуков в пределах от порога слышимости до болевого порога:

	L, дБ	I, Вт/м²	$\Delta p_{эф}$, Па
Порог слышимости	0	10^{-12}	$2 \cdot 10^{-5}$
Тихий шепот	20	10^{-10}	$2 \cdot 10^{-4}$
Громкая речь	70	10^{-5}	$6{,}3 \cdot 10^{-2}$
Оркестр (громко)	100	10^{-2}	2
Болевой порог	130	10	63

Мощность, развиваемая источником звука, может быть представлена в зависимости от интенсивности звука, звукового давления, а также от громкости. При разговоре средней громкости мощность, развиваемая голосовыми связками человека, составляет около 100 эрг/с; при переходе к громкой речи эта мощность увеличивается в десятки и сотни раз.

Для звукового восприятия в помещениях большое значение имеет *реверберация звука*, т. е. постепенное ослабление его интенсивности вследствие поглощения при многократных отражениях от стен, потолка, предметов и т. д. Каждый звук существует в помещении некоторое время, пока его интенсивность не уменьшится до порога слышимости. Слишком медленное затухание звука (в пустых помещениях) создает «гулкость» помещения. При очень быстром затухании звуки получаются приглушенными (в комнатах, обвешанных коврами). *Временем реверберации* называют время, в течение которого интенсивность звука в помещении ослабляется в миллион раз. Это время различно для различных частот и зависит от объема помещения и формы, размеров и акустических свойств находящихся в них тел. Для стандартной частоты 512 Гц это время должно быть для небольших помещений $\approx$1 с (оптимум реверберации); для больших концертных залов и театров (при условии заполнения их людьми) оно приближается к 2 с.

Время реверберации концертного зала зависит от того, пуст он или заполнен слушателями. В одном оперном театре восприятие музыки (акустическое качество зала) заметно ухудшилось вследствие изменения моды: женщины сменили пышную одежду на плотно облегающую и тем вызвали некоторое изменение времени реверберации в зале. Утверждают, что, измеряя время реверберации, можно установить, находятся ли в концертном зале мужчины или женщины, разумеется, если есть заметное различие в акустических свойствах их одежды и прически.

§ 28. УЛЬТРАЗВУКИ И ИХ ПРИМЕНЕНИЯ

Ультразвуки имеют частоты колебаний свыше 20 000 Гц, поэтому длина ультразвуковых волн в различных средах мала, а скорости и ускорения колеблющихся частиц среды и возникающие избыточные давления велики. Интенсивность ультразвукового излучения некоторых источников достигает миллионов ватт на квадратный метр. Однако, прежде чем рассчитывать важнейшие характеристики ультразвуковой волны по формулам (5.25), необходимо проверить, соблюдается ли условие малости $\Delta p \ll p_0$, на основании которого были получены эти формулы. При частоте $\nu = 160$ кГц ($\omega = 2\pi\nu = 10^6$ 1/с) и интенсивности $I = 5$ Вт/см^2 $= 5 \cdot 10^4$ Вт/м^2 по формуле (5.24) получаем для воздуха ($\rho v \approx 430$ кг/(м$^2 \cdot$с))

$$\Delta p_0 = \sqrt{2I\rho v} \approx 6500 \text{ Па} = 0{,}065 \text{ атм}.$$

Рассчитаем остальные величины:

1) длина волны $\lambda = v/\nu = 332 : 160\,000 = 0{,}0021$ м $= 2{,}1$ мм;

2) амплитуда смещения частиц среды

$$y_0 = \frac{1}{\omega}\sqrt{\frac{2I}{\rho v}} = \frac{1}{10^6}\sqrt{\frac{2 \cdot 5 \cdot 10^4}{430}} \approx 1{,}54 \cdot 10^{-5} \text{ м} = 0{,}0154 \text{ мм};$$

3) амплитуда скорости колебательного движения частиц среды $u_0 = \omega y_0 = 15{,}4$ м/с;

4) амплитуда ускорений колебательного движения частиц среды $a_0 = \omega u_0 = 15{,}4 \cdot 10^6$ м/с^2. При больших частотах и интенсивностях можно получить ускорения, в миллионы раз превышающие ускорение силы тяжести.

Ультразвуки могут быть получены при помощи известных источников слышимых частот, имеющих соответствующие размеры или параметры (миниатюрные камертоны, короткие струны, свистки, сирены с большим числом оборотов ротора и т. д.). Большое применение получили источники ультразвуков, основанные на использовании магнитострикции и электрострикции — изменения размеров тел, помещенных в магнитное и электрическое поля. Если в быстропеременное магнитное поле поместить, например, никелевый стержень, то длина его будет изменяться (на несколько тысячных долей процента) в соответствии с частотой поля. При резонансе между внешним воздействием и собственными колебаниями стержня можно получить большие амплитуды колебаний и, следовательно, большие интенсивности излучаемой волны. Таким же образом получают ультразвуковые волны от пластинки кварца (или другого диэлектрика), помещенного в высокочастотное электрическое поле.

Различные источники ультразвука характеризуются *потоком звуковой энергии* (*звуковой мощностью*) (от долей ватта до десятков киловатт) и коэффициентом полезного действия, т. е. отношением этой мощности к потребляемой. В хороших образцах ультразвуковых генераторов этот коэффициент достигает 60—70%. Мощность излучения данного источника зависит от акустического сопротивления среды, в которой возбуждаются звуки. Например, кварцевая пластинка при одинаковой частоте и амплитуде колебаний будет, согласно формуле (5.15), потреблять и излучать в воде в 3500 раз бóльшую мощность, чем в воздухе.

Благодаря малой длине волны ультразвуков их можно фокусировать при помощи вогнутых отражателей или соответствующих «ультразвуковых линз» (например, алюминиевых). При этом можно получить большую концентрацию мощности в единице объема среды. Например, если вся энергия, излучаемая кварцевой пластинкой диаметром $d = 1{,}15$ см при интенсивности излучения $I = 10$ Вт/см^2, фокусируется в 1 см^3 среды, то ежесекундно в этот объем будет поступать $I(\pi d^2/4) = 10$ Дж энергии. При полном поглощении эта энергия может вызвать нагревание одного грамма воды со скоростью до 2,5° С в секунду.

Кроме зеркал и линз увеличение амплитуды колебаний и интенсивности ультразвука достигается также применением стержней с уменьшающимися сечениями, которые присоединяются к излучающей

поверхности источника ультразвука (рис. I.68). Можно выбрать вещество стержня, слабо поглощающее энергию колебаний, а при подходящем законе убывания сечения довести до минимума излучение энергии через боковую поверхность стержня. Тогда ежесекундные количества энергии, проходящие через два сечения S_1 и S_2, будут равны и, следовательно,

$$I_1 S_1 = I_2 S_2; \quad I_2 = I_1 \frac{S_1}{S_2}.$$

Так как интенсивность звука пропорциональна квадрату амплитуд смещения, скорости и т. д., то для круглых сечений интенсивность I будет обратно пропорциональна квадрату диаметра, а y_0, u_0, Δp_0 — первой степени диаметра стержня.

Высокие значения акустических скоростей, ускорений, избыточных давлений и плотностей, а также хорошо разработанные методы излучения, приема, измерения интенсивности и скорости распространения ультразвуков позволили использовать их для решения многих технических задач. Перечислим важнейшие применения ультразвука:

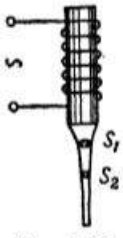


Рис. I.68

1) использование ультразвука как средства связи и обнаружения; определение местонахождения предметов и неоднородностей в акустически прозрачных средах; в морях — *акустическая локация* косяков рыб, подводных лодок, определение глубины; в массивных металлических поковках и отливках — обнаружение внутренних трещин и раковин (дефектоскоп С. Я. Соколова);

2) изучение физических свойств различных твердых, жидких и газообразных веществ (скорость распространения, коэффициент поглощения и т. д.);

3) воздействие на различные физико-химические процессы: кристаллизацию, намагничивание, диффузию, различные электрохимические процессы и т. д.; образование эмульсий;

4) механическая обработка очень твердых или очень хрупких тел; очистка мелких предметов (деталей часовых механизмов и т. д.), помещенных в жидкость; обезгаживание;

5) воздействие на биологические объекты, различные применения в медицине (хирургия, лечение некоторых заболеваний и т. п.).

Некоторые животные и насекомые испускают и воспринимают ультразвуковые колебания различных частот: дельфины — до 50 кГц, пчелы — до 22 кГц, собаки и мыши слышат ультразвуки до 100 кГц и т. д. Летучие мыши испускают ультразвуки короткими импульсами; продолжительность каждого импульса составляет тысячные доли секунды, число таких импульсов в секунду от 5 до 60, частота колебаний от 30 до 120 кГц. Интересно, что ночные бабочки, являющиеся пищей для летучих мышей, воспринимают ультразвуки с частотами от 10 до 200 кГц и благодаря этому могут обнаружить грозящую им опасность.

Глава 6

МЕХАНИКА ЖИДКОСТЕЙ И ГАЗОВ

§ 29. ГИДРОСТАТИЧЕСКОЕ ДАВЛЕНИЕ. БАРОМЕТРИЧЕСКАЯ ФОРМУЛА

Каждая масса жидкости при отсутствии внешних сил (под действием одних только сил притяжения между молекулами) имеет определенный объем и приобретает форму шара. Внешними силами можно изменить величину объема (т. е. произвести сжатие или расширение жидкости), а также изменить форму, которую имеет этот объем. Под действием распределенной по всему объему силы тяготения к Земле жидкость принимает форму сосуда, причем ее открытая поверхность располагается горизонтально. Капля жидкости, падающая в воздухе, под действием распределенной вдоль ее поверхности силы сопротивления воздуха принимает удобообтекаемую форму, при которой сопротивление движению минимальное; чем меньше скорость падения, тем ближе форма капли к сферической. В состоянии невесомости (при падении в безвоздушном пространстве, а также в космических кораблях) свободная жидкость принимает форму шара.

Газы занимают определенный объем только при наличии внешнего сжимающего давления со стороны стенок сосуда, в котором находится данная масса газа. Газы и жидкости могут подвергаться только *всестороннему* (объемному) сжатию или расширению. Сжимаемость жидкостей и газов оценивается *изотермическим* (т. е. при постоянной температуре) *коэффициентом всестороннего сжатия* (*сжимаемости*):

$$k = -\frac{1}{V}\left(\frac{dV}{dp}\right)_{T=\text{const}}, \qquad (6.1)$$

где dV — изменение объема V при изменении давления на dp при постоянной температуре T. Коэффициент сжимаемости зависит от температуры и давления. При 0° С и $p = 1$ атм вода имеет $k \approx \approx 0{,}5 \cdot 10^{-10}$ м²/Н, ртуть — $3{,}8 \cdot 10^{-11}$ м²/Н. Для газов k сильно зависит от того, как изменяется температура при изменении объема. Обратная величина $1/k = D$ называется *модулем объемной упругости* и измеряется в паскалях (Н/м²) или атмосферах (кгс/см²).

Если внутрь покоящейся жидкости или газа поместить очень тонкую твердую пластинку, то части жидкости или газа, расположенные по обе стороны ее, будут действовать на площадку ΔS с силами ΔF, которые независимо от ориентации пластинки равны по величине и направлены перпендикулярно площадке (рис. 1.69, *а*). Эти силы существуют и при отсутствии пластинки, для любой мысленно проведенной площадки ΔS. Предел

$$p = \lim_{\Delta S \to 0} \frac{\Delta F}{\Delta S}$$

является скалярной величиной и называется *давлением* в данном месте жидкости или газа.

Согласно з а к о н у П а с к а л я, *внешнее давление p_0, приложенное к поверхности жидкости, передается всему объему*, чем вызывается

всестороннее (объемное) сжатие. Например, в гидравлическом прессе при помощи малой силы F_1, приложенной к поршню с малой площадью S_1, создается большое внешнее давление на жидкость $p = F_1/S_1$; оно передается всему объему жидкости и на стенки сосуда.

В условиях земного тяготения открытая поверхность покоящейся жидкости (небольших размеров) — горизонтальная. На любой горизонтальной поверхности, взятой внутри жидкости, давление везде одинаковое. В смеси неоднородных жидкостей внизу располагается жидкость с большей плотностью; границей раздела между ними также является горизонтальная поверхность.

Согласно закону Архимеда, *равнодействующая всех сил давления, приложенных к поверхности тел, погруженных в жидкость* (или газ), *направлена вертикально вверх и равна весу жидкости* (или газа) *в объеме данного тела*. Тело может находиться внутри жидкости (газа) в равновесии, если сила Архимеда равна весу тела; у тел, плавающих на поверхности жидкости, равенство этих сил достигается при определенном погружении.

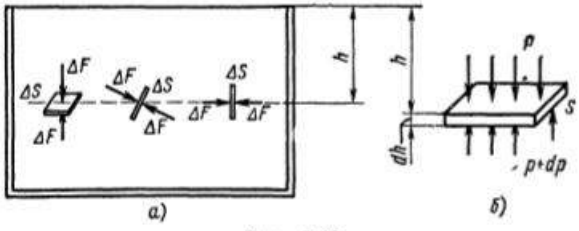


Рис. 1.69

Напишем условия равновесия горизонтального слоя жидкости (газа) толщиной dh. Для элемента с площадью S (рис. 1.69, *б*) сила давления снизу должна уравновесить вес элемента и силу давления сверху; обозначим через ρ плотность жидкости (газа); тогда

$$(p + \mathrm{d}p)\, S = \rho g S\, \mathrm{d}h + pS,$$

откуда

$$\mathrm{d}p = \rho g\, \mathrm{d}h.$$

Если от поверхности $h = 0$ до глубины h величины ρ и g не изменяются, то, интегрируя, получим:

$$p - p_0 = \rho g h; \quad p = p_0 + \rho g h. \tag{6.2}$$

По этой формуле рассчитывается давление внутри покоящейся жидкости на глубине h от открытой горизонтальной поверхности; величина $\rho g h$ называется *гидростатическим давлением*.

При расчете давления внутри атмосферы удобнее вместо h отсчитывать высоту H от поверхности Земли до интересующего нас места. Так как $dh = -dH$, то $dp = -\rho g dH$. Если плотность воздуха прямо пропорциональна его давлению (что имеет место вблизи поверхности Земли при одинаковой температуре воздуха), то $\rho = \mathrm{const} \cdot p$ и инте-

грирование дает так называемую *барометрическую формулу*

$$p = p_0 e^{-\text{const}\cdot gH}; \quad \rho = \rho_0 e^{-\text{const}\cdot gH}, \tag{6.3}$$

где p_0 — давление воздуха на поверхности Земли. На небольшой высоте $p \approx p_0 - \rho_0 gH$. Значение const можно найти, полагая воздух идеальным газом. Согласно формуле (2.11), которая указана в § 9 ч. II курса, у идеального газа $p = nkT$, где n — число частиц в единице объема газа; k — постоянная Больцмана; T — абсолютная температура газа. Так как $\rho = nm$, где m — масса одной частицы газа, то $\rho = mp/kT$, т. е. константа в формуле (6.3) равна m/kT, следовательно,

$$p = p_0 e^{-\frac{mgH}{kT}}; \quad \rho = \rho_0 e^{-\frac{mgH}{kT}}.$$

Заметим также, что mgH есть работа, совершаемая при подъеме одной частицы газа на высоту H.

§ 30. ЛАМИНАРНОЕ И ТУРБУЛЕНТНОЕ ТЕЧЕНИЯ. ФОРМУЛА БЕРНУЛЛИ

Жидкость (газ) может быть приведена в движение различными силами: силой тяжести, разностью давлений в различных местах объема, силами трения (вязкости) между слоями, движущимися с различными скоростями, и т. п. Проведем в потоке жидкости л и н и ю т о к а так, чтобы вектор скорости в каждой точке лежал на касательной к этой линии. Течение называется *установившимся* (стационарным), когда форма и расположение линий тока, а также значения скоростей в каждой ее точке со временем не изменяются. Течение называется *ламинарным* (слоистым), если вдоль потока каждый выделенный тонкий слой (или струя) скользит относительно соседних, не перемешиваясь с ними, и *турбулентным* (вихревым), если вдоль потока происходит интенсивное вихреобразование и перемешивание жидкости (газа).

Если можно пренебрегать сжимаемостью жидкости (газа) и силами вязкости, то для установившегося течения в каждой точке линии тока соблюдается соотношение (у р а в н е н и е Д. Б е р н у л л и)

$$\frac{\rho v^2}{2} + \rho gh + p = \text{const}. \tag{6.4}$$

Для вывода этого уравнения рассмотрим струю жидкости между сечениями S_1 и S_2, боковая поверхность которой образована линиями тока (рис. I.70). За малое время Δt сквозь сечение S_1 пройдет элементарный объем жидкости в форме цилиндра с основанием S_1 и высотой $v_1\Delta t$; каждая единица объема прошедшей через S_1 жидкости вносит кинетическую энергию $\rho_1 v_1^2/2$ и потенциальную энергию $\rho_1 g h_1$. Внешняя сила $p_1 S_1$, действующая в сечении S_1, смещает указанный объем жидкости на $v_1\Delta t$ и поэтому совершает положительную работу, равную $p_1 S_1 v_1 \Delta t$. За то же время Δt через второе сечение S_2 выйдет жидкость в объеме цилиндра $S_2 v_2 \Delta t$, а внешняя сила $p_2 S_2$ совершит отрицательную работу, равную $p_2 S_2 v_2 \Delta t$. При установившемся состоянии течения

полная энергия жидкости в объеме струи между сечениями S_1 и S_2 должна оставаться постоянной, поэтому сумма изменений всех видов энергии и работ внешних сил должна равняться нулю:

$$S_1 v_1 \Delta t \left(\frac{\rho_1 v_1^2}{2} + \rho_1 g h_1\right) + p_1 S_1 v_1 \Delta t - S_2 v_2 \Delta t \left(\frac{\rho_2 v_2^2}{2} + \rho_2 g h_2\right) - p_2 S_2 v_2 \Delta t = 0.$$

Предположим, что жидкость несжимаема ($\rho_1 = \rho_2$) и струя не имеет разрывов; тогда объемы жидкости, ежесекундно поступающей через S_1 и выходящей через S_2, будут равны:

$$S_1 v_1 = S_2 v_2.$$

Произведя сокращения, мы получаем уравнение (6.4).

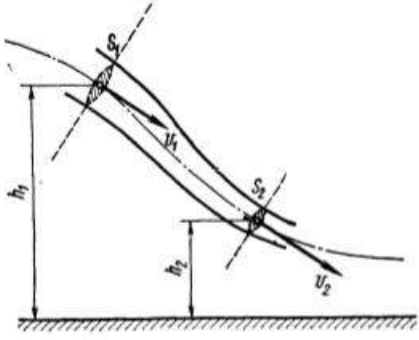


Рис. I.70

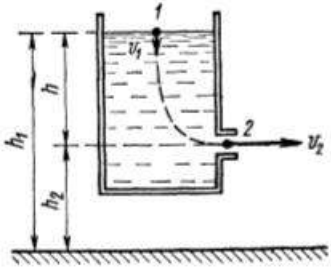


Рис. I.71

Уравнение Бернулли применяется для решения некоторых задач гидро- и аэромеханики. Приведем два примера:

1) расчет скорости истечения жидкости из резервуара. Для двух точек линии тока *1* и *2* имеем (рис. I.71)

$$\frac{\rho v_1^2}{2} + \rho g h_1 + p_1 = \frac{\rho v_2^2}{2} + \rho g h_2 + p_2.$$

Так как p_1 и p_2 равны наружному давлению воздуха p_0, а v_1 очень мала, то

$$v_2 = \sqrt{v_1^2 + 2g(h_1 - h_2)} \approx \sqrt{2gh}. \qquad (6.5)$$

Для струи газа, выходящего из сосуда с высоким давлением p_1 в сосуд с низким давлением p_2, пренебрегая $\rho g h$, получим

$$v = \sqrt{\frac{2(p_1 - p_2)}{\rho}}; \qquad (6.6)$$

2) расчет давления в узком сечении потока (рис. I.72). Полагая $h_1 = h_2$, найдем

$$p_2 = p_1 - \frac{1}{2}\rho(v_2^2 - v_1^2).$$

Так как $v_2 > v_1$, то $p_2 < p_1$ и может быть сделано меньше атмосферного. Этим можно воспользоваться для всасывания газа или жидкости в трубу через боковые отверстия C.

Кроме сил тяжести и разности давлений движение жидкости (газа) можно вызвать или затормозить силами трения между ее слоями. В ламинарном потоке жидкости (газа) сила трения между двумя соседними слоями, движущимися со скоростями v и $v + dv$, равна (рис. I.73)

$$F = \eta \frac{dv}{dn} S, \tag{6.7}$$

где η — *коэффициент вязкости* (динамический); dv/dn — *градиент скорости потока* в данном месте, т. е. быстрота изменения скорости в направлении, перпендикулярном самой скорости, следовательно, и площадке S (S — также площадь соприкосновения). Характер изменения скорости по мере удаления от стенок трубы или канала показан кривой a; у самих стенок вследствие прилипания жидкости и торможения ее шероховатой поверхностью стенки скорость течения равна нулю. Коэффициент вязкости η измеряется в ньютон-секундах на квадратный метр (Н·с/м²), а в системе СГС — в пуазах [г/(см·с)]. Коэффициент вязкости жидкостей зависит заметно от температуры: у воды при 0° С $\eta_0 = 0{,}0018$, а при 15° С уже делается равной 0,0011 Н·с/м². У воздуха соответственно: $\eta_0 = 1{,}71 \cdot 10^{-4}$, $\eta_{15} = 1{,}81 \cdot 10^{-4}$ Н·с/м². В некоторые формулы входит отношение $\eta/\rho = \nu$,

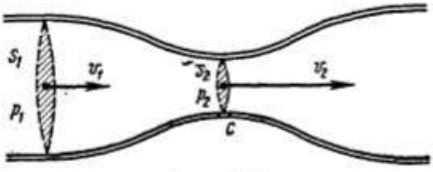


Рис. I.72

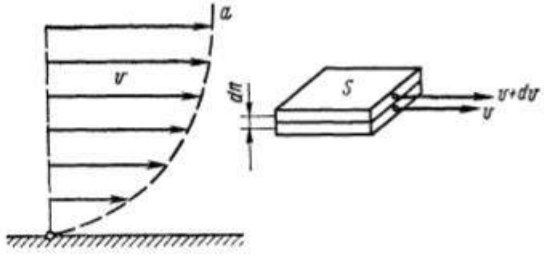


Рис. I.73

которое называется *кинематической вязкостью* жидкости (газа) и измеряется в квадратных метрах на секунду (м²/с).

На тело, движущееся внутри жидкости или газа, а также на тело, обтекаемое движущимся потоком жидкости или газа, действуют:

1) касательные силы трения между телом и обтекаемой жидкостью (газом). Для каждого элемента поверхности ΔS (рис. I.74) эта сила ΔF_t определяется градиентом скорости потока в пограничном слое и коэффициентом вязкости; она направлена по касательной к обтекаемому элементу поверхности;

2) нормальные силы давления ΔF_n, зависящие от значения давления в потоке возле каждого элемента ΔS и направленные перпендикулярно ему ($\Delta F_n = p\Delta S$).

Векторное сложение этих сил для всей обтекаемой поверхности определяет величину, направление и точку приложения полной силы, действующей на тело в потоке жидкости (газа):

$$\mathbf{R} = \sum (\Delta \mathbf{F}_t + \Delta \mathbf{F}_n).$$

Проекция R на направление потока называется *лобовым сопротивлением*; другая составляющая, если она направлена вертикально вверх,

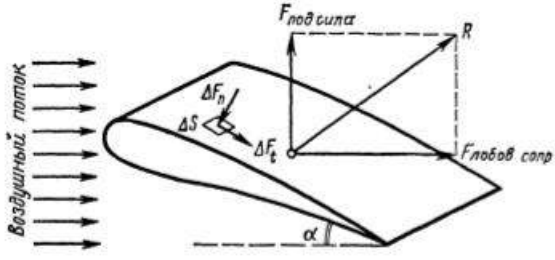


Рис. 1.74

называется *подъемной силой*. Величина этих сил зависит от формы (обтекаемость) и размеров тел, их расположения относительно потока (угол атаки α, рис. 1.74), состояния обтекаемой поверхности (шероховатость) и, кроме того, определяется скоростью движения тела или обтекающего потока и свойствами жидкости (плотностью, коэффициентом вязкости и т. д.). Измерение и вычисление этих сил составляют важную задачу аэрогидродинамики. Выдающиеся исследования в этой области принадлежат Н. Е. Жуковскому, С. А. Чаплыгину и др.

СЖИМАЕМОСТЬ И ВЯЗКОСТЬ В ПОТОКЕ

Для упрощения некоторых аэрогидродинамических расчетов и, в частности, при выводе уравнения Бернулли пренебрегают сжимаемостью жидкостей и газов, а также работой сил вязкости. Выясним, при каких условиях это возможно. Допустим, что в горизонтальном потоке максимальная скорость на линии тока равна v, а минимальная — нулю. Для этих двух точек уравнение Бернулли дает:

$$\frac{\rho v^2}{2} + p_1 = p_2; \quad v = \sqrt{\frac{2(p_2 - p_1)}{\rho}}. \tag{6.8}$$

Если $p_2 \gg p_1$, то $v \approx \sqrt{2p_2/\rho}$; для газов это близко к скорости распространения звука. Следовательно, чтобы разность давлений была небольшой и сжимаемостью можно было пренебречь, максимальная скорость потока газа должна быть значительно меньше скорости звука в этом газе. Рассчитаем, например, какое изменение давления могло получиться в потоке воздуха, движущегося со звуковой ско-

ростью $v = 340$ м/с, если затормозить поток до $v = 0$ (плотность воздуха возьмем при 0°С):

$$p_2 - p_1 = \frac{\rho v^2}{2} = \frac{1{,}29 \cdot 340^2}{2} = 0{,}645 \cdot 10^5 \text{ Па},$$

или около 0,65 атм. При таком изменении давления, очевидно, нельзя пренебрегать ни сжимаемостью воздуха, ни изменением его температуры. Поэтому для дозвуковых и сверхзвуковых течений газов потребовалась разработка специальной области механики — *газодинамики*.

Работой сил вязкости можно пренебрегать при следующих условиях; рассмотрим течение жидкости (газа) по трубе радиуса r и допустим, что труба расположена горизонтально, а на участке l скорость течения не изменяется. Тогда, согласно уравнению Бернулли, при отсутствии трения между жидкостью (газом) и стенками сосуда давление на концах участка l должно быть одинаковым. При наличии трения появится разность давлений $p_2 - p_1$, которую можно приравнять отношению силы трения F между жидкостью (газом) и стенками трубы к сечению трубы πr^2:

$$p_2 - p_1 = \frac{F}{\pi r^2}.$$

Согласно формуле (6.7),

$$F = \eta \frac{dv}{dn} S = \eta \frac{dv}{dn} \cdot 2\pi r l.$$

Градиент скорости dv/dn заменим средней величиной v/r, полагая, что у стенки трубы $v = 0$, а скорость по оси трубы равна v. Следовательно,

$$p_2 - p_1 = 2\eta \frac{dv}{dn} \frac{l}{r} = \frac{2\eta v l}{r^2}.$$

Силами трения можно пренебрегать, если эта разность давления достаточно мала по сравнению с $\rho v^2/2$:

$$\frac{\rho v^2}{2} \gg \frac{2\eta v l}{r^2}; \quad \frac{\rho v r^2}{4\eta l} \gg 1.$$

Это условие безусловно соблюдается при очень малых l, т. е. когда рассматривается очень короткий участок трубы.

При изучении ламинарных и турбулентных течений важное значение имеет безразмерная величина

$$\mathrm{Re} = \frac{\rho v r}{\eta} = \frac{v r}{\nu}, \tag{6.9}$$

называемая *числом Рейнольдса* для цилиндрической трубы. Аналогичные формулы имеются для труб или каналов разных сечений. Число Рейнольдса является весьма важной характеристикой потока при данных условиях; например, если при данных условиях это число превышает определенное критическое значение, то поток становится турбулентным; при меньших значениях течение будет ламинарным.

РАСПРЕДЕЛЕНИЕ СКОРОСТЕЙ В ЛАМИНАРНОМ ПОТОКЕ

При движении вязкой жидкости по трубам малого диаметра (когда течение ламинарное, например в капиллярах) скорость течения различна в различных местах сечения трубки. Возле стенок вследствие прилипания жидкости эта скорость равна нулю; вдоль оси трубки она имеет максимальное значение v_0. На рис. 1.75 показано распределение скоростей по диаметру трубки. Для нахождения этого распределения рассмотрим элементарный полый цилиндр малой толщины dr, выделенный в текущей жидкости. Силы вязкости F_1 и F_2, действующие по внутренней и наружной поверхностям этого цилиндра, согласно формуле (6.7), равны:

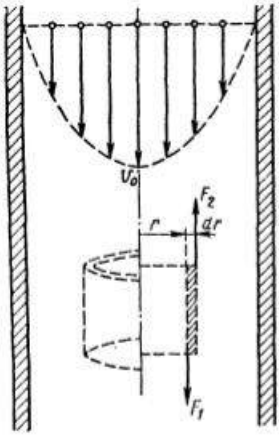


Рис. 1.75

$$F_1 = \eta \left(\frac{dv}{dr}\right)_1 S_1 = \eta \left(\frac{dv}{dr}\right)_1 \cdot 2\pi r l;$$

$$F_2 = \eta \left(\frac{dv}{dr}\right)_2 S_2 = \eta \left(\frac{dv}{dr}\right)_2 \cdot 2\pi (r + dr) l.$$

Сила F_1 направлена по течению, так как жидкость, движущаяся внутри цилиндра, имеет большую скорость. Сила же F_2 направлена против течения, так как рассматриваемый цилиндр движется быстрее, чем соприкасающиеся с ним слои жидкости, расположенные ближе к стенкам трубки. Разность этих сил $dF = F_2 - F_1$ направлена против течения и должна преодолеваться теми внешними силами, которые вызывают движение жидкости. Для расчета dF предварительно найдем связь между градиентами скорости на наружной и внутренней поверхности цилиндра:

$$\left(\frac{dv}{dr}\right)_2 = \left(\frac{dv}{dr}\right)_1 + \frac{d^2v}{dr^2} dr.$$

Поэтому

$$dF = \eta \cdot 2\pi l \left(\frac{d^2v}{dr^2} r + \frac{dv}{dr}\right) dr.$$

Если давления жидкости в сечениях A и B равны p_A и p_B, то разность «внешних» сил, действующих на рассматриваемый цилиндр,

$$dF_{\text{внеш}} = (p_A - p_B) \cdot 2\pi r \, dr = \Delta p \cdot 2\pi r \, dr.$$

При установившемся течении жидкости по трубке силы dF и $dF_{\text{внеш}}$ должны быть равны; в противном случае рассматриваемый цилиндр перемещается либо ускоренно, либо замедленно. Приравнивая эти силы, получаем дифференциальное уравнение

$$\frac{d^2v}{dr^2} + \frac{1}{r} \frac{dv}{dr} - \frac{\Delta p}{\eta l} = 0.$$

Это уравнение имеет решение (с учетом того, что у стенок трубки $r = R$, $v = 0$, а вдоль оси $r = 0$, $v = v_0$) в виде

$$v = v_0 \left(1 - \frac{r^2}{R^2}\right), \qquad (6.10)$$

т. е. распределение скоростей течения вязкой жидкости в трубке малого радиуса является *параболическим*. При этом максимальная скорость течения (вдоль оси трубки) получается равной

$$v_0 = \frac{\Delta p}{l}\frac{R^2}{4\eta}, \tag{6.11}$$

где $\Delta p/l$ — перепад давления, приходящийся на единицу длины трубки; если течение жидкости вызвано силой тяжести, то $\Delta p = \rho g l$ и тогда $v_0 = (\rho g R^2)/(4\eta)$.

Жидкость, содержащаяся в объеме рассматриваемого цилиндра, пройдет через сечение B за время $t = l/v$, следовательно, ежесекундный расход, приходящийся на долю этого цилиндра, будет равен $v \cdot 2\pi r\, \mathrm{d}r$. Общий расход жидкости через сечение трубки в единицу времени

$$Q = \int_0^R v \cdot 2\pi r\, \mathrm{d}r = 2\pi v_0 \frac{R^2}{4} = \frac{\Delta p \pi R^4}{8\eta l}. \tag{6.12}$$

Рассчитаем полную работу сил вязкости за единицу времени на единичной длине трубки. Обозначим эту величину через P. За время Δt рассматриваемый цилиндр сместится вниз на расстояние Δx и разность сил F_1 и F_2, действующих вдоль внутренней и наружной поверхности цилиндра, совершит работу

$$\mathrm{d}A = (F_2 - F_1)\,\Delta x = \eta \frac{\mathrm{d}v}{\mathrm{d}r} \cdot 2\pi r l\, \mathrm{d}r\, \Delta x.$$

Тогда

$$\mathrm{d}P = \frac{\mathrm{d}A}{l\,\Delta t} = \eta \frac{\mathrm{d}v}{\mathrm{d}r} \cdot 2\pi r\, \mathrm{d}r \frac{\Delta x}{\Delta t}.$$

Так как $\Delta x/\Delta t$ есть скорость v движения цилиндра, то, учитывая выражения (6.10) и (6.11) и интегрируя от $r = 0$ до $r = R$, получим

$$P = \frac{\pi}{30\eta}\left(\frac{\Delta p}{l}\right)^2 R^5. \tag{6.13}$$

Это есть ежесекундное количество механической энергии текущей жидкости, которое превращается в теплоту на единице длины трубки. Интересно отметить, что эта *удельная мощность диссипации* для других распределений скоростей по сечению, отличных от выражения (6.10), имеет бóльшее значение, чем для параболического. Так как распределение скоростей по формуле (6.10) соответствует установившемуся (*стационарному*) течению жидкости, то можно утверждать, что при всяком отклонении от стационарности удельная мощность диссипации будет возрастать. В гидродинамике доказано, что это правило соблюдается и для других, более сложных видов течения вязких жидкостей. Оно является частным случаем более общего закона, согласно которому *при переходе от неравновесных* (нестационарных) *состояний к равновесному* (установившемуся) *общая мощность диссипации энергии в системе стремится к минимальному значению.*

Часть II

МОЛЕКУЛЯРНАЯ ФИЗИКА И ТЕРМОДИНАМИКА

Глава 1

ОСНОВЫ МОЛЕКУЛЯРНО-КИНЕТИЧЕСКОЙ ТЕОРИИ И ТЕРМОДИНАМИКИ

§ 1. ТЕПЛОВОЕ ДВИЖЕНИЕ

Атомы и молекулы, из которых состоят различные вещества, находятся в состоянии непрерывного теплового движения.

Первой особенностью теплового движения является его *хаотичность*; ни одно направление движения молекул не выделяется среди других направлений. Поясним это: если проследить за движением одной молекулы, то с течением времени вследствие столкновений с другими молекулами величина скорости и направление движения этой молекулы изменяются совершенно беспорядочно; далее, если в какой-нибудь момент времени зафиксировать скорости движения всех молекул, то по направлению эти скорости оказываются равномерно разбросанными в пространстве, а по величине — имеют самые разнообразные значения.

Второй особенностью теплового движения является существование *обмена энергией* между молекулами, а также между различными видами движения; энергия поступательного движения молекул может переходить в энергию их вращательного или колебательного движения и обратно.

Обмен энергией между молекулами, а также между различными видами их теплового движения происходит благодаря взаимодействию молекул (столкновениям между ними). На больших расстояниях силы взаимодействия между молекулами очень малы и ими можно пренебрегать; на малых расстояниях эти силы оказывают заметное действие. В газах молекулы большую часть времени пребывают на сравнительно больших расстояниях друг от друга; лишь в течение весьма малых промежутков времени, оказавшись достаточно близко друг к другу, они взаимодействуют между собой, изменяя скорости своих движений и обмениваясь энергиями. Такие кратковременные взаимодействия молекул называются *столкновениями*. Различают два вида столкновений между молекулами:

1) *столкновения*, или удары, *первого рода*, в результате которых изменяются только скорости и кинетические энергии соударяющихся частиц; состав или структура самих молекул не испытывают никаких изменений;

2) *столкновения*, или удары, *второго рода*, в результате которых происходят изменения внутри молекул, например изменяется их состав или относительное расположение атомов внутри этих молекул. При этих столкновениях часть кинетической энергии молекул затрачивается на совершение работы против сил, действующих внутри молекул. В некоторых случаях, наоборот, может выделиться некоторое количество энергии за счет уменьшения внутренней потенциальной энергии молекул.

В дальнейшем мы будем иметь в виду только столкновения первого рода, происходящие между молекулами *газов*. Обмен энергиями при тепловых движениях в твердых и жидких телах является более сложным процессом и рассматривается в специальных разделах физики. Столкновения второго рода используются для объяснения электропроводности газов и жидкостей, а также теплового излучения тел.

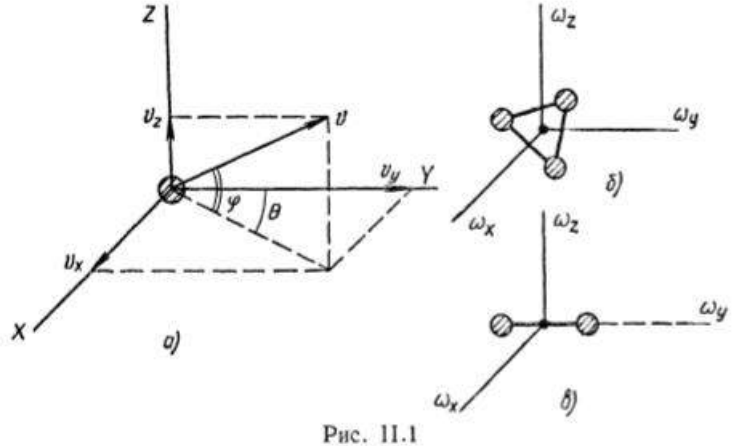


Рис. II.1

Для описания каждого вида теплового движения молекул (поступательного, вращательного или колебательного) необходимо задать ряд величин. Например, для поступательного движения молекулы необходимо знать величину и направление ее скорости. Для этой цели достаточно указать три величины: значение скорости v и два угла φ и θ между направлением скорости и координатными плоскостями или же три проекции скорости на координатные оси: v_x, v_y и v_z (рис. II.1, *а*). Заметим, что эти три величины *независимы*: при данном v углы φ и θ могут иметь любые значения и, наоборот, при заданном, например, угле φ значения v и θ могут быть любыми. Точно так же задание определенного значения v_x не накладывает никаких ограничений на значения v_y и v_z, и наоборот. Таким образом, для описания поступательного движения молекулы в пространстве необходимо задать три независимые друг от друга величины: v, φ и θ или v_x, v_y и v_z. Энергия поступательного движения молекулы будет состоять из трех независимых компонент:

$$E_{\text{поступ}} = \frac{mv_x^2}{2} + \frac{mv_y^2}{2} + \frac{mv_z^2}{2}. \qquad (1.1)$$

Для описания вращательного движения молекулы вокруг своей оси необходимо указать величину и направление угловой скорости вращения ω, т. е. опять-таки три *независимые* друг от друга величины: ω, φ и θ или ω_x, ω_y и ω_z (рис. II.1, *б*). Энергия вращательного движения молекулы также будет состоять из трех независимых компонент:

$$E_{\text{вращ}} = \frac{J_x\omega_x^2}{2} + \frac{J_y\omega_y^2}{2} + \frac{J_z\omega_z^2}{2}, \qquad (1.2)$$

где J_x, J_y и J_z — моменты инерции молекулы относительно трех взаимно перпендикулярных координатных осей. У одноатомной молекулы все эти моменты инерции очень малы, поэтому энергией ее вращательного движения пренебрегают. У двухатомной молекулы (рис. II.1, *в*) пренебрегают энергией вращательного движения относительно оси, проходящей через центры атомов, поэтому, например,

$$E_{\text{вращ}} = \frac{J_x\omega_x^2}{2} + \frac{J_z\omega_z^2}{2}. \qquad (1.3)$$

Для описания колебательного движения атомов в молекуле необходимо сначала разделить это движение на простые колебания, происходящие вдоль определенных направлений. Сложное колебание удобно разложить на простые прямолинейные колебания, происходящие по трем взаимно перпендикулярным направлениям. Эти колебания независимы друг от друга, т. е. частоте и амплитуде колебаний в одном из этих направлений могут соответствовать любая частота и амплитуда колебаний в других направлениях. Если каждое из этих прямолинейных колебаний гармоническое, то его можно описать при помощи формулы

$$x = x_0 \sin \omega t. \qquad (1.4)$$

Таким образом, для описания отдельного прямолинейного колебания атомов необходимо задать две величины: частоту колебания ω и амплитуду колебания x_0. Эти две величины также *независимы* друг от друга: при данной частоте амплитуда колебания не связывается никакими условиями, и наоборот. Следовательно, для описания сложного колебательного движения молекулы вокруг точки (т. е. своего положения равновесия) необходимо задать шесть независимых друг от друга величин: три частоты и три амплитуды колебаний по трем взаимно перпендикулярным направлениям.

Независимые друг от друга величины, определяющие состояние данной физической системы, называются степенями свободы *этой системы.* При изучении теплового движения в телах (для расчета энергии этого движения) определяют число степеней свободы каждой молекулы этого тела. При этом подсчитываются только те степени свободы, между которыми происходит обмен энергиями. Молекула одноатомного газа обладает тремя степенями свободы поступательного движения; двухатомная молекула имеет три степени свободы поступательного и две степени свободы вращательного движения (третья степень свободы, соответствующая вращению вокруг оси, проходящей через центры атомов, не учитывается). Молекулы, содержащие три

атома и больше, обладают тремя поступательными и тремя вращательными степенями свободы. Если в обмене энергиями участвует и колебательное движение, то на каждое независимое прямолинейное колебание добавляют две степени свободы.

Рассматривая раздельно поступательное, вращательное и колебательное движения молекул, можно найти *среднюю энергию, которая приходится на каждую степень свободы* этих видов движения. Рассмотрим сначала поступательное движение молекул: допустим, i-я молекула обладает кинетической энергией $mv_i^2/2$ (m — масса молекулы). Сумма $\Sigma mv_i^2/2 = E_к$ есть энергия поступательного движения всех N молекул. Разделив $E_к$ на $3N$ степеней свободы, получим среднюю энергию, приходящуюся на одну степень свободы поступательного движения молекул:

$$\varepsilon_{\text{поступ}} = \frac{E_к}{3N} = \frac{1}{3N}\sum\frac{mv_i^2}{2}. \qquad (1.5)$$

Так же можно рассчитать средние энергии, приходящиеся на одну степень свободы вращательного $\varepsilon_{\text{вращ}}$ и колебательного $\varepsilon_{\text{колеб}}$ движений. Если каждая молекула обладает $i_п$ степенями свободы поступательного, $i_в$ степенями свободы вращательного и $i_к$ степенями свободы колебательного движений, то полная энергия теплового движения всех N молекул будет равна

$$U = N(i_п\varepsilon_{\text{поступ}} + i_в\varepsilon_{\text{вращ}} + i_к\varepsilon_{\text{колеб}}). \qquad (1.6)$$

В теоретической физике (где разработаны основы молекулярно-кинетической теории) установлено, что средние энергии всех видов теплового движения молекул связаны с температурой, причем для разреженных газов с достаточным приближением можно полагать, что на каждую степень свободы поступательного и вращательного движений молекул в среднем приходится одна и та же энергия, равная

$$\varepsilon = \frac{kT}{2}, \qquad (1.7)$$

где $k = 1{,}38 \cdot 10^{-23}$ Дж/К (*постоянная Больцмана*), T — *абсолютная температура газа*. Связь между средней энергией простого колебательного движения молекул (т. е. происходящего в одном из направлений) и абсолютной температурой имеет более сложный вид:

$$\varepsilon = \frac{h\nu}{e^{h\nu/kT} - 1}, \qquad (1.8)$$

где $h = 6{,}626 \cdot 10^{-34}$ Дж·с (*постоянная Планка*), ν — частота колебаний; k — постоянная Больцмана. При высоких температурах и малых частотах колебаний, когда произведение kT значительно больше, чем $h\nu$, можно разложить экспоненциальную функцию в ряд и тогда формула (1.8) дает приближенное выражение $\varepsilon \approx kT$. Так как простое колебание обладает двумя степенями свободы, то на каждую степень свободы придется энергия, равная $kT/2$. Таким образом, при высоких температурах и малых частотах колебаний на каждую степень свободы колебательного движения приходится столько же энергии ($kT/2$),

сколько и на одну степень свободы поступательного или вращательного движения. При таких условиях можно очень просто рассчитать суммарную энергию беспорядочного движения частиц системы, содержащей N молекул:

$$U = N\frac{i}{2}kT, \tag{1.9}$$

где i — число степеней свободы, которым обладает одна молекула данной системы.

§ 2. ВЗАИМОДЕЙСТВИЕ МОЛЕКУЛ; ЭНЕРГИЯ СВЯЗИ МОЛЕКУЛ

Характер и величина силы взаимодействия между молекулами зависят от расстояния между ними. Установлено, что на малых расстояниях молекулы отталкиваются друг от друга, а на больших — притягиваются. На рис. II.2 показано изменение сил взаимодействия F двух изолированных молекул в зависимости от расстояния r между

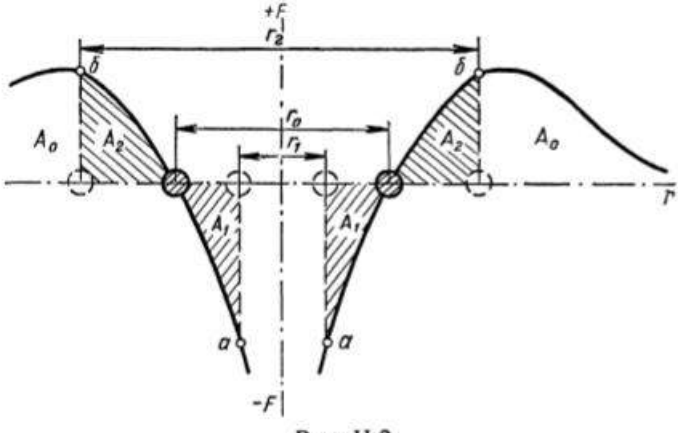


Рис. II.2

ними (для простоты рассуждений молекулы предполагаются одинаковыми). Такой характер зависимости F от r можно объяснить, если предположить, что между молекулами одновременно существуют *силы отталкивания*, быстро убывающие с расстоянием, и *силы притяжения*, также ослабляющиеся с расстоянием, но медленнее. На расстоянии $r = r_0$ эти силы равны, т. е. их равнодействующая равна нулю. При $r < r_0$ преобладают силы отталкивания, при $r > r_0$ — силы притяжения.

Легко показать, что при таком характере сил взаимодействия молекулы будут совершать колебательные движения. Действительно, допустим, что при $r = r_0$, когда силы взаимодействия между молекулами равны нулю, молекулы движутся навстречу друг другу со скоростями, равными v. Тогда по мере уменьшения расстояния между молекулами силы отталкивания будут совершать отрицательную работу, уменьшая кинетическую энергию их, и на некотором расстоянии

$r = r_1$ молекулы, потеряв свою энергию, остановятся. Можно утверждать, что при этом кинетическая энергия обеих молекул перешла в потенциальную энергию их взаимодействия.

В масштабе чертежа, приведенного на рис. II.2, эта энергия, равная работе сил отталкивания на участке от r_1 до r_0, изобразится заштрихованными площадками A_1. В дальнейшем силы отталкивания отбрасывают молекулы друг от друга и потенциальная энергия этой системы переходит обратно в кинетическую энергию. На расстоянии $r = r_0$ потенциальная энергия системы становится равной нулю (так как равны нулю силы взаимодействия), т. е. целиком переходит в кинетическую энергию удаляющихся друг от друга молекул. При увеличении расстояния между ними ($r > r_0$) начнут действовать силы притяжения, которые остановят эти молекулы на некотором расстоянии $r = r_2$. В этом состоянии опять-таки кинетическая энергия молекул целиком превращена в потенциальную; на рис. II.2 эти энергии изобразятся заштрихованными площадками A_2 (очевидно, $A_1 = A_2$). В последующем начнется обратное движение молекул и процесс повторяется: расстояние между молекулами периодически изменяется от r_1 до r_2.

Описанное выше колебательное движение молекул не является, вообще говоря, гармоническим, так как действующие на молекулы силы не удовлетворяют условию линейности $F = -kx$ (k — коэффициент упругой силы, $x = r - r_0$ — смещение молекул от положения равновесия). Лишь при малых амплитудах колебаний, когда $r_2 - r_0$ и $r_0 - r_1$ достаточно малы и отрезок кривой $F(r)$ на участке ab может быть заменен отрезком прямой, колебания будут близки к гармоническим.

Мы рассматривали колебательное движение двух изолированных молекул, находящихся достаточно далеко от других молекул. Действительная картина колебательных движений в веществе гораздо сложнее, так как каждая молекула одновременно взаимодействует со многими другими молекулами. При этом в конечном счете движение молекул является *беспорядочным*, т. е. направления колебательных движений изменяются с течением времени совершенно беспорядочно.

Для характеристики различных состояний вещества важно установить, насколько прочно связана между собой каждая пара взаимодействующих друг с другом молекул. Очевидно, эту связь можно оценить по величине работы, которую необходимо затратить, чтобы «оторвать» молекулы друг от друга, или, как говорят, удалить их на бесконечно большие расстояния друг от друга. На рис. II.2 эта работа в масштабе чертежа изображается площадками A_0; она зависит от величины наибольшего расстояния r_2, на которое при колебаниях удаляются друг от друга взаимодействующие молекулы. Чем больше r_2, тем слабее силы притяжения между молекулами и, следовательно, тем меньшую работу необходимо затратить для разъединения этих молекул.

Работа, необходимая для отрыва молекул друг от друга, может быть сообщена им различными способами; в частности, связь между данной парой взаимодействующих молекул может быть разрушена

тепловым столкновением с другими молекулами. Вероятность такого события особенно велика, если средняя кинетическая энергия теплового движения молекул превосходит по величине работу отрыва молекул друг от друга. Так как средняя кинетическая энергия теплового движения пропорциональна температуре, то при очень высоких температурах тепловые столкновения разрушают всякую связь между молекулами; это означает, что при достаточно высоких температурах вещество будет находиться в газообразном состоянии. При низких температурах, когда работа отрыва молекул друг от друга значительно больше средней энергии их теплового движения, вероятность разрушения связи между молекулами очень мала и эта связь может сохраняться долго: вещество находится в твердом состоянии. Жидкое же состояние является промежуточным.

§ 3. АГРЕГАТНЫЕ СОСТОЯНИЯ И ФАЗОВЫЕ ПЕРЕХОДЫ. РАВНОВЕСНЫЕ И НЕРАВНОВЕСНЫЕ СОСТОЯНИЯ И ПРОЦЕССЫ

Всевозможные состояния, в которых может находиться то или иное вещество, прежде всего разбиваются на так называемые *агрегатные состояния*: твердое, жидкое и газообразное. Эти состояния выделяются по основным физическим свойствам вещества.

Следует учесть, что у некоторых веществ нет резкой границы между различными агрегатными состояниями. Например, при нагревании стекла происходит постепенное его размягчение и невозможно установить, когда оно переходит из твердого состояния в жидкое. При очень большом внешнем давлении твердые металлы начинают «течь», т. е., подобно жидкости, принимают форму сосуда, в котором они находятся.

Различные состояния одного и того же вещества можно отличать друг от друга также и по значениям физических величин, которые характеризуют эти состояния, например по значениям объема, температуры и давления. Поэтому каждому агрегатному состоянию вещества соответствует бесконечное множество различных состояний, которые отличаются друг от друга различными значениями объема, давления, температуры и других физических величин; при изменении этих величин вещество переходит из одного состояния в другое, оставаясь твердым, жидким или газообразным.

УРАВНЕНИЕ СОСТОЯНИЯ

Физические величины, характеризующие то или иное состояние вещества, иногда называют *параметрами состояния*. Основными параметрами являются объем V, внешнее давление p и температура T. Если между параметрами состояния существует какое-нибудь определенное однозначное соотношение, которое сохраняется при переходе из одного состояния в другое, то это соотношение называется у р а в н е н и е м с о с т о я н и я. Например, для разреженных газов соблюдается уравнение

$$\frac{pV}{T} = \text{const}. \qquad (1.10)$$

Это соотношение связывает между собой значения объема, давления и температуры для множества отличающихся друг от друга состояний данной массы газообразного вещества. Для других агрегатных состояний — твердого и жидкого — такие простые соотношения между параметрами не найдены.

Следует различать два вида переходов вещества из одного состояния в другое:

1) переходы, при которых меняются только численные значения параметров, характеризующих состояние вещества (объема, давления температуры и др.), но состав, строение вещества, его агрегатное состояние не изменяются. Такими переходами являются сжатие, расширение, нагревание газов, а также твердых и жидких тел при условии, что в них не происходит изменений в составе, структуре и физических свойствах;

2) переходы, при которых или происходит изменение агрегатного состояния вещества или существенно изменяются его состав, строение и физические свойства. Такие переходы называются фазовыми переходами; к ним относятся испарение и конденсация, плавление и отвердевание, кристаллизация и перекристаллизация, переходы металлов в сверхпроводящее состояние и т. д.

При некоторых фазовых переходах происходит изменение агрегатного состояния вещества, например вещество из твердого состояния переходит в жидкое или газообразное и наоборот. При этом изменяются взаимное расположение молекул, расстояние между ними, характер их теплового движения. Постоянство состава вещества при фазовых переходах такого рода можно иллюстрировать на примере воды: таяние льда или испарение воды не сопровождается изменением состава молекулы воды (H_2O).

При других фазовых переходах агрегатное состояние вещества сохраняется, но в веществе происходят существенные изменения в строении или структуре (взаимном расположении молекул), вследствие чего заметно изменяются физические свойства вещества. К таким переходам относятся: потеря ферромагнитных свойств железом при нагревании его до температуры 780 °С и выше, переход некоторых металлов в сверхпроводящее состояние при очень низких температурах (когда электрическое сопротивление этих металлов скачком уменьшается до нуля), переход кристаллического вещества из одной модификации в другую и т. д.

СВОЙСТВА ТЕРМОДИНАМИЧЕСКИХ СИСТЕМ

Термодинамической системой называется всякая физическая система, состоящая из большого числа частиц — атомов и молекул, которые совершают бесконечное тепловое движение и, взаимодействуя между собой, обмениваются энергиями. Такими термодинамическими системами, и притом простейшими, являются газы, молекулы которых совершают беспорядочное поступательное и вращательное движения и при столкновениях обмениваются кинетическими энергиями. Термодинамическими системами являются также твердые

и жидкие вещества. Молекулы твердых тел совершают беспорядочные колебания вокруг своих положений равновесия; обмен энергиями между молекулами происходит благодаря их непрерывному взаимодействию, вследствие чего смещение одной молекулы от своего положения равновесия немедленно отражается на расположении и скорости движения соседних молекул. Так как средняя энергия теплового движения молекул, согласно формулам (1.7) и (1.8), связана с температурой, то температура является важнейшей физической величиной, характеризующей различные состояния термодинамических систем. Кроме температуры состояния таких систем определяются также и объемом, который они занимают, и внешним давлением или внешними силами, действующими на систему.

Важным свойством термодинамических систем является существование у них *равновесных состояний*, в которых они могут пребывать сколь угодно долго. Если на термодинамическую систему, находящуюся в одном из равновесных состояний, оказать некоторое внешнее воздействие и затем прекратить его, то система самопроизвольно переходит в новое равновесное состояние. Однако следует подчеркнуть, что тенденция к переходу в равновесное состояние действует всегда и непрерывно, даже в течение того времени, когда система подвергается внешнему воздействию. Эта тенденция или, точнее, постоянное существование процессов, ведущих к достижению равновесных состояний, является важнейшей особенностью термодинамических систем.

Для газа, заключенного в некотором сосуде, равновесным является состояние, в котором температура, давление и плотность (или число молекул в единице объема) в пределах объема газа везде одинаковы. Если в каком-нибудь месте этого объема вызвать местное нагревание или сжатие, то в системе начнется процесс выравнивания температуры и давления; этот процесс будет происходить и в течение того времени, пока имеется внешнее воздействие, однако только после прекращения этого воздействия процесс выравнивания приведет систему к новому равновесному состоянию.

Состояния изолированных термодинамических систем, которые, несмотря на отсутствие внешних воздействий, не сохраняются в течение конечных промежутков времени, называются *неравновесными*. Система, первоначально находящаяся в неравновесном состоянии, с течением времени переходит в равновесное состояние. Время перехода из неравновесного состояния в равновесное называется *временем релаксации*. Обратный переход из равновесного состояния в неравновесное может быть осуществлен при помощи внешних воздействий на систему. Неравновесным является, в частности, состояние системы с различными температурами в различных местах; выравнивание температуры в газах, твердых и жидких телах есть переход этих тел в равновесное состояние с одинаковой температурой в пределах объема тела. Другой пример неравновесного состояния можно привести, рассматривая двухфазные системы, состоящие из жидкости и ее пара. Если над поверхностью жидкости, находящейся в закрытом сосуде, имеется ненасыщенный пар, то состояние системы неравновесное: число молекул N_1, вылетающих в единицу времени из жидкости, больше, чем число моле-

кул N_2, возвращающихся за это же время из пара в жидкость. Вследствие этого с течением времени число молекул в парообразном состоянии увеличивается (т. е. увеличивается плотность пара) до тех пор, пока не установится равновесное состояние с $N_1 = N_2$.

Переход от неравновесного состояния в равновесное в большинстве случаев происходит *непрерывно*, причем скорость этого перехода можно при помощи соответствующего внешнего воздействия плавно регулировать, сделав процесс релаксации либо очень быстрым, либо очень медленным. Так, например, путем механического перемешивания можно заметно повысить скорость выравнивания температуры в жидкостях или газах; охлаждая жидкость, можно сделать очень медленным процесс диффузии растворенного в ней вещества, и т. п.

Для некоторых систем существуют такие состояния, называемые *метастабильными*, в которых эти системы могут находиться относительно долгое время, но как только на систему будет оказано внешнее воздействие определенного характера, происходит самопроизвольный скачкообразный переход к равновесному состоянию. В этих случаях внешнее воздействие лишь открывает возможность к переходу в равновесное состояние. Например, достаточно чистая вода при медленном подводе тепла может быть нагрета до температуры на несколько градусов выше температуры кипения. Это состояние воды является метастабильным; если встряхнуть такую воду (или внести небольшое число пылинок — центров образования пузырьков пара), она со взрывом закипает и ее температура скачком понижается до температуры кипения. Таким образом, метастабильное состояние характеризуется тем, что при выводе из этого состояния система не только не возвращается к ней, но, наоборот, еще более отходит от нее, скачком переходя в существующее для этой системы равновесное состояние.

ВИДЫ ТЕРМОДИНАМИЧЕСКИХ ПРОЦЕССОВ

Переход физической системы из одного («начального») состояния в другое («конечное») через какую-то последовательность промежуточных состояний называется п р о ц е с с о м. Однако при классификации процессов, происходящих в объеме данной термодинамической системы, необходимо учитывать также и те изменения, которые происходят в окружающих телах (взаимодействующих с данной системой). Процесс называется *обратимым*, если выполняются два условия:

1) если изменения в системе можно провести в обратном направлении через те же промежуточные состояния, через которые проходила система в прямом направлении;

2) если при обратном переходе не только сама система, но и все связанные с нею окружающие тела в точности возвращаются в первоначальное состояние.

Процесс называется *равновесным*, если начальное, конечное и все промежуточные состояния системы являются равновесными. Таким образом, для равновесности процесса, происходящего в н у т р и термодинамической системы, существование или отсутствие «остаточных изменений» в окружающих телах не имеет значения; важно только,

чтобы каждое из промежуточных состояний системы было равновесным.

Для иллюстрации рассмотрим процесс расширения и сжатия газа, заключенного в цилиндре с поршнем (рис. II.3).

Если поршень смещается вправо или влево очень медленно, то давление и температура газа в различных местах объема газа успевают выравниваться: $p_1 = p_2$; $T_1 = T_2$ и, следовательно, каждое промежуточное состояние можно считать с удовлетворительной точностью равновесным. Такие процессы можно провести как в одном (например, расширение), так и в обратном (сжатие) направлениях через одни и те же промежуточные состояния с одинаковыми давлениями и температурами по всему объему тела.

При быстром сжатии и расширении промежуточные состояния не будут равновесными. При быстром сжатии давление и температура вблизи поршня (p_1 и T_1) больше, чем вдали от поршня (p_2 и T_2), так как для выравнивания давления и температуры всегда требуется некоторое время. При быстром расширении, наоборот, давление и температура вблизи поршня меньше, чем вдали. Таким образом, промежуточные состояния в обоих процессах оказываются неравновесными вследствие того, что процессы выравнивания температур и давлений не происходят «мгновенно».

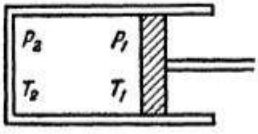


Рис. II.3

Скорость изменения состояния термодинамической системы определяется не только скоростью внешнего воздействия (в данном примере — скоростью поршня, изменяющего объем газа), но и скоростью внутренних процессов выравнивания температур и давлений (т. е. скоростью релаксации). Вопрос о том, является ли изучаемый процесс «медленным» или «быстрым», зависит от соотношения между скоростями внешнего воздействия и релаксации. Промежуточные состояния могут быть равновесными только в двух предельных случаях: 1) если скорость внешнего воздействия бесконечно мала и 2) если скорость процессов релаксации бесконечно велика.

Примером необратимых процессов являются процессы расширения или сжатия, происходящие при наличии т р е н и я. Рассмотрим еще раз расширение и сжатие газов в цилиндре с поршнем (рис. II.3). Если бы эти процессы происходили равновесно и без трения, то работа, совершаемая газом при расширении (A_0), в точности равнялась бы внешней работе, необходимой для сжатия. При наличии же трения (даже если оба процесса происходят достаточно медленно) работа, совершаемая газом при расширении, будет меньше A_0, а работа внешних сил, затрачиваемая на сжатие газа, будет больше, чем A_0. Обозначим через Q количество теплоты, которое выделилось при трении поршня о стенки цилиндра в процессе расширения. Для простоты рассуждений допустим, что эта теплота идет только на нагревание цилиндра и поршня. Для того чтобы процесс сжатия был в точности обратным процессу расширения, необходимо, чтобы при сжатии теплота Q была отнята от цилиндра и поршня, превращена в механическую энергию

и передана тому «механизму», который производит сжатие газа. Такой способ возвращения к первоначальному состоянию оказывается невозможным; поршень и цилиндр нагреваются также и при сжатии, а в окружающей среде фиксируются «остаточные изменения» — превращение некоторого количества механической энергии в теплоту (важно подчеркнуть, что теплота, выделившаяся при трении, не может быть превращена в механическую энергию без новых «остаточных изменений» в окружающей среде; см. ч. II, § 7).

Таким образом, все процессы, происходящие при наличии трения, являются необратимыми. Превращение механической энергии в тепловую при трении является односторонним процессом; его невозможно провести в обратном направлении, при котором теплота, выделившаяся при трении, могла бы превратиться в механическую работу без каких-либо остаточных изменений в системе и в окружающих телах.

Другим важным примером необратимых процессов является теплообмен между телами, имеющими различные температуры. Допустим, что в течение «прямого» процесса между двумя какими-нибудь телами, входящими в состав системы, существует конечная разность температур и теплота переходит от тела с высокой температурой к телу с низкой температурой. При «обратном» процессе теплота, полученная холодным телом, должна быть возвращена горячему телу, с тем чтобы было восстановлено первоначальное состояние системы. Путем одной только теплопроводности такая передача теплоты от холодных тел к горячим невозможна.

Обратимые процессы имеют большое значение в теоретической термодинамике как идеальные процессы перехода систем из одного состояния в другое. Перечислим основные условия, необходимые для того, чтобы процесс был обратимым:

1) каждое промежуточное состояние системы должно быть равновесным;

2) в системе должно отсутствовать внутреннее трение, т. е. одностороннее превращение механической энергии в тепловую;

3) в системе не должны происходить односторонние химические реакции, например горение;

4) разность температур между соприкасающимися телами внутри системы, а также между системой и окружающими телами должна быть бесконечно малой. В частности, если система получает теплоту из окружающей среды, то температура источника тепла должна быть больше температуры системы также на бесконечно малую величину. Благодаря этому процесс теплопередачи протекает бесконечно медленно и поэтому будет равновесным и обратимым процессом.

§ 4. ВНУТРЕННЯЯ ЭНЕРГИЯ ТЕРМОДИНАМИЧЕСКОЙ СИСТЕМЫ. ПЕРВЫЙ ЗАКОН ТЕРМОДИНАМИКИ

Термодинамическая система, как и любая другая физическая система, обладает некоторым запасом энергии, который обычно называют *внутренней энергией системы*. Внутренняя энергия системы есть

сумма всех видов кинетической и потенциальной энергии всех составных частей системы: молекул, атомов, электронов и т. д. Таким образом, в состав внутренней энергии входит кинетическая энергия поступательного и вращательного движений атомов и молекул, энергия их колебательного движения, потенциальная энергия взаимодействия атомов и молекул, кинетическая и потенциальная энергия электронов в атомах, внутриядерная энергия. Однако в большинстве физических явлений, в которых участвуют термодинамические системы, не все перечисленные виды энергии испытывают изменения. Например, при сжатии, расширении или нагревании газообразных тел изменяются только интенсивности поступательного и вращательного движений их молекул; внутриатомная энергия в таких процессах не участвует. В химических процессах остается без изменения внутриядерная энергия; ее изменения наблюдаются только в явлениях радиоактивности и в ядерных реакциях.

Поэтому очень часто, употребляя понятие внутренней энергии, имеют в виду не полную энергию данной системы, а только ту ее часть, которая участвует и изменяется в рассматриваемых явлениях.

Внутренняя энергия системы является однозначной функцией ее состояния, т. е. в каждом определенном состоянии система обладает вполне определенным значением внутренней энергии. Однако при данной внутренней энергии система может находиться в различных состояниях. Внутренняя энергия системы, в частности термодинамической, может быть выражена (рассчитана) в зависимости от значений всех физических величин (параметров), определяющих это состояние: объема, давления, температуры и т. д. Расчет внутренней энергии тел, находящихся в твердом или жидком состоянии, затруднен и требует использования ряда упрощающих предположений. Имеется довольно простая формула только для расчета внутренней энергии разреженного газа в зависимости от его температуры (для сильно сжатых газов эта энергия зависит еще и от объема). Ее можно получить на основании следующих рассуждений.

Допустим, газ сильно разрежен, так что его молекулы в среднем находятся далеко друг от друга и весьма слабо взаимодействуют между собой. При этих условиях потенциальной энергией взаимодействия молекул можно пренебречь и тогда внутренняя энергия газа определяется только кинетической энергией теплового движения его молекул. Если газ не только разрежен, но и имеет высокую температуру, то можно воспользоваться формулами (1.7) и (1.9):

$$U = N \frac{i}{2} kT. \qquad (1.11)$$

Эта формула применима также для расчета энергии теплового движения молекул в простых кристаллических веществах (алмаз, металлы), но только при высоких температурах и низких частотах колебаний молекул, когда сохраняет свою силу формула (1.7); в этом случае $i = 6$ (так как молекулы совершают независимые колебания по трем взаимно перпендикулярным направлениям в пространстве, а каждому из этих колебаний приписываются две степени свободы).

Допустим, что система переходит из состояния с энергией U_1 в состояние с энергией U_2. Изменение внутренней энергии системы может произойти, если:

1) система получает извне или отдает окружающим телам некоторую энергию ΔE в какой-нибудь форме или

2) система совершает работу A против действующих на нее внешних сил (или внешние силы, изменяя относительное расположение составных частей системы, совершают некоторую работу A).

Условимся *внешней работой* называть ту работу, которую совершают силы, приложенные системой к окружающим телам. Эта работа может быть и положительной и отрицательной. Например, если газ, заключенный в цилиндре с поршнем (рис. II.3), расширяется, то сила давления, приложенная со стороны газа к внешнему телу — поршню, совершает *положительную* работу. При сжатии газа сила давления на поршень совершает *отрицательную* работу, так как эта сила действует в направлении, обратном движению поршня (в этом случае положительную работу совершают внешние силы, действующие на систему).

На основании закона сохранения энергии можно написать:

$$U_2 - U_1 = \Delta E - (\pm A). \tag{1.12}$$

Внутренняя энергия системы уменьшается ($U_2 < U_1$), если система отдает в окружающую среду энергию ΔE, а также если система совершает положительную работу A. Внутренняя энергия системы увеличивается, если она получает энергию извне и если положительную работу совершают внешние силы, действующие на систему.

В технической термодинамике рассматриваются такие процессы, в результате которых система получает извне или отдает в окружающую среду некоторое количество энергии в виде т е п л о т ы Q, а также совершается положительная или отрицательная внешняя работа A. Тогда в уравнении (1.12) можно заменить ΔE на Q и переписать его в виде (знаки «$\pm$» обычно не пишут)

$$Q = (U_2 - U_1) + A, \tag{1.13}$$

или для бесконечно малых значений этих величин

$$\mathrm{d}Q = \mathrm{d}U + \mathrm{d}A. \tag{1.14}$$

Следует иметь в виду, что в этих формулах теплота должна быть подставлена с положительным знаком, если она поступает в систему извне, и с отрицательным знаком, если она отдается системой другим телам.

Соотношения (1.13) и (1.14), выражающие закон сохранения энергии в применении к термодинамическим процессам, являются математической формулировкой п е р в о г о з а к о н а т е р м о д и н а м и к и: *при переходе термодинамической системы из одного состояния в другое изменение ее внутренней энергии равно разности между количеством получаемой или отдаваемой теплоты* ($\pm Q$) *и внешней работы* ($\pm A$), *совершаемой при этом системой*:

$$U_2 - U_1 = Q - A; \quad \mathrm{d}U = \mathrm{d}Q - \mathrm{d}A. \tag{1.15}$$

Заметим, что в этих формулах теплота Q, внутренняя энергия U и внешняя работа A должны выражаться в одних и тех же единицах (джоулях). Если теплота выражена в калориях, то для перевода в джоули пользуются соотношением

$$1 \text{ кал} = 4{,}1868 \text{ Дж}.$$

§ 5. ВНЕШНЯЯ РАБОТА СИСТЕМЫ И ТЕПЛООБМЕН С ОКРУЖАЮЩЕЙ СРЕДОЙ. ТЕПЛОЕМКОСТЬ ТЕЛ

Первый закон термодинамики связывает между собой: 1) изменение внутренней энергии системы, 2) количество поступающей в систему или отводимой от нее теплоты и 3) совершаемую внешнюю работу. Эти величины должны быть выражены в зависимости от изменения параметров системы при переходе от начального состояния в конечное. Так как внутренняя энергия системы есть однозначная функция состояния, то, зная параметры начального и конечного состояний системы,

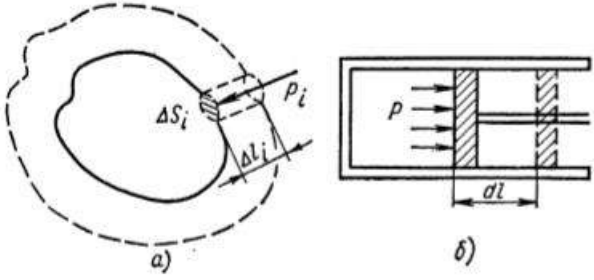


Рис. II.4

можно рассчитать энергию системы U_1 и U_2 в этих состояниях и затем найти изменение этой энергии $U_2 - U_1$. Такой расчет может быть легко произведен только для разреженных газов, для которых существует простое выражение (1.11) для внутренней энергии.

Внешняя работа, совершаемая системой, может быть вычислена в зависимости от значения внешних сил, действующих на систему, и от величины деформации системы — изменения ее формы и размеров. Если внешние силы приложены по поверхности в виде, например, внешнего давления, сжимающего систему, то расчет внешней работы может быть произведен в зависимости от изменения объема системы.

Допустим, что внешнее давление на систему в месте расположения площадки ΔS_i равно p_i (рис. II.4, *а*). Если при расширении системы эта площадка сместилась на Δl_i, то в этом месте совершается работа $\Delta A_i = p_i \Delta S_i \Delta l_i = p_i \Delta V_i$, где ΔV_i — увеличение объема системы, вызванное смещением площадки ΔS_i на Δl_i. Вся работа расширения системы равна сумме этих элементарных работ: $\Delta A = \Sigma p_i \Delta V_i$. Если во всех участках поверхности системы внешнее давление одно и то же, как это, например, имеет место при расширении газа, заключенного

в цилиндре с поршнем (рис. II.4, *б*), то внешняя работа

$$dA = F\,dl = pS\,dl = p\,dV, \tag{1.16}$$

где S — площадь поршня, а $Sdl = dV$ — изменение объема системы. При расширении системы (например, движении поршня в цилиндре) внешнее давление не всегда остается постоянным, поэтому внешняя работа, совершаемая системой при изменении ее объема от V_1 до V_2, должна рассчитываться как сумма:

$$A = \sum_{V_1}^{V_2} p_i \Delta V_i \quad \text{или} \quad A = \int_{V_1}^{V_2} p\,dV. \tag{1.17}$$

Лишь в частном случае, когда внешнее давление все время поддерживается постоянным, работа будет равна

$$A = p\,(V_2 - V_1). \tag{1.18}$$

Заметим, что величина внешней работы, которую совершает система при определенном изменении ее объема, зависит от того, происходит это изменение объема равновесным или неравновесным образом. Допустим, например, что расширение газа в цилиндре с поршнем (см. рис. II.3) происходит медленно (равновесно) и газ совершает работу A_0. При быстром (неравновесном) расширении давление газа на поршень в каждом его промежуточном положении будет меньше, чем при равновесном расширении, и поэтому совершаемая газом работа будет меньше, чем A_0. При быстром сжатии все промежуточные значения давления газа на поршень будут больше, чем при равновесном сжатии, и поэтому затрачиваемая работа внешних сил будет больше, чем A_0.

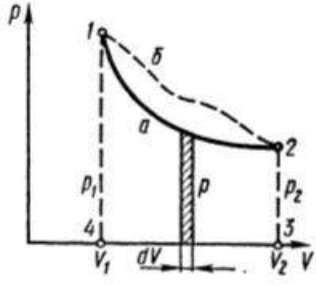


Рис. II.5

Следует подчеркнуть также, что при равновесном переходе системы из одного определенного состояния в другое может быть совершена различная работа в зависимости от того, как изменяется давление при изменении объема, т. е. через какую последовательность промежуточных состояний осуществляется этот переход. На рис. II.5 в координатах p и V изображены два различных процесса, переводящие систему из определенного начального состояния (*1*) с давлением p_1 и объемом V_1 в одно и то же конечное состояние (*2*) с давлением p_2 и объемом V_2. Заштрихованная площадка в масштабе чертежа изображает элементарную работу, совершаемую системой при расширении на dV. Площадь, охватываемая контуром *1 — a — 2 — 3 — 4*, показывает работу $A = \int_{V_1}^{V_2} p\,dV$, совершаемую системой при переходе из первого состояния во второе через последовательность состояний *1 — a — 2*. Если же переход осуществляется через последовательность состояний *1 — б —*

— 2, то внешняя работа, изображаемая площадью *1 — 6 — 2 — 3 — 4*, будет больше. Таким образом, параметры начального (p_1, V_1) и конечного (p_2, V_2) состояний системы не определяют величину совершаемой внешней работы; необходимо знать еще и функцию $p\,(V)$, показывающую, как изменяется давление в процессе перехода.

Имеется бесчисленное множество различных равновесных процессов, переводящих систему из одного состояния в другое; каждому из этих процессов соответствует определенная работа. Ввиду этого величину элементарной (бесконечно малой) работы $p\mathrm{d}V$ иногда обозначают не $\mathrm{d}A$, а δA, подчеркивая этим зависимость работы (см. формулу (1.17)) от характера изменения давления в изучаемом процессе $p = p\,(V)$. Однако если эту зависимость всегда иметь в виду, то можно избежать недоразумений и не вводя нового обозначения.

Теплообмен между системой и окружающей средой также зависит не только от параметров начального и конечного состояний системы, но и от той последовательности промежуточных состояний, через которые пробегает система. Это следует из первого закона термодинамики $Q = U_2 - U_1 + A$, где U_1 и U_2 определяются только заданием параметров начального и конечного состояний, а внешняя работа A зависит, кроме того, еще и от самого процесса перехода. Вследствие этого теплота Q не может быть выражена в зависимости только от температуры начального и конечного состояний системы.

Допустим, что некоторое количество газа переходит от состояния с параметрами p_1, V_1 и T_1 в состояние с параметрами p_2, V_2 и T_2, причем газ получил (или отдал) теплоту Q и совершил внешнюю работу A. Можно ли утверждать, что изменение температуры газа от T_1 до T_2 произошло только вследствие теплообмена с окружающей средой? Очевидно, нет, так как изменение температуры могло быть частично вызвано тем, что газ совершил внешнюю работу (изменение температуры может быть вызвано и без теплообмена с окружающей средой: сжимая газ, можно его нагреть, расширяя — охладить). Поэтому, наблюдая изменение температуры газа, мы должны еще выяснить, какая часть этого изменения вызвана притоком или отводом теплоты, а какая часть — внешней работой расширения или сжатия. Таким образом, при переходе системы из одного состояния в другое теплообмен между системой и окружающей средой зависит еще и от величины внешней работы, совершаемой при данном переходе. Ввиду этого элементарное количество теплоты при теплообмене иногда обозначают не $\mathrm{d}Q$, а δQ и записывают первый закон термодинамики (1.14) в виде

$$\delta Q = \mathrm{d}U + \delta A.$$

Однако в таком изменении записи нет необходимости, если во всех дальнейших расчетах и рассуждениях всегда будет учитываться зависимость Q и A от той последовательности состояний, через которые проходит система. Если система не совершает внешней работы, то количество поступающей в систему или отводимой от нее теплоты будет равно изменению внутренней энергии системы:

$$A = 0; \quad Q = U_2 - U_1; \quad \mathrm{d}Q = \mathrm{d}U. \tag{1.19}$$

В этом случае количество теплоты может быть представлено в зависимости от изменения температуры системы; для однородного вещества пользуются формулой

$$Q = Mc(T_2 - T_1); \quad dQ = Mc\,dT, \tag{1.20}$$

где M — масса вещества, а c — величина, характеризующая данное вещество и называемая *удельной теплоемкостью* этого вещества:

$$c = \frac{1}{M}\frac{dQ}{dT}. \tag{1.21}$$

Измерения показали, что теплоемкость для данного вещества не есть постоянная величина, а заметно изменяется вместе с температурой (возрастает с увеличением температуры). Теоретический расчет теплоемкости твердых и жидких веществ представляет значительные трудности; просто удается рассчитывать теплоемкости только для газов.

Иногда формулу (1.20), выражающую количество теплоты в зависимости от изменения температуры системы, применяют и в тех случаях, когда внешняя работа A не равна нулю, т. е. когда система (например, газ) сжимается или расширяется и, следовательно, часть изменения температуры вызывается совершаемой при этом внешней работой. В этом случае теплоемкость системы c уже не характеризует только физические свойства системы, но зависит еще и от того, какая внешняя работа совершается при данном изменении температуры системы.

§ 6. ИЗОПРОЦЕССЫ; АДИАБАТИЧЕСКИЙ ПРОЦЕСС. ЭНТРОПИЯ СИСТЕМЫ. ЗАМКНУТЫЕ (КРУГОВЫЕ) ПРОЦЕССЫ

Среди всевозможных равновесных процессов, происходящих с термодинамическими системами, выделяют так называемые *изопроцессы*, при которых один из основных параметров состояния (объем, давление, температура) сохраняется постоянным:

1) изохорический процесс, при котором объем системы остается постоянным ($V = \text{const}$). При изохорическом процессе система не совершает внешней работы, поэтому, согласно формуле (1.13), изменение внутренней энергии системы равно количеству теплоты, полученной или отданной системой;

2) изобарический процесс, при котором давление, оказываемое со стороны системы на окружающие тела, остается постоянным ($p = \text{const}$). При изобарическом процессе внешняя работа рассчитывается по простой формуле (1.18);

3) изотермический процесс, при котором температура системы остается постоянной ($T = \text{const}$). Если такой процесс происходит с газом, то постоянство температуры означает, согласно формуле (1.11), постоянство внутренней энергии; однако для других систем при $T = \text{const}$ $U \neq \text{const}$.

К этим основным процессам добавляется еще и

4) адиабатический процесс, при котором на протяжении всего процесса теплообмен с окружающей средой отсутствует ($dQ = 0$; $Q = 0$). Из первого закона термодинамики следует, что при

адиабатических процессах

$$Q = U_2 - U_1 + A = 0; \quad U_1 - U_2 = A, \tag{1.22}$$

т. е. внешняя работа совершается за счет внутренней энергии системы. Если система совершает положительную внешнюю работу, то ее внутренняя энергия уменьшается на эквивалентную величину, и, наоборот, если система совершает отрицательную работу, то ее внутренняя энергия увеличивается (за счет работы внешних сил).

ЭНТРОПИЯ

Адиабатические процессы в термодинамических системах могут быть равновесными и неравновесными. Для характеристики равновесного адиабатического процесса можно ввести некоторую физическую величину, которая оставалась бы *постоянной* в течение всего процесса; ее назвали *энтропией*. Условились обозначать энтропию через S и определять ее следующим образом: *энтропия есть такая функция состояния системы, элементарное изменение которой при равновесном переходе системы из одного состояния в другое равно полученному или отданному количеству теплоты, деленному на температуру, при которой произошел этот процесс*; для бесконечно малого изменения состояния системы

$$dS = \frac{dQ}{T}. \tag{1.23}$$

При переходе системы из одного равновесного состояния *1* в другое состояние *2* через последовательность промежуточных равновесных же состояний изменение энтропии рассчитывается по формуле

$$S_2 - S_1 = \int_1^2 \frac{dQ}{T}. \tag{1.24}$$

При равновесных адиабатических процессах $dQ = 0$, $S_2 = S_1$; следовательно, равновесные адиабатические процессы есть *изоэнтропические* процессы, при которых энтропия не изменяется ($S = \text{const}$). При других процессах, в частности при изохорических, изобарических и изотермических процессах, энтропия изменяется. Для равновесных изотермических процессов $T = \text{const}$ и расчет изменения энтропии дает простой результат:

$$S_2 - S_1 = \frac{1}{T} \int_1^2 dQ = \frac{Q}{T}. \tag{1.25}$$

Во всех случаях, когда система получает извне теплоту, Q положительно, следовательно, $S_2 > S_1$ и энтропия системы увеличивается. Если же система отдает теплоту, то Q имеет отрицательный знак и, следовательно, $S_2 < S_1$; энтропия системы уменьшается.

Если термодинамическая система переходит из одного состояния в другое через *неравновесные* промежуточные состояния (например,

при расширении газа в вакуум), то изменение энтропии возможно и при отсутствии теплообмена. Ввиду этого изменение энтропии в общем случае записывается в виде неравенства Клаузиуса:

$$S_2 - S_1 \geqslant \int_1^2 \frac{dQ}{T},$$

где знак равенства относится к равновесным процессам. Например, при адиабатических процессах ($dQ = 0$) энтропия остается постоянной только в том случае, если этот процесс протекает равновесно; в неравновесных адиабатических процессах энтропия всегда возрастает.

Заметим, что энтропия системы пропорциональна массе (или числу частиц) этой системы. Масса системы представляется в виде суммы масс ее составных частей, поэтому *энтропия всей системы будет равна сумме энтропий ее составных частей*, т. е. энтропия есть аддитивная величина.

Формулы (1.23) и (1.24) дают возможность рассчитывать только изменения энтропии системы, а не саму энтропию. Однако для термодинамических расчетов этого вполне достаточно; нас почти всегда интересует переход систем из одного состояния в другое, и поэтому в окончательные расчетные формулы входит не абсолютное значение энтропии, а лишь ее изменение в рассматриваемом процессе. То же самое можно сказать и о внутренней энергии системы: точное значение внутренней энергии термодинамических систем с учетом внутриатомной и внутриядерной энергии почти нигде не используется; практическое значение имеют только изменения внутренней энергии при переходе системы из одного состояния в другое.

ГРАФИЧЕСКОЕ ИЗОБРАЖЕНИЕ ИЗОПРОЦЕССОВ

Изопроцессы могут быть изображены графически в координатных системах, по осям которых отложены параметры состояния. В частности, если эти изопроцессы совершаются газообразными телами, то можно ограничиться плоской координатной системой, по осям которой откладывают давление и объем, или температуру и объем, или, наконец, температуру и давление. Третий параметр состояния газа (при постоянной массе) можно рассчитать при помощи уравнения состояния ($pV/T = \text{const}$), поэтому задание только двух параметров газа полностью определяют его состояние. На рис. II.6 приведены графики изопроцессов в координатах p и V, проведенные из одного и того же начального состояния. Кривая адиабатического процесса («адиабата») идет круче, чем кривая изотермического процесса («изотерма»). Это обстоятельство можно легко объяснить на основании

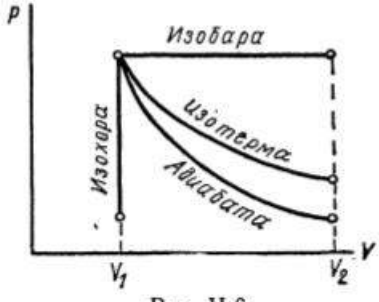


Рис. II.6

уравнения состояния для газов

$$\frac{p_1V_1}{T_1} = \frac{p_2V_2}{T_2} = \text{const}.$$

Разность давлений при расширении газа от объема V_1 до объема V_2 равна

$$p_1 - p_2 = \text{const}\left(\frac{T_1}{V_1} - \frac{T_2}{V_2}\right). \qquad (1.26)$$

При адиабатическом расширении внешняя работа совершается только за счет внутренней энергии газа, вследствие чего внутренняя энергия, а вместе с ней и температура газа уменьшаются ($T_2 < T_1$);

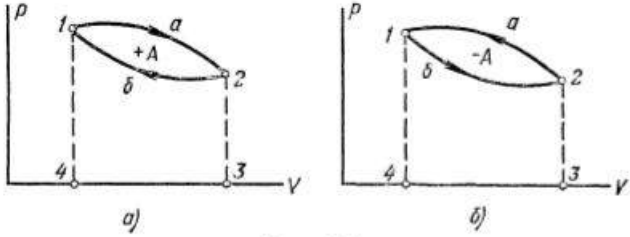


Рис. II.7

при изотермическом процессе $T_2 = T_1$. Поэтому в формуле (1.26) разность давлений $p_1 - p_2$ будет при адиабатическом расширении больше, чем при изотермическом.

КРУГОВЫЕ ПРОЦЕССЫ

Важное значение в термодинамике имеют так называемые «круговые» или «замкнутые» процессы, в результате которых система в точности возвращается в исходное состояние. Графически в координатных системах, по осям которых отложены параметры состояния, круговые процессы изображаются замкнутыми кривыми; так, например, для газообразных тел в координатной системе p, V круговой процесс изобразится плоской замкнутой кривой (рис. II.7).

Круговой процесс называется *равновесным*, если все промежуточные состояния, через которые пробегает система, являются равновесными, и *обратимым*, если его можно провести в обратном направлении через ту же самую последовательность состояний, как и в прямом направлении, и притом без каких-либо «остаточных» изменений в окружающей среде. В дальнейшем рассматриваются только равновесные и обратимые круговые процессы.

Применим к круговым процессам первый закон термодинамики; так как начальное и конечное состояния системы совпадают, то $U_1 = U_2$ и $U_2 - U_1 = 0$, т. е. при круговых процессах изменение внутренней энергии системы равно нулю. Тогда из первого закона термодинамики (1.13) следует:

$$Q = A, \qquad (1.27)$$

т. е., *выполняя круговой (замкнутый) процесс, система совершает внешнюю работу за счет полученной извне теплоты,* т. е. превращает теплоту, полученную извне, в механическую работу, совершаемую против внешних сил. В этом и заключается важное значение круговых процессов в термодинамике. Так как, совершая круговой процесс, система в точности возвращается в исходное состояние с тем же значением внутренней энергии, то такой процесс можно многократно повторять, каждый раз превращая некоторое количество теплоты, полученное извне, в механическую работу без каких-либо изменений самой системы. Круговые процессы иногда называются *циклами*.

Круговой процесс можно вести как в одном, так и в противоположном направлении. Рассмотрим сначала круговой процесс, изображенный на рис. II.7, *а*; стрелками показано направление этого процесса. На этом графике площадь *1 — а — 2 — 3 — 4 —1* в масштабе чертежа изображает положительную внешнюю работу (A_1), которую совершает система над внешними силами при расширении, а площадь *2 — б — — 1 — 4 — 3 — 2* изображает отрицательную работу (A_2), совершаемую системой при сжатии (напомним, что отрицательная работа системы означает, что положительную работу совершают внешние силы). Суммируя эти две работы A_1 и A_2, получим, что в результате такого кругового процесса система совершает положительную работу $A = A_1 - A_2$ (так как $A_1 > A_2$), которая графически изображается площадью *1 — а — 2 — б — 1*. Согласно же равенству (1.27), система для совершения положительной работы должна получить извне эквивалентное количество теплоты $Q = A$. Таким образом, замкнутые процессы, идущие в направлении *1 — а — 2 — б — 1*, могут быть использованы в тепловых машинах, предназначенных для превращения теплоты в механическую работу.

Весьма важно подчеркнуть, что в соотношении $Q = A$, полученном из первого закона термодинамики для круговых процессов, Q есть разность между теплотой Q_1, полученной извне, и теплотой Q_2, отданной обратно в окружающую среду:

$$Q = Q_1 - Q_2.$$

Установлено (см. ч. II, § 7), что, совершая круговой процесс *1 — а — — 2 — б — 1*, термодинамическая система не может полностью превратить всю полученную извне теплоту Q_1 в механическую работу; часть этой теплоты система обязательно возвращает обратно в окружающую среду. В связи с этим условились называть *коэффициентом полезного действия кругового процесса* (вернее, коэффициентом использования теплоты) отношение того количества теплоты, которое превращается в механическую работу ($A = Q = Q_1 - Q_2$), ко всему количеству теплоты Q_1, получаемому системой извне:

$$\eta = \frac{A}{Q_1} = \frac{Q_1 - Q_2}{Q_1} = 1 - \frac{Q_2}{Q_1}. \qquad (1.28)$$

Коэффициент полезного действия неравновесных круговых процессов меньше, чем равновесных. Действительно, при неравновесном

расширении, как это было показано в § 5, система совершает меньшую работу A_1, чем при равновесном расширении; далее, на неравновесное сжатие системы внешние силы должны затратить бо́льшую работу A_2, чем на равновесное сжатие. Вследствие этого полезная работа $A = A_1 - A_2$ при неравновесном круговом процессе оказывается меньше, чем при равновесном.

Рассмотрим теперь круговой процесс (рис. II.7, *б*), протекающий в обратном направлении: *1 — б — 2 — а — 1*. Расширение системы происходит вдоль *1 — б — 2*, а сжатие — вдоль *2 — а — 1*, поэтому положительная работа A_1, совершаемая системой при расширении (площадь *1 — б — 2 — 3 — 4 — 1*), будет меньше, чем отрицательная работа A_2 при сжатии (площадь *2 — а — 1 — 4 — 3 — 2*). Для такого кругового процесса результирующая работа $A = A_1 - A_2$ отрицательна. Согласно же первому закону термодинамики для круговых процессов (1.27) отрицательной внешней работе должен соответствовать отрицательный теплообмен ($-A = -Q$), т. е. система, совершающая круговой процесс в направлении *1 — б — 2 — а — 1*, должна отдать в окружающую среду большее количество теплоты ($-Q_2$), чем она получает извне ($+Q_1$):

$$Q = Q_1 - Q_2; \quad Q_1 < Q_2.$$

Таким образом, при обратном круговом процессе система получает извне теплоту Q_1 и механическую работу A, а отдает в окружающую среду теплоту $Q_2 = Q_1 + A$. Так как, затрачивая внешнюю механическую работу A, можно при сжатии нагреть систему (например, газ) до высокой температуры, то теплоту Q_2 можно отдавать окружающим телам, имеющим высокую температуру, тогда как теплоту Q_1 можно отнимать у тел, имеющих низкую температуру. Многократно повторяя такой цикл, вызывают постепенное охлаждение тел, от которых система отнимает теплоту Q_1. Следовательно, круговые процессы, идущие в направлении *1 — б — 2 — а — 1*, могут быть использованы в холодильных машинах, в которых затрачивается механическая работа для охлаждения тел.

При подробном исследовании термодинамических процессов необходимо интересоваться, при какой температуре происходит теплообмен между системой и окружающими телами. Разделим круговой процесс на элементарные участки; на некоторых участках система получает теплоту извне (эти теплоты обозначим через ΔQ_1), на других — отдает теплоту в окружающую среду (эти теплоты обозначим через ΔQ_2). Температуры, при которых система получает теплоты ΔQ_1, обозначим через T_1, а температуры, при которых система отдает наружу теплоты ΔQ_2, — через T_2. Изучение всевозможных равновесных круговых процессов показало, что сумма отношений $\Delta Q_1/T_1$, подсчитанных для участков процесса, при которых система получает теплоту, всегда равна сумме отношений $\Delta Q_2/T$, вычисленных для участков процесса, при которых система отдает теплоту:

$$\sum \frac{\Delta Q_1}{T_1} = \sum \frac{\Delta Q_2}{T_2}. \qquad (1.29)$$

Допустим для простоты, что система, совершающая равновесный круговой процесс, получает теплоту Q_1 при постоянной температуре T_1 и отдает теплоту Q_2 при постоянной температуре T_2, т. е. те участки кругового процесса, в течение которых происходит теплообмен между системой и окружающими телами, являются изотермами. Тогда

$$\sum \frac{\Delta Q_i}{T_1} = \frac{Q_1}{T_1}; \quad \sum \frac{\Delta Q_2}{T_2} = \frac{Q_2}{T_2},$$

а из соотношения (1.29) следует:

$$\frac{Q_1}{T_1} = \frac{Q_2}{T_2}; \quad \frac{Q_2}{Q_1} = \frac{T_2}{T_1}.$$

Подставив этот результат в выражение (1.28) для коэффициента полезного действия кругового процесса, получим для этого частного случая

$$\eta = 1 - \frac{T_2}{T_1} = \frac{T_1 - T_2}{T_1}. \tag{1.30}$$

Так как $Q_2 < Q_1$, то $T_2 < T_1$, т. е. система должна получать теплоту Q_1 от тел, имеющих высокую температуру T_1, и отдавать теплоту Q_2 окружающим телам, имеющим низкую температуру T_2.

§ 7. ВТОРОЙ ЗАКОН ТЕРМОДИНАМИКИ. ЗАКОН ВОЗРАСТАНИЯ ЭНТРОПИИ

В § 4 была указана математическая формулировка первого закона термодинамики в виде соотношений (1. 13) и (1.14):

$$dU = dQ - dA.$$

Эквивалентным и более удобным выражением этого закона является утверждение:

внутренняя энергия термодинамической системы есть однозначная функция ее состояния.

В каждом определенном состоянии термодинамическая система имеет одно определенное значение внутренней энергии. Поэтому если система, имея начальное значение энергии U, совершает круговой процесс и в точности возвращается в исходное состояние, то сумма элементарных изменений энергии dU, положительных на одних участках ($dU > 0$) и отрицательных на других участках ($dU < 0$), будет для всего кругового процесса равна нулю. Это утверждение символически записывается в виде

$$\oint dU = 0; \quad \oint (dQ - dA) = 0. \tag{1.31}$$

В конце § 7 ч. II было приведено соотношение (1.29), которое также соблюдается в круговых процессах; перепишем его в виде

$$\sum \left(\frac{\Delta Q_1}{T_1} + \frac{\Delta Q_2}{T_2} \right) = \sum \frac{\Delta Q_i}{T_i} = 0. \tag{1.32}$$

Разбивая круговой процесс на бесконечно малые участки, можем записать это соотношение в более точном виде:

$$\oint \frac{dQ}{T} = 0. \tag{1.33}$$

Напомним, что в формулах (1.32) и (1.33) элементарные количества теплоты dQ положительны, если они поступают в систему извне, и отрицательны, если отнимаются от нее. Первоначально соотношение (1.33) было получено для некоторых (сравнительно несложных) круговых процессов, совершаемых простыми термодинамическими системами. Например, в следующем параграфе оно будет получено для цикла Карно, совершаемого идеальным газом. Однако во всех случаях, где можно было проверить формулы (1.32) и (1.33) или вытекающие из них следствия, никаких расхождений не было обнаружено. Это дало основание для предположения, что, по-видимому, формулы (1.32) или (1.33) выражают некоторый закон термодинамики. Называя dQ/T элементарным приращением энтропии (см. формулу (1.23)), можно утверждать, что в результате кругового процесса суммарное изменение энтропии системы, согласно (1.33), равно нулю. Таким образом, был установлен второй закон термодинамики:

энтропия термодинамической системы есть однозначная функция ее состояния.

В каждом определенном состоянии термодинамическая система имеет одно определенное значение энтропии, поэтому при круговых процессах, когда система в точности возвращается в исходное состояние, изменение энтропии равно нулю, т. е.

$$\oint dS = \oint \frac{dQ}{T} = 0; \quad \oint \frac{dU + dA}{T} = 0. \tag{1.34}$$

На тех участках процесса, на которых система получает теплоту извне, dQ имеет положительный знак и энтропия увеличивается; на других участках, где теплота отнимается от системы ($dQ < 0$), энтропия уменьшается. В любом круговом процессе независимо от последовательности пробегаемых состояний увеличение энтропии от притока теплоты полностью компенсируется уменьшением энтропии от отдачи теплоты системой.

Таким образом, первый и второй законы термодинамики имеют аналогичные формулировки: если система совершает замкнутый процесс (и, следовательно, в точности возвращается в первоначальное состояние), то

$$\begin{aligned} &\oint dU = 0 - \text{первый закон}, \\ &\oint dS = 0 - \text{второй закон}. \end{aligned} \tag{1.35}$$

Из соотношений (1.32) или (1.33) можно получить ряд следствий; приведем некоторые из них. Прежде всего заметим, что *круговые процессы, в течение которых система только получает, но не отдает теплоту, невозможны.* Действительно, для таких круговых процессов

все элементарные количества теплоты ΔQ были бы положительными и тогда в противоречии с соотношением (1.32) сумма $\Sigma \Delta Q/T$ была бы отлична от нуля (отрицательные члены этой суммы — $\Delta Q/T$, соответствующие отдаче тепла от системы в окружающую среду, отсутствовали бы). Заметим, что такие процессы не противоречили бы первому закону термодинамики, т. е. закону сохранения энергии, так как система, получив извне теплоту Q_1, могла бы совершить эквивалентное количество положительной внешней работы $A = Q_1$ и вернуться в исходное состояние с тем же значением внутренней энергии.

Очевидно, *невозможны также и круговые процессы, в течение которых система только отдавала бы, но не получала теплоты*; тогда в сумме $\Sigma \Delta Q/T$ присутствовали бы только отрицательные и отсутствовали положительные члены и эта сумма не равнялась бы нулю. Такие процессы также не противоречили бы закону сохранения энергии, так как работой внешних сил можно было бы полностью компенсировать убыль энергии системы от отдачи теплоты в окружающую среду ($-A = -Q_2$) и система опять-таки могла бы вернуться в исходное состояние с общим балансом энергии, равным нулю.

Таким образом, мы видим, что осуществить такой круговой процесс, в результате которого система полностью превращала бы в механическую работу всю полученную извне теплоту Q_1, невозможно; часть этой теплоты Q_2 система обязательно возвращает в окружающую среду. Пользуясь понятием коэффициента использования теплоты при круговых процессах (см. выражение (1.28)), можно утверждать, что этот коэффициент никогда не может быть равен единице.

Допустим, что получение теплоты извне и отдача ее окружающим телам происходят изотермически. В этом случае, как было показано в предыдущем параграфе, из соотношения (1.29) следует:

$$Q_1/T_1 = Q_2/T_2 \quad \text{или} \quad Q_1/Q_2 = T_1/T_2,$$

т. е. система, совершающая круговой процесс в прямом направлении ($Q_1 > Q_2$), получает теплоту Q_1 от тел, имеющих высокую температуру T_1, и отдает теплоту Q_2 телам, имеющим низкую температуру T_2. Таким образом, для превращения теплоты в механическую работу при помощи круговых процессов необходимо существование *разности температур* между окружающими систему телами. Существенно, что в результате кругового процесса не только происходит превращение некоторого количества теплоты $Q = Q_1 - Q_2$ в механическую работу, но обязательно имеет место «непроизводительный» переход теплоты Q_2 от горячих тел к холодным. Количество этой «неиспользованной теплоты» зависит от разности температур между телами, участвующими в теплообмене с системой. Чем больше разность между температурой T_1, при которой система получает теплоту Q_1, и температурой T_2, при которой отдается неиспользуемая теплота Q_2, тем выше коэффициент полезного действия кругового процесса $\eta = (Q_1 - Q_2)/Q_1 = (T_1 - T_2)/T_1$ и, следовательно, тем меньше Q_2. В частности, если холодное тело («холодильник»), которому система могла бы отдать теплоту Q_2, имеет все время температуру, равную абсолютному нулю (т. е. $T_2 = 0$), то $\eta = 1$ и $Q_2 = 0$.

Из этих рассуждений можно сделать заключение, что возможны только такие процессы, в результате которых система получает от окружающих тел, имеющих высокие температуры, некоторое количество теплоты Q_1, часть этой теплоты превращает в механическую работу A, а остальную часть $Q_2 = Q_1 - A$ возвращает в окружающую среду — телам, имеющим низкие температуры.

Выше речь шла о круговых процессах, идущих в прямом направлении (см. рис. II.7, *а*). В результате круговых процессов, идущих в обратном направлении (см. рис. II.7, *б*), система получает теплоту Q_1 от окружающих тел, имеющих низкие температуры, использует работу A, совершаемую внешними силами, для повышения температуры и затем передает теплоту $Q_2 = Q_1 + A$ телам, имеющим высокие температуры.

Заметим, что теплота связана с беспорядочным движением частиц системы, а механическая работа обусловлена упорядоченным движением этих частиц. Если, например, газ, перемещая поршень в цилиндре (см. ч. II, § 3, рис. II.3), совершает механическую работу, то это означает, что из кинетической энергии беспорядочного движения молекул газа выделяется некоторая часть, которая передается поршню, т. е. переходит в энергию упорядоченного движения поршня. Можно утверждать вообще, что превращение теплоты в механическую работу есть переход энергии беспорядочного движения молекул системы в энергию упорядоченного движения. В связи с этим второй закон термодинамики можно понимать иначе:

при помощи круговых процессов, совершаемых какой-нибудь термодинамической системой, невозможно полностью превратить энергию беспорядочного теплового движения частиц в энергию упорядоченного движения тел.

Формулы (1.31), (1.33) и (1.34) иллюстрируют существенное отличие величин U и S от величин Q и A. Внутренняя энергия U и энтропия S характеризуют *состояние* термодинамической системы, тогда как Q и A характеризуют *процесс* обмена энергией между данной системой и окружающей средой и вовсе не являются функциями состояния этой системы. Нельзя утверждать, что, находясь в определенном состоянии, термодинамическая система имеет определенный «запас теплоты» или «запас работы». Если бы это было так, то при совершении кругового процесса система должна была бы вернуться в первоначальное состояние с теми же «запасами» Q и A, следовательно, суммарные изменения этих величин должны были бы равняться нулю. Однако в круговых процессах

$$\oint dQ \neq 0; \quad \oint dA \neq 0.$$

Далее следует отметить, что при переходе из одного состояния в другое изменения внутренней энергии и энтропии

$$\int_1^2 dU = U_2 - U_1; \quad \int_1^2 dS = S_2 - S_1$$

не зависят от последовательности промежуточных состояний, через которые происходит этот переход, тогда как результирующий теплообмен и суммарная внешняя работа

$$\int_1^2 dQ = Q; \quad \int_1^2 dA = A$$

зависят от этой последовательности. Переход термодинамической системы из одного фиксированного состояния в другое может быть осуществлен различным образом, через множество различных процессов перехода; для всех этих процессов изменения внутренней энергии и энтропии *одинаковы*, а теплообмен и внешняя работа *различны*. Таким образом, по изменениям $U_2 - U_1$ и $S_2 - S_1$ нельзя отличить один процесс перехода от другого; для этого необходимо дополнительно указать величину теплообмена или внешней работы.

Рассмотрим два типа задач, решаемых с помощью первого и второго законов термодинамики:

1) допустим, что нам известны внутренняя энергия и энтропия в начальном и конечном состояниях системы: U_1, S_1 и U_2, S_2. Необходимо рассчитать теплообмен Q и внешнюю работу A при переходе из первого состояния во второе. Очевидно, что одного первого закона термодинамики $Q = U_2 - U_1 + A$ недостаточно, так как в это соотношение входят обе неизвестные величины. Задача может быть решена, если, например, воспользоваться вторым законом термодинамики и рассчитать теплообмен; так как $dS = dQ/T$, $dQ = TdS$, то

$$Q = \int_1^2 T\, dS.$$

Однако для вычисления этого интеграла недостаточно задать S_1 и S_2; необходимо знать все промежуточные значения температуры и элементарные изменения энтропии *в течение всего процесса*, т. е. необходимо выделить рассматриваемый процесс перехода среди множества других возможных процессов (разумеется, если бы нам было известно, как изменяются давление и объем в рассматриваемом переходе, то можно было бы определить сначала внешнюю работу $A = \int_1^2 p\, dV$, а затем, воспользовавшись первым законом термодинамики, рассчитать теплообмен);

2) допустим, что нам известны начальные значения внутренней энергии U_1, энтропии S_1, теплообмен Q и внешняя работа A; требуется определить конечное состояние системы. В этом случае U_2 найдем из первого закона термодинамики ($U_2 = U_1 + Q - A$). Однако этого недостаточно для полной характеристики конечного состояния (при данной внутренней энергии U система может находиться в различных состояниях). Можно дополнительно определить энтропию конечного состояния S_2, воспользовавшись вторым законом термодинамики:

$$S_2 = S_1 + \int_1^2 \frac{dQ}{T}.$$

Для вычисления интеграла здесь также необходимо знание промежуточных этапов перехода из начального состояния в конечное. Однако при этом нет необходимости знать изменения всех параметров, характеризующих систему, достаточно значений теплообмена dQ и температуры T. Это обстоятельство облегчает исследование и расчеты в тех случаях, когда известно, например, поступление теплоты в систему (нагрев электрическим током, сгорание топлива или химические реакции, сопровождающиеся выделением теплоты, и т. п.) и фиксируется изменение только одного параметра системы — температуры.

ТЕПЛОВЫЕ И ХОЛОДИЛЬНЫЕ МАШИНЫ

Первый и второй законы термодинамики определяют работу периодически действующих тепловых и холодильных машин, предназначенных для превращения теплоты в механическую работу и обратно. Поэтому иногда этим законам дают другие формулировки, которые, по существу, совпадают с приведенными выше, но в отличие от них содержат конкретные упоминания области применения этих законов.

Первый закон термодинамики утверждает:

невозможно построить такую периодически действующую тепловую машину, которая, не получая никакой энергии извне, совершала бы некоторое количество внешней работы и возвращалась точно в исходное состояние.

Многократно повторяя такой цикл, можно было бы совершить в окружающей среде любое количество внешней работы без каких-либо последствий для машины и без какого-либо расхода энергии из окружающей среды. Машины, которые могли бы работать таким образом, называются *вечными двигателями* (или «перпетуум мобиле») *первого рода*. Первый закон термодинамики утверждает, что *перпетуум мобиле первого рода невозможен*.

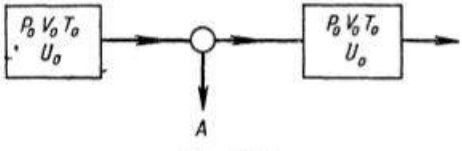


Рис. 11.8

Такой невозможный цикл работы тепловой машины схематически изображен на рис. 11.8. Здесь p_0, V_0, T_0 — основные параметры состояния машины, U_0 — ее внутренняя энергия. При этом предполагается, что в результате каждого такого цикла внутри машины никаких остаточных изменений нет и машина вместе со всеми находящимися внутри нее телами в точности возвращается в исходное состояние.

Однако первый закон термодинамики допускает существование любых тепловых машин, которые совершали бы механическую работу за счет полученной извне теплоты. Второй закон термодинамики ограничивает возможности превращения теплоты в механическую работу:

невозможно построить такую периодически действующую тепловую машину, которая, получив извне некоторое количество теплоты при любой температуре, целиком превращала бы ее в механическую работу и при этом возвращалась точно в исходное состояние.

Многократно повторяя такой цикл, можно было бы превратить в механическую работу огромные запасы теплоты, имеющиеся в морях и океанах. Машины, которые могли бы выполнять такую задачу, называются *вечными двигателями* (или «перпетуум мобиле») *второго рода*. Цикл работы таких машин изображен схематически на рис. II.9.

При этом предполагается, что в результате каждого такого цикла внутри машины никаких остаточных изменений не происходит, а в окружающей среде только исчезает теплота Q и совершается механическая работа $A = Q$. Таким образом, второй закон термодинамики утверждает, что *перпетуум мобиле второго рода невозможен*.

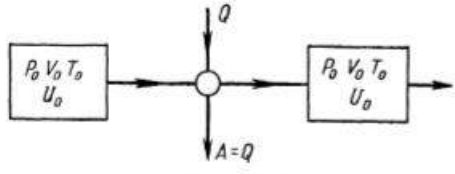


Рис. II.9

Первый и второй законы термодинамики допускают периодическую работу тепловых машин только по схеме, приведенной на рис. II.10.

Источник теплоты, откуда машина берет энергию Q_1, должен иметь более высокую температуру, чем холодильник, которому машина отдает неиспользованную часть энергии Q_2.

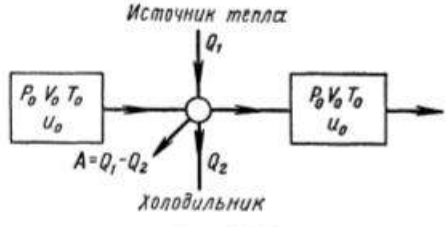


Рис. II.10

Идеальными тепловыми машинами называются такие машины, которые работают на равновесных и обратимых циклах. Такие машины, если их работу провести в обратном направлении, могут отнять у холодильника теплоту Q_2, израсходовать ту же самую работу A, совершаемую теперь внешними силами, и вернуть нагревателю теплоту $Q_1 = Q_2 + A$, которая была взята у него при работе в прямом направлении. При этом изменения, происходящие в машине и в окружающей среде при обратном цикле, точно повторяют изменения, которыми сопровождался прямой цикл. Можно показать, что при заданных температурах источника теплоты T_1 и холодильника T_2 коэффициент полезного действия

$$\eta = \frac{Q_1 - Q_2}{Q_1}$$

всех идеальных тепловых машин, работающих как в прямом, так и в обратном направлениях, независимо от их конструкции, состава рабочего вещества и характера совершаемого в машине цикла *одинаков*

и равен

$$\eta = \frac{T_1 - T_2}{T_1}.$$

Если бы имелись две идеальные тепловые машины с различными коэффициентами полезного действия, то, заставляя одну из этих машин работать в прямом, а вторую — в обратном направлении, можно было бы осуществить перпетуум мобиле второго рода, что невозможно.

Машины, работающие по обратному циклу (т. е. забирающие теплоту от тел с низкой температурой и отдающие ее телам с высокой температурой), позволяют более экономно расходовать энергию, предназначенную для нагревания, например для отопления жилых и производственных зданий и т. д. Допустим, что для этой цели имеется некоторое количество электрической энергии W. Если всю эту энергию превратить в теплоту непосредственно, то будет получено $Q = W$ джоулей теплоты. Однако можно этой энергией привести в действие машину, работающую по обратному циклу. Такая машина получит из окружающей среды теплоту Q_2 при низкой температуре и передаст потребителю $Q_1 = Q_2 + A = Q_2 + W$ джоулей теплоты при необходимой высокой температуре. Таким образом, без машины потребитель получил бы W, а при наличии машины он получает $W + Q_2$ джоулей теплоты, причем на работу самой машины никакая другая энергия, кроме W, не расходуется. Можно утверждать, что в данном случае машина работает как «тепловой насос, выкачивающий теплоту из резервуара с низкой температурой и накачивающий его в резервуар с высокой температурой».

ЭНТРОПИЯ И ВЕРОЯТНОСТЬ

Второй закон термодинамики имеет также более общую формулировку, применимую не только к равновесным и обратимым, но и к любым другим термодинамическим процессам. Эта формулировка основана на новом определении понятия энтропии, которое охватывает все возможные состояния термодинамических систем, а для равновесных состояний совпадает с выражениями (1.23) или (1.35). Такое определение энтропии было предложено Л. Больцманом:

$$S = k \ln W, \tag{1.36}$$

где k — постоянная Больцмана; W — число, иногда называемое *термодинамической вероятностью* данного состояния системы. Поясним смысл этого числа на одном сравнительно простом примере. Допустим, что имеется идеальный газ, содержащий N пронумерованных частиц. В некотором состоянии этого газа N_1 частиц, имеющие номера от единицы до N_1 и разбросанные по всему объему газа, обладают одинаковыми энергиями ε_1; следующие N_2 частиц (с номерами от $N_1 + 1$ до $N_1 + N_2$) имеют энергии ε_2 и т. д. Таким образом, общее число частиц газа $N = \Sigma N_i$, а полная энергия $U = \Sigma \varepsilon_i N_i$. Состояние, при котором в каждом определенном месте объема системы находится частица с определенным номером и энергией, будем называть *микро-*

состоянием. Допустим теперь, что в системе произошел следующий процесс: вследствие теплового движения две частицы, имеющие разные номера, но одинаковые энергии, поменялись местами, которые они занимали в объеме системы. При этом микросостояние системы изменится, однако при помощи приборов (термометров, манометров и т. п.), для которых номер частицы безразличен, такое изменение не может быть обнаружено. Следовательно, несмотря на наличие в системе теплового движения, общее термодинамическое состояние системы, которое мы характеризуем распределением температуры, давления и плотности по объему, при таких перестановках сохранится неизменным. Состояние, фиксируемое термодинамическими величинами, называют *макросостоянием*. Очевидно, одному определенному макросостоянию системы может соответствовать большое число микросостояний, которые отличаются друг от друга перестановками частиц, имеющих одинаковые энергии. Это число определяется по формуле

$$W = \frac{N!}{N_1!\,N_2!\ldots}. \tag{1.37}$$

В частности, если макросостояние является равновесным, т. е. температура, давление и плотность газа в пределах его объема имеют одинаковые значения, то W имеет *наибольшее* значение по сравнению с любыми другими (неравновесными) состояниями. Поэтому переход газа из неравновесного состояния в равновесное сопровождается увеличением W, а следовательно, и энтропии.

Доказательство того, что определение энтропии на основании формул (1.36) и (1.37) совпадает для равновесных состояний с термодинамическим определением (1.23) и (1.24), требует сложных вычислений и приводится в статистической физике.

Формулировка второго закона термодинамики дается в виде двух взаимосвязанных утверждений:

---

1) *в каждом определенном состоянии — равновесном или неравновесном — термодинамическая система имеет только одно значение энтропии.*

Изменение энтропии при равновесных и обратимых переходах может быть вычислено по формулам:

$$dS = \frac{dQ}{T}; \quad S_2 - S_1 = \int_1^2 \frac{dQ}{T}.$$

Энтропия есть аддитивная величина: энтропия системы равна сумме энтропий ее составных частей;

---

2) *при переходе термодинамических систем из неравновесного состояния в равновесное энтропия увеличивается.*

В равновесном состоянии энтропия достигает наибольшего значения и в дальнейшем, при отсутствии внешних воздействий, сохраняется

постоянной. Это утверждение является весьма важной частью второго закона термодинамики и называется з а к о н о м в о з р а с т а н и я э н т р о п и и.

Необходимо иметь в виду, что энтропия термодинамической системы изменяется, во-первых, вследствие теплообмена между системой и окружающей средой и, во-вторых, под действием тех внутренних (односторонних, необратимых) процессов, которые переводят систему из неравновесных состояний в равновесное. Так как эти внутренние («релаксационные», т. е. ведущие к равновесным состояниям) процессы действуют всегда, то тенденция к увеличению энтропии существует не только у изолированной системы, но и при наличии любых внешних воздействий на систему.

Таким образом, закон возрастания энтропии выражает известную из опыта необратимость реальных термодинамических процессов. Важное значение имеет скорость возрастания энтропии: в изолированных системах она характеризует интенсивность происходящих внутри них необратимых процессов.

Глава 2

ИДЕАЛЬНЫЙ ГАЗ

§ 8. ОСНОВНОЕ УРАВНЕНИЕ КИНЕТИЧЕСКОЙ ТЕОРИИ ГАЗОВ. СРЕДНИЕ СКОРОСТИ МОЛЕКУЛ

Газообразное состояние вещества характеризуется тем, что мельчайшие частицы вещества — атомы или молекулы — бóльшую часть времени пребывают сравнительно далеко друг от друга. Силы взаимодействия между ними оказывают свое действие только в течение весьма коротких промежутков времени, когда частицы газа сталкиваются между собой. Поэтому действие молекулярных сил выражается только в обмене энергиями при столкновениях. Чем меньше плотность газа, тем больше свободный пробег его молекул и, следовательно, тем меньше влияния оказывают молекулярные силы на общее поведение газа при тех или иных изменениях его состояния.

Вследствие этого при изучении свойств разреженных газов можно пренебрегать действием молекулярных сил; так как размеры молекул очень малы, то можно пренебрегать также и собственным объемом частиц газа по сравнению с объемом сосуда, в котором находится газ. Таким образом, мы приходим к представлению об *идеальном газе*, теория которого оказывается очень простой. Теоретически обосновав свойства идеального газа, можно затем перейти к теории реальных газов, внеся поправки, учитывающие действие молекулярных сил и влияние собственного объема частиц газа.

И д е а л ь н ы м г а з о м называется газ, который удовлетворяет следующим условиям:

1) собственный объем частиц (молекул или атомов) газа очень мал по сравнению с объемом сосуда, в котором находится газ;

2) силы взаимодействия между частицами идеального газа отсутствуют (или настолько слабы, что ими можно пренебрегать);

3) столкновения частиц газа между собой и со стенками сосуда являются абсолютно упругими ударами.

Иногда в связи с первым предположением считают, что частицы идеального газа есть «материальные точки». Тогда для возможности обмена энергиями между точечными телами нельзя игнорировать действие молекулярных сил (вероятность столкновения между невзаимодействующими материальными точками равна нулю). Однако если допустить, что частицы идеального газа есть шарики очень малых, но конечных размеров, то вероятность столкновения между ними будет отлична от нуля и при отсутствии сил взаимодействия; в этом случае столкновения можно рассматривать как простые механические удары упругих тел.

Частицы газа могут иметь различные скорости; иногда для упрощения рассуждений предполагают, что частицы газа имеют одинаковые скорости. Направления скоростей частиц ориентированы в пространстве совершенно беспорядочно; это означает, что число частиц, движущихся в одном направлении, равно числу частиц, движущихся в противоположном, а также и в любом другом направлении. Таким образом, к указанным выше трем предположениям необходимо добавить четвертое:

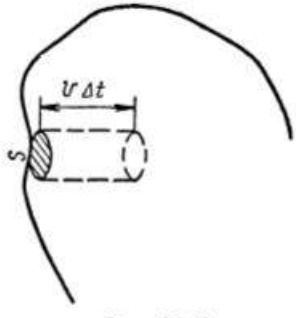


Рис. II.11

4) частицы идеального газа совершают беспорядочное тепловое движение.

Выведем теперь «основное уравнение кинетической теории газов», связывающее давление газа со скоростями движения его частиц. При беспорядочном движении частицы газа сталкиваются между собой и со стенками сосуда. Механическое действие этих ударов о стенки сосуда воспринимается как давление на эти стенки. Рассчитаем это давление, пользуясь сделанными выше предположениями. Выделим на стенке сосуда (рис. II.11) некоторую элементарную площадку S и найдем давление, оказываемое на эту площадку. Обозначим через v некоторую среднюю скорость частиц, N — полное число частиц газа, V — объем газа (сосуда), n — число частиц в единице объема ($n = N/V$).

Построим на площадке S цилиндр с высотой $v\Delta t$ и объемом $Sv\Delta t$, где Δt — некоторый малый промежуток времени (Δt должно быть значительно больше среднего времени свободного пробега частицы от одного столкновения до другого). Число частиц в этом объеме равно $\Delta N = nSv\Delta t$; из них за время Δt некоторая часть дойдет до S и ударится об нее. При каждом ударе одной частицы на площадке S произойдет некоторое изменение импульса; нам необходимо рассчитать суммарное изменение импульса всех частиц, которые за единицу времени ударились и отскочили от площадки S, и приравнять эту величину силе, действующей на площадку S. Расчет несколько усложняется

тем, что частицы движутся к площадке S под различными углами и имеют различные скорости. Однако можно получить правильный результат, если для упрощения рассуждений и расчетов заменить беспорядочное движение частиц следующим: допустим, что движение ΔN частиц происходит по трем взаимно перпендикулярным направлениям, так что одна шестая часть движется перпендикулярно площадке, одна шестая — в противоположном направлении, а остальные частицы движутся по четырем остальным направлениям параллельно площадке S. Таким образом, за время Δt число ударов частиц, двигающихся перпендикулярно стенке, о площадку S будет равно $\frac{1}{6} nSv\,\Delta t$. При каждом (абсолютно упругом) ударе частицы о стенку импульс этой частицы изменяется на $2m_0v$, где m_0 — масса частицы. Действительно, на первой стадии удара, когда летящая к стенке частица останавливается, она полностью теряет свой импульс m_0v, направленный к стенке. Во второй стадии удара, когда частица отбрасывается обратно, она получает импульс m_0v, направленный в обратную сторону. Таким образом, полное (суммарное) изменение импульса одной частицы будет равно $m_0v - (-m_0v) = 2m_0v$. За время Δt на площадке S изменяют направление своего движения $nSv\Delta t/6$ частиц; общее изменение импульса частиц газа на этой площадке равно

$$\Delta I = 2m_0v \cdot \frac{1}{6} nSv\,\Delta t = \frac{1}{3} nm_0v^2S\,\Delta t.$$

Согласно второму закону Ньютона, это изменение должно быть вызвано силой, приложенной со стороны стенки к частицам газа. Так как сила, действующая на систему, равна изменению импульса этой системы в единицу времени, то площадка S должна действовать на газ с силой

$$F = \frac{\Delta I}{\Delta t} = \frac{1}{3} nm_0v^2S.$$

Тогда давление на газ, т. е. внешняя сила, действующая на газ со стороны единичной площадки, равна $p = F/S$, т. е.

$$p = \frac{1}{3} nm_0v^2. \qquad (2.1)$$

На основании третьего закона Ньютона, действующего в процессе столкновений между частицами и стенкой, можно утверждать, что эта формула выражает также давление газа на поверхность стенки.

Формула (2.1) называется основным уравнением кинетической теории газов. Можно заменить n на N/V, и тогда

$$pV = \frac{1}{3} Nm_0v^2. \qquad (2.2)$$

Мы предполагали, что скорости молекул одинаковы; если же в газе имеется N_1 молекул со скоростью v_1, N_2 — со скоростью v_2 и т. д., то

$$pV = \sum\left(\frac{1}{3} m_0N_1v_1^2 + \frac{1}{3} m_0N_2v_2^2 + \ldots\right) = \frac{1}{3} m_0 \sum N_iv_i^2 = \frac{1}{3} m_0Nc^2, \qquad (2.3)$$

где

$$\sum N_i v_i^2 = Nc^2; \quad c = \sqrt{\frac{N_1 v_1^2 + N_2 v_2^2 + \dots}{N}}; \tag{2.4}$$

c называется *средней квадратичной скоростью* молекул газа; она отличается от *средней арифметической скорости* $\langle v \rangle$:

$$\langle v \rangle = \frac{N_1 v_1 + N_2 v_2 + \dots}{N}. \tag{2.5}$$

§ 9. УРАВНЕНИЕ СОСТОЯНИЯ ИДЕАЛЬНОГО ГАЗА

Допустим, что в объеме V при давлении p и температуре T находятся N молекул газа. Количество газа можно измерять либо его массой $m = Nm_0$ (m_0 — масса одной молекулы), либо числом молей:

$$\nu = \frac{m}{M}, \tag{2.6}$$

где M — молярная масса (масса одного моля вещества). Масса m выражается в килограммах (кг), M — в килограммах на моль (кг/моль), а ν — в молях. Число молекул в газе можно выразить через число молекул в одном моле вещества N_A (*постоянная Авогадро*, равная $6{,}022 \cdot 10^{23}$ 1/моль):

$$N = \nu N_A = N_A \frac{m}{M}. \tag{2.7}$$

Объем газа измеряется в кубических метрах ($м^3$), давление — в ньютонах на квадратный метр ($Н/м^2$), температура — по *абсолютной шкале Кельвина* (см. ч. II, § 18):

$$T = 273{,}15° + t,$$

где t — температура по шкале Цельсия.

Измерения показали, что у газов при равновесных переходах из одного состояния в другое изменения параметров p, V и T с некоторым приближением удовлетворяют так называемому *объединенному закону Бойля — Мариотта и Гей-Люссака*: для данной массы газа отношение произведения объема газа на его давление к абсолютной температуре сохраняется постоянным при переходе газа из одного равновесного состояния в другое:

$$\frac{p_1 V_1}{T_1} = \frac{p_2 V_2}{T_2}; \quad \frac{pV}{T} = \text{const}. \tag{2.8}$$

Поэтому идеальный газ определяется как такой газ, который в точности подчиняется законам Бойля — Мариотта и Гей-Люссака, а следовательно, объединенному закону (2.8). Так как при данном давлении и температуре объем газа V пропорционален его массе, то отношение pV/T должно быть также пропорционально массе газа, а следовательно числу частиц газа N. Обозначим коэффициент пропорциональности через k; тогда

$$\frac{pV}{T} = kN, \quad \text{или} \quad pV = NkT. \tag{2.9}$$

Величину k можно найти, если по массе газа найти число молекул N и измерить объем, давление и температуру газа в каком-нибудь его равновесном состоянии. Расчет $k = pV/(TN)$ дает для разнородных газов при различных (равновесных) значениях параметров одно и то же значение этой величины, равное постоянной Больцмана.

Заменим в формуле (2.9) число молекул N на $N_A \frac{m}{M}$. Произведение двух постоянных величин kN_A есть *молярная* (*универсальная*) *газовая постоянная*, равная $R = 8{,}31$ Дж/(К·моль).

Тогда формулу (2.9) можно переписать в виде

$$pV = \frac{m}{M} RT. \tag{2.10}$$

Эта формула, связывающая между собой параметры равновесного состояния идеального газа, называется *уравнением состояния идеального газа* (Клапейрона — Менделеева).

В формуле (2.9) можно разделить обе части ее на объем; получим:

$$p = \frac{N}{V} kT; \quad p = nkT. \tag{2.11}$$

Таким образом, давление газа прямо пропорционально числу молекул в единице объема и температуре газа.

Сравним уравнение (2.9) с основным уравнением кинетической теории газов (2.2), которое было выведено в предыдущем параграфе; из равенства правых частей этих формул следует:

$$NkT = \frac{1}{3} Nm_0 v^2; \quad \frac{m_0 v^2}{2} = \frac{3}{2} kT. \tag{2.12}$$

В основном уравнении теории газов (2.1) v есть средняя квадратичная скорость c; поэтому энергия, рассчитанная по этой скорости,

$$\frac{m_0 c^2}{2} = \frac{m_0}{2} \left(\frac{v_1^2 + v_2^2 + \ldots}{N} \right) = \frac{1}{N} \left(\frac{m_0 v_1^2}{2} + \frac{m_0 v_2^2}{2} + \ldots \right)$$

есть средняя кинетическая энергия поступательного движения молекул газа; согласно формуле (2.12), она оказывается прямо пропорциональной абсолютной температуре газа.

Вспомним, что поступательному движению молекул соответствуют три степени свободы. Тогда из формулы (2.12) следует, что на каждую степень свободы поступательного движения молекулы в среднем приходится энергия, равная

$$\varepsilon = \frac{1}{2} kT,$$

о чем было сказано в § 1.

Состояние идеального газа характеризуется не только занимаемым им объемом V, давлением p и температурой T, но и распределением молекул по скоростям. В неравновесных состояниях газа это распреде-

ление может быть любым, но в равновесном состоянии устанавливаются вполне определенные соотношения между числами молекул, имеющими различные скорости. При теоретическом выводе равновесного распределения молекул по скоростям используются следующие положения:

1) в газе не существует молекул, имеющих в точности одинаковые скорости;

2) число молекул ΔN, скорости которых лежат в узких пределах между v и $v + \Delta v$ (например, от 100 до 101 м/с или от 500 до 501 м/с), прямо пропорционально общему числу молекул N, ширине интервала Δv и зависит от величины скорости v. Зависимость ΔN от скорости представим в виде множителя $f(v)$:

$$\Delta N = N f(v)\, \Delta v. \qquad (2.13)$$

Функцию $f(v) = \Delta N/N\Delta v$, показывающую относительное число молекул, приходящихся на единицу интервала скоростей, называют *функцией распределения молекул газа по скоростям*. Таким образом, число частиц со скоростями, лежащими между 100 и 101 м/с, не равно числу частиц, скорости которых лежат между 200 и 201 м/с или 500 и 501 м/с. Функция $f(v)$ для равновесных состояний газа вычислена теоретически и проверена путем измерений; по Максвеллу, она имеет вид

$$f(v) = Av^2 e^{-Bv^2}, \qquad (2.14)$$

где A и B — некоторые постоянные величины, зависящие от массы молекул m_0 и от температуры газа T:

$$A = 4\pi\left(\frac{m_0}{2\pi kT}\right)^{3/2}; \quad B = \frac{m_0}{2kT} \qquad (2.15)$$

(k — постоянная Больцмана). Графически функция распределения молекул по скоростям имеет вид, представленный на рис. II.12. Заштрихованная площадка $f(v)\,dv$ равна dN/N, т. е. относительному числу частиц, скорости которых лежат в пределах v, $v + dv$. Скорость, соответствующая максимуму функции распределения и обозначенная на рисунке $v_в$, называется *средней наивероятной* (или наиболее вероятной) *скоростью* молекул газа. Большинство молекул имеют скорости, лежащие вблизи $v_в$.

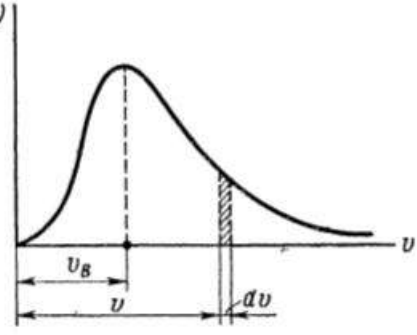


Рис. II.12

Пользуясь формулой Максвелла (2.14), можно рассчитать средние скорости молекул:

$$c = \int_0^\infty v^2 f(v)\, dv; \quad \langle v\rangle = \int_0^\infty v f(v)\, dv.$$

Средняя наивероятная скорость $v_в$ определяется как то значение скорости, для которого функция распределения (2.14) имеет максимум.

Расчеты дают формулы:

$$c = \sqrt{\frac{3kT}{m_0}} = \sqrt{\frac{3RT}{M}} \approx 1{,}73\sqrt{\frac{RT}{M}},$$

$$\langle v \rangle = \sqrt{\frac{8kT}{\pi m_0}} = \sqrt{\frac{8RT}{\pi M}} \approx 1{,}60\sqrt{\frac{RT}{M}}, \qquad (2.16)$$

$$v_в = \sqrt{\frac{2kT}{m_0}} = \sqrt{\frac{2RT}{M}} \approx 1{,}41\sqrt{\frac{RT}{M}}.$$

Измерение скоростей молекул было произведено Штерном при помощи установки, схема которой изображена на рис. II.13. Внутри цилиндра малого диаметра *I* со щелью *A* помещается платиновая проволока *P*, покрытая слоем серебра. При нагревании электрическим током серебро испаряется, атомы серебра вылетают через щель и попадают на внутреннюю поверхность второго цилиндра *II*. Если оба цилиндра неподвижны, то все атомы независимо от их скоростей попадут в одно и то же место *B* большого цилиндра. При одновременном вращении обоих цилиндров с угловой скоростью ω атомы серебра в зависимости от своих скоростей попадут в точки B', B'' и т. д. По величине ω, расстоянию l и смещению $x = BB'$ можно вычислить скорость атомов, попавших в точку B':

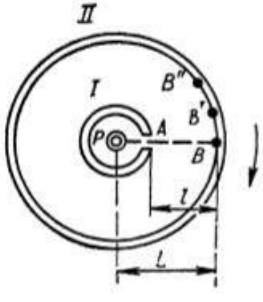


Рис. II.13

$$x = \omega \frac{l}{v} L; \quad v = \frac{\omega l}{x} L,$$

где L — расстояние от оси вращения до поверхности цилиндра *II*. Эти измерения подтвердили теоретические расчеты скоростей.

Однако следует заметить, что в подобных опытах определяется функция распределения молекул по скоростям не в самом сосуде, а в молекулярном пучке, вылетающем из сосуда. Очевидно, что наиболее вероятной скоростью молекул в пучке будет средняя скорость потока. Скорость истечения газа из отверстия в сосуде, в котором имеется давление p_1, в окружающую среду с давлением p_2, будет, согласно формуле (6.6), приведенной в § 30 ч. I, равна

$$v = \sqrt{\frac{2(p_1 - p_2)}{\rho}}.$$

В цилиндре *I* пары серебра можно считать идеальным газом и поэтому $p_1 = nkT$; в цилиндре *II* поддерживается вакуум и можно полагать $p_2 = 0$. Тогда ($\rho = nm_0$)

$$v = \sqrt{\frac{2kT}{m_0}}, \qquad (2.17)$$

т. е. средняя скорость молекул в потоке газа, выходящего из сосуда в вакуум, в точности равна средней наивероятной скорости молекул, вычисленной по формуле (2.14). Следовательно, измерения показывают, что распределение молекул по скоростям в потоке газа из сосуда

в вакуум соответствует формуле (2.14). Однако это вовсе не означает, что такое же распределение имеется в самом сосуде. Если, например, в цилиндре II давление было бы отлично от нуля, то средняя скорость потока отличалась бы от v; поэтому для проверки формулы (2.14) необходим более подробный анализ результатов измерений. В частности, производя измерения при различных значениях p_2, приближающихся к p_1, можно получить распределение молекул по скоростям в сосуде как предельное при $p_2 = p_1$.

§ 10. ИЗОПРОЦЕССЫ В ИДЕАЛЬНОМ ГАЗЕ; ТЕПЛОЕМКОСТИ ГАЗОВ

Согласно первому закону термодинамики, если газ при переходе из одного состояния в другое получает извне (или отдает в окружающую среду) энергию Q в виде теплоты и при этом совершается положительная (или отрицательная) внешняя работа A, то

$$Q = (U_2 - U_1) + A, \tag{2.18}$$

а для бесконечно близких состояний

$$\mathrm{d}Q = \mathrm{d}U + \mathrm{d}A. \tag{2.19}$$

Изменение внутренней энергии идеального газа можно рассчитать на основании формулы (1.9). Обычно полагают, что молекулы идеального газа обладают только энергией поступательного и вращательного движений (энергией колебательного движения обладают молекулы реальных газов, жидкостей и твердых тел). Доказано, что в равновесном состоянии газа поступательные и вращательные степени свободы являются равноценными, т. е. на каждую из этих степеней свободы в среднем приходится одна и та же энергия $kT/2$; тогда молекула идеального газа, обладающая i степенями свободы, имеет среднюю энергию $ikT/2$, а энергия N молекул будет равна

$$U = N\,\frac{i}{2}\,kT = \frac{m}{M}\,\frac{iR}{2}\,T. \tag{2.20}$$

Следовательно,

$$U_2 - U_1 = \frac{m}{M}\,\frac{iR}{2}\,(T_2 - T_1); \quad \mathrm{d}U = \frac{m}{M}\,\frac{iR}{2}\,\mathrm{d}T. \tag{2.21}$$

Для одноатомного газа $i = 3$, для двухатомного $i = 5$, для молекул, состоящих из трех атомов и более, $i = 6$. Существенно, что *внутренняя энергия идеального газа зависит только от температуры и при данной температуре не зависит от объема и давления газа.* Эта энергия прямо пропорциональна массе газа (или числу молекул в газе), числу степеней свободы каждой молекулы и абсолютной температуре газа. При переходе из одного состояния в другое внутренняя энергия идеального газа изменяется только в том случае, если при этом переходе изменяется температура газа.

Внешняя работа, совершаемая газом, рассчитывается на основании общих формул (1.16) и (1.17):

$$\mathrm{d}A = p\,\mathrm{d}V; \quad A = \int_{V_1}^{V_2} p\,\mathrm{d}V. \tag{2.22}$$

Напомним, что эта работа в зависимости от знака изменения объема $V_2 - V_1$ может быть положительной (при расширении газа) или отрицательной (при сжатии газа, когда положительную работу совершают внешние силы, сжимающие газ). Для подстановки в формулу (2.18) работа должна быть измерена в тех же единицах (например, в джоулях), в которых измеряются количество теплоты и внутренняя энергия.

Теплообмен между газом и окружающей средой можно рассчитывать по той же формуле, по которой определяется количество теплоты, необходимое для нагревания (или отнимаемое при охлаждении) твердых и жидких веществ:

$$Q = mc\,(T_2 - T_1); \quad \mathrm{d}Q = mc\,\mathrm{d}T. \tag{2.23}$$

Однако измерения показали, что удельная теплоемкость для данного вещества (c) не есть постоянная величина, а заметно зависит от того, как протекает самый процесс нагревания или охлаждения. Например, при постоянном объеме теплоемкость одноатомного газа почти на 70% меньше, чем при постоянном давлении; для двухатомного газа эти теплоемкости отличаются на 40%. Более того, можно сообщить газу теплоту и одновременно изменять его объем так, чтобы газ всю получаемую извне теплоту целиком расходовал на внешнюю работу и температура газа оставалась постоянной; тогда формула (2.23) вообще теряет смысл ($Q \neq 0$; $T_2 - T_1 = 0$). Таким образом, если формула (2.23) применяется к газам, то удельная теплоемкость зависит не только от вещества газа, но и от условий нагревания — насколько изменяется объем газа при сообщении или отнятии от него теплоты. Впрочем, зависимость удельной теплоемкости от изменения объема существует и для жидких и твердых тел, однако эта зависимость (ввиду малого изменения объема таких тел при нагревании) слабая и в приближенных технических расчетах не принимается во внимание.

Подставим в уравнение (2.18) первого закона термодинамики указанные выше выражения (2.20) и (2.23) для внутренней энергии и для теплообмена; тогда

$$mc\,(T_2 - T_1) = \frac{m}{M}\frac{iR}{2}(T_2 - T_1) + A; \tag{2.24}$$

$$c = \frac{1}{M}\frac{iR}{2} + \frac{A}{m\,(T_2 - T_1)}.$$

Так как $\frac{1}{M}\frac{iR}{2}$ — постоянная величина, то удельная теплоемкость идеального газа зависит от величины совершаемой газом внешней работы A. Рассмотрим два простейших, но важных способа нагревания газа. Допустим, газу сообщается теплота Q при постоянном объеме; тогда газ работы не совершает ($A = 0$) и вся сообщаемая теплота целиком идет на изменение внутренней энергии газа, т. е. на его нагревание. Из формулы (2.24) получаем

$$c_V = \frac{1}{M}\frac{iR}{2}, \tag{2.25}$$

где индекс V показывает, что сообщение или отнятие теплоты происходит при постоянном объеме. Произведение удельной теплоемкости газа

на массу одного моля называется *молярной теплоемкостью*: $C_m = cM$. Для идеального газа при постоянном объеме молярная теплоемкость равна

$$C_V = \frac{iR}{2} \tag{2.26}$$

и одинакова для всех газов, имеющих одинаковое число степеней свободы молекул; удельная же теплоемкость обратно пропорциональна массе одного моля газа.

Допустим теперь, что газу сообщается теплота, но при этом его давление поддерживается постоянным. Тогда газ должен расширяться и будет совершать положительную внешнюю работу

$$A = \int_{V_1}^{V_2} p\,dV = p(V_2 - V_1) = pV_2 - pV_1 = \frac{m}{M}RT_2 - \frac{m}{M}RT_1 = \frac{m}{M}R(T_2 - T_1)$$

(здесь мы воспользовались уравнением состояния идеального газа $pV_1 = \frac{m}{M}RT_1$ и $pV_2 = \frac{m}{M}RT_2$, чтобы выразить совершаемую газом внешнюю работу через изменение температуры газа). Подставим это выражение для работы в формулу (2.24) и снабдим удельную теплоемкость при постоянном давлении индексом p:

$$c_p = \frac{1}{M}\frac{iR}{2} + \frac{R}{M} = \frac{R}{M}\frac{i+2}{2}. \tag{2.27}$$

Тогда мольная теплоемкость при постоянном давлении

$$C_p = Mc_p = \frac{i+2}{2}R. \tag{2.28}$$

Таким образом, теплоемкости идеального газа при постоянном объеме и давлении зависят только от числа степеней свободы молекул; *для всех газов, молекулы которых имеют одинаковое число степеней свободы, молярные теплоемкости одинаковы*, но их удельные теплоемкости различны, так как они обратно пропорциональны молекулярной массе M.

Сравнивая выражение для теплоемкостей идеального газа при постоянном объеме и давлении, получим:

$$C_p = C_V + R; \quad c_p = c_V + \frac{R}{M};$$

$$\gamma = \frac{C_p}{C_V} = \frac{c_p}{c_V} = \frac{i+2}{i}. \tag{2.29}$$

Мы нашли два значения теплоемкости идеального газа, соответствующие двум различным условиям нагревания: при постоянном объеме и при постоянном давлении. Однако можно сообщить или отнимать теплоту и при других условиях нагревания; таких условий может быть бесконечное множество; среди них следует выделить еще два «крайних» случая:

1) допустим, что газ получает теплоту Q и одновременно совершает внешнюю работу, численно равную $A = Q$; тогда изменение внутренней энергии газа $U_2 - U_1 = Q - A$ равно нулю, следовательно, температура газа остается постоянной. Применяя к этому процессу формулу (2.23), получим для теплоемкости газа $c = \frac{1}{m}\frac{Q}{T_2 - T_1}$ бесконечно большое значение:

$$Q \neq 0, \quad T_2 = T_1, \quad c = \infty.$$

Таким образом, *при изотермическом* (T = const) *сообщении или отнятии теплоты газ ведет себя как вещество, обладающее бесконечно большой теплоемкостью*;

2) допустим, что теплообмен между газом и окружающей средой чрезвычайно мал (близок к нулю), но путем сильного сжатия или расширения вызывается очень большое изменение температуры. Применяя к такому процессу формулу (2.23), получим для теплоемкости газа очень малые значения. В пределе, когда $Q \to 0$, $T_2 \neq T_1$, $c \to 0$, т. е. *при адиабатическом* ($Q = 0$) *изменении температуры, газ ведет себя как вещество, обладающее нулевой теплоемкостью.*

Таким образом, применяя формулу (2.23) или (2.24) для вычисления теплоемкости газа при всех условиях, когда газу сообщается или от него отнимается теплота, мы получаем для этой теплоемкости все возможные значения от нуля до бесконечности. Среди этих значений теплоемкости особо важную роль в термодинамике играют C_p и C_V. В частности, пользуясь формулой (2.26), можно выразить внутреннюю энергию идеального газа через C_V:

$$U = \frac{m}{M}\frac{iR}{2}T = \frac{m}{M}C_V T. \tag{2.30}$$

Рассмотрим теперь простейшие газовые процессы. Для того чтобы можно было использовать уравнение состояния идеального газа $pV = \frac{m}{M}RT$ (применимое только для равновесных состояний), предположим, что все рассматриваемые процессы протекают достаточно медленно, поэтому́ в каждом из промежуточных состояний давление, плотность и температура одинаковы в пределах всего объема.

Отложим по оси абсцисс объем, а по оси ординат — давление газа. В этой координатной системе каждая точка *1* или *2* соответствует определенному равновесному состоянию газа, так как по данным двум параметрам состояния p и V, известной массе газа и массе одного моля можно по уравнению состояния рассчитать третий неизвестный параметр — температуру T. Вследствие этого любая линия, например *1—2* (рис. II.14), изображает некоторый равновесный процесс изменения состояния идеального газа, а площадь *1—2—3—4—1* — внешнюю работу, совершаемую при этом процессе. Для характеристики каждого из таких процессов необходимо знать уравнение процесса, связывающее параметры состояния газа (или уравнение кривой, изображающей этот процесс в выбранной координатной системе), и, кроме того, вычислить следующие величины: 1) изменение внутренней энергии газа; 2) теплообмен с окружающей средой и 3) совершаемую газом

внешнюю работу. Найдем эти величины для простейших процессов, полагая, что они протекают равновесно.

I. Изохорический процесс, при котором объем газа в течение всего процесса сохраняется постоянным (рис. II.15):

а) *уравнение процесса* $V = \text{const}$ можно получить, использовав уравнение состояния идеального газа: для начального состояния $p_1V = \frac{m}{M}RT_1$; для конечного состояния $p_2V = \frac{m}{M}RT_2$. Разделив их почленно, получим:

$$\frac{p_1}{p_2} = \frac{T_1}{T_2} \quad \text{или} \quad \frac{p}{T} = \text{const};$$
$$p = \text{const} \cdot T, \tag{2.31}$$

т. е. при изохорическом процессе давление газа прямо пропорционально абсолютной температуре;

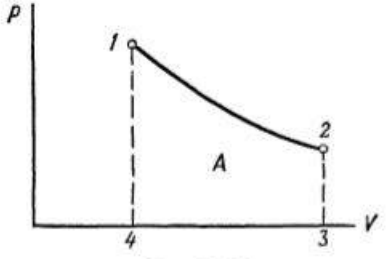


Рис. II.14

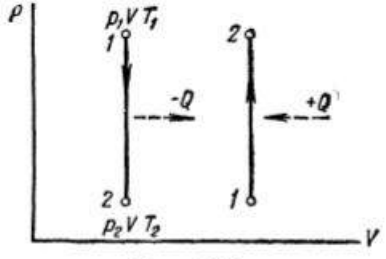


Рис. II.15

б) *изменение внутренней энергии* газа можно рассчитать по формуле

$$U_2 - U_1 = \frac{m}{M}\frac{iR}{2}(T_2 - T_1) = \frac{m}{M}C_V(T_2 - T_1). \tag{2.32}$$

Заметим, что эта формула применима для всех без исключения процессов, так как внутренняя энергия идеального газа зависит только от температуры и изменяется прямо пропорционально температуре;

в) *внешняя работа* равна нулю, так как $V = \text{const}$;

г) *теплообмен* с окружающей средой

$$Q = mc_V(T_2 - T_1) = \frac{m}{M}C_V(T_2 - T_1). \tag{2.33}$$

Согласно первому закону термодинамики,

$$Q = U_2 - U_1 + A; \quad A = 0; \quad Q = U_2 - U_1,$$

поэтому, если газ получает теплоту, то его температура, а по формуле (2.31) также и давление повышаются; если газ отдает теплоту, то температура и давление понижаются. На рис. II.15 горизонтальными стрелками условно показан теплообмен и его знак.

II. Изобарический процесс, при котором давление газа в течение всего процесса поддерживается постоянным (рис. II.16):

а) *уравнение процесса* $p = \text{const}$ можно получить, опять-таки использовав уравнение состояния:

$$\left.\begin{aligned} pV_1 &= \frac{m}{M} RT_1; \\ pV_2 &= \frac{m}{M} RT_2; \end{aligned}\right| \frac{V_1}{V_2} = \frac{T_1}{T_2};$$

$$\frac{V}{T} = \text{const}; \quad V = \text{const} \cdot T, \tag{2.34}$$

т. е. при изобарическом процессе объем газа прямо пропорционален его абсолютной температуре; большим объемам соответствуют высокие температуры;

б) *изменение внутренней энергии* газа рассчитывается по той же общей формуле (2.32). Следует иметь в виду, что в этой формуле молярная теплоемкость при постоянном объеме является просто постоянной величиной, заменяющей $iR/2$;

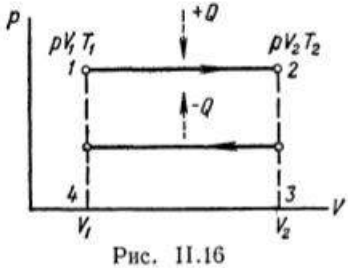


Рис. II.16

в) *внешняя работа* может быть рассчитана аналитически по формуле (2.22):

$$A = \int_{V_1}^{V_2} p \, dV = p(V_2 - V_1) \tag{2.35}$$

или же графически, как площадь *1—2—3—4—1*. Воспользовавшись уравнениями состояния, приведенными выше, можно заменить pV на $\frac{m}{M} RT$ и выразить внешнюю работу в зависимости от изменения температуры газа:

$$A = \frac{m}{M} R(T_2 - T_1); \tag{2.36}$$

г) *теплообмен* с окружающей средой рассчитывается по формуле

$$Q = mc_p(T_2 - T_1) = \frac{m}{M} C_p(T_2 - T_1), \tag{2.37}$$

где c_p — удельная, а C_p — молярная теплоемкости при постоянном давлении. Применение первого закона термодинамики к изобарическому процессу дает уже полученное выше соотношение между теплоемкостями при постоянном объеме и постоянном давлении:

$$Q = U_2 - U_1 + A;$$

$$\frac{m}{M} C_p(T_2 - T_1) = \frac{m}{M} C_V(T_2 - T_1) + \frac{m}{M} R(T_2 - T_1); \tag{2.38}$$

$$C_p = C_V + R.$$

Это соотношение называется *уравнением Р. Майера.*

III. Изотермический процесс, при котором температура газа в течение всего процесса поддерживается постоянной:

а) *уравнение процесса* ($T = \text{const}$) можно на основании уравнения состояния представить в следующем виде:

$$pV = \frac{m}{M} RT;$$

$$T = \text{const}; \quad pV = \text{const}; \quad p_1 V_1 = p_2 V_2; \quad \frac{p_1}{p_2} = \frac{V_2}{V_1}, \tag{2.39}$$

т. е. при изотермическом процессе давление газа изменяется обратно пропорционально его объему. Графически такой процесс изобразится гиперболой, показанной на рис. II.17;

б) *изменение внутренней энергии*, очевидно, равно нулю ($T_2 = T_1$; $U_2 = U_1$);

в) *внешняя работа* (площадь *1—2—3—4—1*) может быть рассчитана следующим образом:

$$A = \int_1^2 p \, dV = \int_1^2 \frac{m}{M} RT \frac{dV}{V} = \frac{m}{M} RT \ln \frac{V_2}{V_1} \tag{2.40}$$

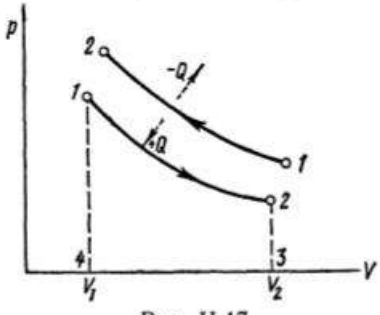


Рис. II.17

(при вычислении была произведена замена: $p = \frac{m}{M} \frac{RT}{V}$). В этой формуле при необходимости можно заменить $\frac{m}{M} RT$ на $p_1 V_1$ или $p_2 V_2$, а отношение V_2/V_1 — на p_1/p_2, так что можно, например, написать:

$$A = p_1 V_1 \ln \frac{p_1}{p_2}; \tag{2.41}$$

г) *теплообмен* между газом и окружающей средой не может быть найден по формулам, аналогичным (2.33) или (2.37), так как $T = \text{const}$. Этот теплообмен должен рассчитываться на основании первого закона термодинамики:

$$Q = U_2 - U_1 + A; \quad U_2 = U_1; \quad Q = A = \frac{m}{M} RT \ln \frac{V_2}{V_1}, \tag{2.42}$$

т. е. при изотермическом процессе, совершаемом идеальным газом, теплообмен между этим газом и окружающей средой по величине и знаку равен внешней работе. При расширении ($V_2 > V_1$) логарифм отношения V_2/V_1 есть положительная величина, следовательно, газ совершает положительную работу и при этом получает извне эквивалентное количество теплоты. При сжатии, наоборот, совершается отрицательная работа, и чтобы газ, сжимаемый внешними силами, не нагревался, от него отнимается теплота, эквивалентная этой работе. Следует предупредить, что для других веществ (реальных газов, жидких и твердых тел), у которых зависимость внутренней энергии от температуры более сложная, чем в выражении (2.20), последние два соотношения в формулах (2.42) неприменимы.

IV. Адиабатический процесс, при котором теплообмен между газом и окружающей средой отсутствует, т. е. $dQ = 0$:

а) *уравнение адиабатического процесса* получим, воспользовавшись первым законом термодинамики в дифференциальном виде:

$$dQ = dU + dA.$$

Имеем:

$$dQ = 0; \quad dU = \frac{m}{M} C_V dT; \quad dA = p\,dV = \frac{m}{M} RT \frac{dV}{V}.$$

Следовательно,

$$\frac{m}{M} C_V dT + \frac{m}{M} RT \frac{dV}{V} = 0; \quad \frac{dT}{T} + \frac{R}{C_V} \frac{dV}{V} = 0.$$

Из этого выражения получается следующее соотношение между объемом и температурой газа:

$$d\left(\ln T + \frac{R}{C_V} \ln V + \text{const}\right) = 0$$

или

$$\ln T + \frac{R}{C_V} \ln V = \text{const}; \quad TV^{R/C_V} = \text{const}.$$

В полученной формуле можно произвести замену, воспользовавшись соотношением $C_p = C_V + R$ и обозначением $C_p/C_V = \gamma$:

$$\frac{R}{C_V} = \frac{C_p - C_V}{C_V} = \gamma - 1.$$

Тогда окончательно уравнение адиабатического процесса примет вид

$$TV^{\gamma-1} = \text{const}; \quad T_1 V_1^{\gamma-1} = T_2 V_2^{\gamma-1}; \quad \frac{T_1}{T_2} = \left(\frac{V_2}{V_1}\right)^{\gamma-1}. \qquad (2.43)$$

Из первой формулы можно исключить температуру, воспользовавшись уравнением состояния: умножим левую часть на pV, а правую — на $\frac{m}{M} RT$ и внесем $\frac{m}{M} R$ в постоянную. После сокращений (температуры) получим:

$$pV^{\gamma} = \text{const}; \quad p_1 V_1^{\gamma} = p_2 V_2^{\gamma}; \quad \frac{p_1}{p_2} = \left(\frac{V_2}{V_1}\right)^{\gamma}. \qquad (2.44)$$

Формулы (2.43) и (2.44), связывающие параметры идеального газа при равновесном адиабатическом процессе, называются уравнениями Пуассона. Отношение $C_p/C_V = \gamma$ называется *показателем адиабатического процесса*; значение γ, согласно формуле (2.29), зависит от числа степеней свободы молекул. Для одноатомного газа $i = 3$, $\gamma = 5/3 \approx 1{,}67$; для двухатомного газа $i = 5$, $\gamma = 1{,}4$; для трехатомного газа $i = 6$, $\gamma = 8/6 = 1{,}33$.

Графически адиабатический процесс, совершаемый идеальным газом, изображается кривой, падающей круче, чем кривая изотермического процесса (рис. II.18). Об этом уже было сказано в § 6 ч. II;

напомним, что при изотермическом процессе подводимая теплота частично компенсирует падение давления, тогда как при адиабатическом процессе подвода теплоты нет, температура газа понижается и давление падает быстрее;

б) *изменение внутренней энергии* подсчитывается по общей формуле (2.32), если известны начальная и конечная температуры газа. Это изменение можно также определить из первого закона термодинамики, если известна внешняя работа, совершаемая газом: $Q = U_2 - U_1 + A = 0$; $U_2 - U_1 = -A$, т. е. при адиабатическом процессе газ совершает внешнюю работу за счет изменения своей внутренней энергии;

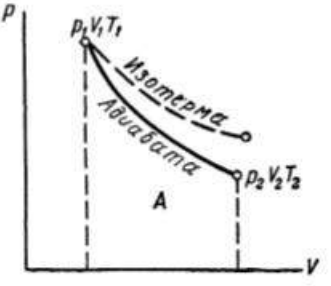


Рис. II.18

в) *внешняя работа* при адиабатическом процессе может быть вычислена непосредственно путем интегрирования общего выражения для внешней работы с использованием уравнения процесса (2.44) или же, проще, на основании первого закона термодинамики (2.18) и уравнения состояния:

$$A = U_1 - U_2 = \frac{m}{M}\frac{iR}{2}(T_1 - T_2) = \frac{i}{2}\left(\frac{m}{M}RT_1 - \frac{m}{M}RT_2\right) = \frac{i}{2}(p_1V_1 - p_2V_2).$$

Так как $i/2 = 1/(\gamma - 1)$, то внешняя работа при равновесном адиабатическом процессе равна

$$A = \frac{1}{\gamma - 1}(p_1V_1 - p_2V_2). \qquad (2.45)$$

Адиабатическими являются процессы сжатия и расширения газа при распространении в нем звуковой волны. Формула (5.12), выведенная в § 22 ч. I,

$$v_{\text{звука}} = \sqrt{\gamma \frac{p}{\rho}}$$

показывает, что скорость звука зависит от отношения давления p к плотности ρ газа, а также от постоянной величины γ, значение которой определяется характером процессов сжатия и расширения. Если бы эти процессы были изотермическими, то γ должно было бы равняться единице, если же адиабатическими, то $\gamma = c_p/c_V$. Измерения показали, что γ равно показателю адиабатического процесса того газа, в котором распространяется звуковая волна.

Если воспользоваться уравнением состояния идеального газа (2.10) и заменить p/ρ на kT/m_0 или RT/M, то формула для скорости распространения звука в газе примет вид

$$v_{\text{звука}} = \sqrt{\gamma \frac{kT}{m_0}} = \sqrt{\gamma \frac{RT}{M}}. \qquad (2.46)$$

Для данного газа скорость звука зависит только от температуры. Поэтому отношение скорости звука к средней квадратичной скорости

теплового движения молекул (см. формулу (2.16)) зависит только от показателя адиабатического процесса:

$$\frac{v_{\text{звука}}}{c} = \sqrt{\frac{\gamma}{3}}.$$

Адиабатический характер процессов сжатия и расширения газа при распространении звука объясняется тем, что эти процессы происходят весьма быстро и поэтому теплообмен между элементами газа (нагревшимися при сжатии и охладившимися при расширении) практически отсутствует. Однако следует отметить и другую особенность адиабатических процессов в газах. Рассмотрим формулу (*уравнение политропы*)

$$pV^n = \text{const},$$

где n — число, которое может принимать всевозможные значения от нуля до бесконечности. При $n = 0$ эта формула описывает изобарический, при $n = 1$ — изотермический, при $n = \gamma$ — адиабатический процесс. Для любого процесса, совершаемого идеальным газом, можно подобрать соответствующее значение показателя n. «Простыми» процессами будут те, для которых n является постоянным числом; у более сложных процессов n будет функцией объема (или давления).

Пользуясь этой формулой, найдем связь между изменениями давления и удельного объема в звуковой волне:

$$\frac{\Delta p}{p} = -n\frac{\Delta V}{V}; \quad \Delta V = -\frac{1}{n}V\frac{\Delta p}{p}.$$

При адиабатическом процессе ($n = \gamma$) изменение удельного объема при данном Δp оказывается меньше, чем при изотермическом ($n = 1$). Если разделить ΔV на половину периода колебаний (когда происходит только сжатие или только расширение газа), то можно найти среднюю скорость изменения удельного объема газа. Эта скорость у адиабатического процесса меньше, чем у изотермического; в этом смысле адиабатический процесс будет более медленным, чем изотермический, и поэтому более близким к равновесному.

§ 11. КРУГОВЫЕ ПРОЦЕССЫ, СОВЕРШАЕМЫЕ ИДЕАЛЬНЫМ ГАЗОМ; ЦИКЛ КАРНО. ЭНТРОПИЯ ИДЕАЛЬНОГО ГАЗА

Допустим, что идеальный газ совершает круговой процесс, т. е. проходит через некоторую последовательность промежуточных состояний и в точности возвращается в первоначальное состояние. Ограничимся рассмотрением только равновесных круговых процессов, состоящих из равновесных промежуточных состояний, так как только к таким процессам можно применить уравнение состояния (2.10).

Равновесные круговые процессы графически (например, в координатах p и V) изображаются замкнутыми линиями.

Круговых процессов, совершаемых идеальным газом, может быть бесконечное множество. В частности, круговые процессы могут состоять из простейших процессов, рассмотренных в предыдущем параграфе. На рис. II.19 приведены несколько таких круговых процессов.

Прежде всего проверим, соблюдается ли для этих процессов соотношение (1.32).

$$\sum \frac{\Delta Q}{T} = 0 \quad \text{или} \quad \oint \frac{dQ}{T} = 0, \tag{2.47}$$

лежащее в основе второго закона термодинамики. Будем рассматривать круговые процессы, идущие в прямом направлении, при котором газ совершает положительную внешнюю работу за счет полученной извне теплоты. Для вычисления интеграла по замкнутому контуру (2.47) расчленим его на сумму интегралов по отдельным участкам:

$$\oint \frac{dQ}{T} = \int_1^2 \frac{dQ}{T} + \int_2^3 \frac{dQ}{T} + \int_3^4 \frac{dQ}{T} + \int_4^1 \frac{dQ}{T}. \tag{2.48}$$

Каждое из этих слагаемых представляет собой изменение энтропии на соответствующем участке кругового процесса. Для сокращения записи

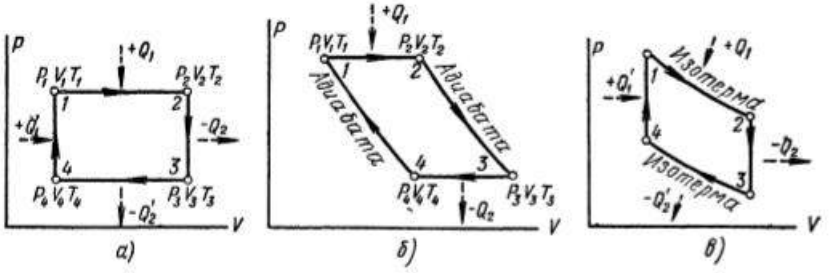


Рис. II.19

допустим, что круговой процесс совершается одним молем газа, т. е. $m/M = 1$. Тогда для процесса, изображенного на рис. II.19, *а*, имеем:

$$dQ_1 = C_p\,dT; \quad dQ_2 = C_V\,dT; \quad dQ_2' = C_p\,dT; \quad dQ_1' = C_V\,dT.$$

Для изобарических процессов *1—2* и *3—4* изменения энтропии равны:

$$\int_1^2 \frac{dQ_1}{T} = \int_1^2 C_p \frac{dT}{T} = C_p \ln \frac{T_2}{T_1};$$

$$\int_3^4 \frac{dQ_2'}{T} = \int_3^4 C_p \frac{dT}{T} = C_p \ln \frac{T_4}{T_3}.$$

Для изохорических процессов *2—3* и *4—1* изменения энтропии равны:

$$\int_2^3 \frac{dQ_2}{T} = \int_2^3 C_V \frac{dT}{T} = C_V \ln \frac{T_3}{T_2}; \quad \int_4^1 \frac{dQ_1'}{T} = \int_4^1 C_V \frac{dT}{T} = C_V \ln \frac{T_1}{T_4}.$$

Подставив эти значения в формулу (2.48), получаем

$$\oint \frac{dQ}{T} = C_p \ln \frac{T_2}{T_1} + C_V \ln \frac{T_3}{T_2} + C_p \ln \frac{T_4}{T_3} + C_V \ln \frac{T_1}{T_4}. \tag{2.49}$$

Но, согласно уравнениям процесса для изобар *1—2* и *4—3*, имеем:

$$\frac{V_2}{V_1} = \frac{T_2}{T_1};\ \frac{V_3}{V_4} = \frac{T_3}{T_4}.$$

Так как $V_1 = V_4$, $V_2 = V_3$, то

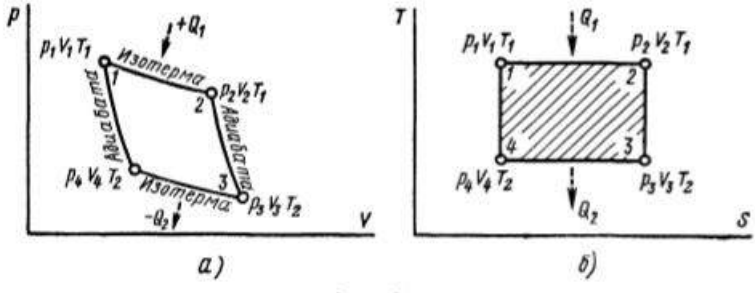
$$\frac{T_2}{T_1} = \frac{T_3}{T_4};\ \frac{T_3}{T_2} = \frac{T_4}{T_1}.$$

Воспользовавшись этими соотношениями, легко показать, что выражение (2.49) равно нулю, т. е. суммарное изменение энтропии идеального газа в результате такого кругового процесса равно нулю.

Такой же результат можно получить и для других процессов (см. рис. II.19, *б*, *в*).

ЦИКЛ КАРНО

Рассмотрим предложенный С. Карно круговой процесс (ц и к л К а р н о), состоящий из двух изотерм и двух адиабат (рис. II.20, *а*). Получение тепла Q_1 извне происходит изотермически (*1—2*), когда газ



а) *б)*

Рис. II.20

находится в тепловом контакте с источником теплоты, имеющим высокую температуру T_1. Затем в состоянии *2* газ отсоединяется от источника теплоты и совершает адиабатическое расширение (*2—3*), при котором теплообмен с окружающей средой отсутствует; при этом газ совершает внешнюю работу за счет своей внутренней энергии, вследствие чего температура газа понижается до T_2. После этого газ приводится в соприкосновение с «холодильником», имеющим температуру T_2, и производится изотермическое сжатие (*3—4*); при этом газ отдает холодильнику теплоту Q_2, эквивалентную работе, совершаемой сжимающими газ внешними силами. В состоянии *4* газ отсоединяется от холодильника и адиабатически сжимается (*4—1*) до первоначального состояния *1*; повышение температуры от T_2 до T_1 производится за счет работы внешних сжимающих сил. На рис. II.20, *б* этот процесс изображен в координатах температура — энтропия.

Изменения энтропии при изотермических процессах *1—2* и *3—4* равны:

$$\int_1^2 \frac{dQ_1}{T_1} = \frac{1}{T_1}\int_1^2 dQ_1 = \frac{Q_1}{T_1} = \frac{1}{T_1}\left(\frac{m}{M} RT_1 \ln \frac{V_2}{V_1}\right) = \frac{m}{M} R \ln \frac{V_2}{V_1};$$

$$\int_3^4 \frac{dQ_2}{T_2} = \frac{Q_2}{T_2} = \frac{1}{T_2}\left(\frac{m}{M} RT_2 \ln \frac{V_4}{V_3}\right) = \frac{m}{M} R \ln \frac{V_4}{V_3} = -\frac{m}{M} R \ln \frac{V_3}{V_4}.$$

Изменения энтропии при адиабатических процессах *2—3* и *4—1* равны нулю; для этих процессов можно написать соотношение (2.43):

$$\left.\begin{aligned} T_1 V_2^{\gamma-1} &= T_2 V_3^{\gamma-1}; \\ T_1 V_1^{\gamma-1} &= T_2 V_4^{\gamma-1}; \end{aligned}\right| \left(\frac{V_2}{V_1}\right)^{\gamma-1} = \left(\frac{V_3}{V_4}\right)^{\gamma-1};\ \frac{V_2}{V_1} = \frac{V_3}{V_4}.$$

Следовательно, суммарное изменение энтропии при цикле Карно также равно нулю:

$$\oint \frac{dQ}{T} = \frac{Q_1}{T_1} + \frac{Q_2}{T_2} = \frac{m}{M} R \ln \frac{V_2}{V_1} - \frac{m}{M} R \ln \frac{V_3}{V_4} = 0.$$

Для цикла Карно коэффициент полезного действия выражается только через температуру источника теплоты и температуру холодильника:

$$\eta = \frac{Q_1 - Q_2}{Q_1} = \frac{T_1 - T_2}{T_1} \qquad (2.50)$$

(в этой формуле Q_2 подставляется по абсолютной величине, так как его знак учитывается уже в самом выражении для η).

Для других круговых процессов коэффициент полезного действия имеет более сложный вид. Внешняя работа, совершаемая при любом круговом процессе, должна определяться как алгебраическая сумма положительных работ расширения и отрицательных работ сжатия. Графически эта работа изображается площадью, охватываемой линией процесса.

ПРОЦЕССЫ, ВЫЗЫВАЮЩИЕ ИЗМЕНЕНИЕ ЭНТРОПИИ

Выведем общее выражение для изменения энтропии идеального газа при его переходе из одного равновесного состояния в другое. Согласно определению,

$$S_2 - S_1 = \int_1^2 \frac{dQ}{T_1}. \qquad (2.51)$$

Подставим сюда $dQ = dU + dA$; $dU = \frac{m}{M} C_V dT$; $dA = pdV$ и на основании уравнения состояния заменим p на $\frac{m}{M}\frac{RT}{V}$. Получаем

$$S_2 - S_1 = \int_1^2 \left(\frac{m}{M} C_V \frac{dT}{T} + \frac{m}{M} R \frac{dV}{V}\right) = \frac{m}{M}\left(C_V \ln \frac{T_2}{T_1} + R \ln \frac{V_2}{V_1}\right). \qquad (2.52)$$

Таким образом, изменение энтропии идеального газа рассчитывается в зависимости от изменения его объема и температуры. Эту формулу можно несколько видоизменить:

$$S_2 - S_1 = \frac{m}{M} C_V \ln \frac{T_2}{T_1} \left(\frac{V_2}{V_1}\right)^{R/C_V} = \frac{m}{M} C_V \ln \frac{T_2 V_2^{\gamma-1}}{T_1 V_1^{\gamma-1}}, \qquad (2.53)$$

так как $\frac{R}{C_V} = \frac{C_p - C_V}{C_V} = \gamma - 1$. Заметим, что при равновесном адиабатическом процессе, согласно уравнению Пуассона, числитель и знаменатель выражения, находящегося под знаком логарифма, равны, и тогда изменение энтропии, как и следовало ожидать, равно нулю. В формуле (2.53) можно сделать замену:

$$\frac{m}{M} C_V = \frac{m}{M} \frac{iR}{2} = \frac{i}{2} Nk; \quad \gamma - 1 = \frac{2}{i},$$

тогда

$$S_2 - S_1 = \frac{i}{2} Nk \ln \frac{T_2}{T_1} \left(\frac{V_2}{V_1}\right)^{2/i} = Nk \ln \left(\frac{T_2}{T_1}\right)^{i/2} \frac{V_2}{V_1}, \qquad (2.54)$$

где i — число степеней свободы одной молекулы данного газа. Здесь важно отметить, что изменение энтропии газа прямо пропорционально числу молекул N, поэтому величина

$$\Delta s = \frac{S_2 - S_1}{N} = k \ln \left(\frac{T_2}{T_1}\right)^{i/2} \frac{V_2}{V_1}$$

есть изменение энтропии, приходящееся на *одну* молекулу газа.

По этим формулам можно рассчитывать изменение энтропии газа, *вызванное только равновесным теплообменом*; при этом энтропия возрастает, когда газ получает теплоту извне, и убывает, когда теплота отнимается от газа.

Однако существуют и другие процессы, которые вызывают изменение энтропии; это: 1) выравнивание температуры и общего давления в пределах объема газа и 2) перемешивание различных молекул, т. е. выравнивание парциальных давлений и плотностей различных компонент газа до достижения равновесного состояния (при котором число молекул каждой компоненты газовой смеси в единице объема будет везде одинаково). Следует отметить два обстоятельства; во-первых, эти процессы не связаны с теплообменом между системой и окружающей средой и могут протекать как при наличии, так и при отсутствии такого теплообмена; во-вторых, эти процессы ведут всегда к *увеличению* энтропии.

В частности, вычислим изменение энтропии, которое происходит при неравновесном адиабатическом расширении идеального газа. Допустим, что в начальном равновесном состоянии N молекул газа занимают часть объема сосуда (V_1), а в другой части сосуда (с объемом V_2), отделенной перегородкой, существует вакуум. Удалив перегородку, можно вызвать расширение газа. Спустя некоторое время газ перейдет в равновесное состояние с одинаковыми повсюду температурой и давлением. При таком «расширении в вакуум» внешняя работа $A = 0$,

теплообмен $Q = 0$ и тогда, согласно первому закону термодинамики, $U_2 = U_1$. Для идеального газа это означает $T_2 = T_1$. Если применить к этому процессу формулу (2.54), то изменение энтропии будет равно

$$\Delta S = Nk \ln \frac{V_1 + V_2}{V_1}. \tag{2.55}$$

Однако формула (2.54) выведена для процессов, протекающих *равновесно и с теплообменом*, тогда как рассматриваемый процесс протекает неравновесно и без теплообмена. Но так как изменение энтропии зависит только от начального и конечного состояния газа и не зависит от процесса перехода, то для равновесного изотермического процесса с теплообменом, для которого получена формула (2.54), и адиабатического расширения в вакуум (с последующим переходом в равновесное состояние) изменения энтропии будут одинаковыми, если одинаковы начальные и конечные состояния газа.

Для расчета изменения энтропии, вызванного процессами теплопроводности и диффузии, рассмотрим простейший пример. Допустим, что в двух частях газа с объемами V_1 и V_2 содержатся одинаковые числа N *различных* молекул при температурах T_1 и T_2 (задание числа молекул, объема и температуры идеального газа определяет также его плотность $n = N/V$ и давление $p = nkT$). Для того чтобы сначала исключить изменение энтропии, вызванное перемешиванием различных молекул, допустим, что эти объемы отгорожены подвижной и теплопроницаемой перегородкой. Тогда выравнивание давлений и плотностей будет осуществляться путем перемещения перегородки, а выравнивание температуры — благодаря теплопроводности через перегородку. Обозначим равновесные значения объема, температуры и плотностей в обеих частях газа через V_0, T_0 и n_0. Изменение энтропии всего газа равно сумме изменений энтропии в каждой части. В этом случае можно применять формулу (2.54), так как между обеими частями газа имеет место равновесный (медленно протекающий) теплообмен; получим

$$\Delta S = \Delta S_1 + \Delta S_2 = Nk \ln \frac{V_0}{V_1} \left(\frac{T_0}{T_1}\right)^{i/2} + Nk \ln \frac{V_0}{V_2} \left(\frac{T_0}{T_2}\right)^{i/2}. \tag{2.56}$$

Заменим отношение объемов на отношение плотностей: $V_0/V_1 = n_1/n_0$, $V_0/V_2 = n_2/n_0$; тогда

$$\Delta S = Nk \ln \frac{n_1 n_2}{n_0^2} \left(\frac{T_0^2}{T_1 T_2}\right)^{i/2}. \tag{2.57}$$

Этот метод расчета применим и в более общем случае, когда в начальном состоянии газа имеется сложное неравновесное распределение температуры, давления и плотности в пределах объема газа. Весь объем газа можно разделить подвижными и теплопроницаемыми перегородками на элементарные объемы ΔV_i, настолько малые, что в их пределах температуру, давление и плотность можно с удовлетворительным приближением считать везде одинаковыми. Разумеется, эти объемы не следует брать слишком маленькими; каждый из них должен содержать достаточно большое число частиц, чтобы можно было говорить о температуре, давлении и плотности газа в различных местах этих

объемов. Допустим теперь, что газ переходит из одного неравновесного состояния в другое, соседнее, неравновесное состояние. В каждом элементарном объеме ΔV_i параметры состояния p, n, и T будут несколько изменяться: предположим по-прежнему, что в пределах каждого объема ΔV эти параметры везде одинаковы. Тогда можно рассчитать по формуле (2.57) изменения энтропии ΔS_i в каждом из элементарных объемов и, суммируя их, найти общее изменение энтропии всего газа.

Если *изолированный* газ переходит из одного неравновесного состояния в другое (следовательно, приближается к равновесному состоянию), то, согласно таким расчетам и соответствующим измерениям, энтропия газа *возрастает*.

ПАРАДОКС ГИББСА

Однако вопрос об изменении энтропии, вызванном *перемешиванием различных молекул*, оказывается более сложным. Рассмотрим опять простейший пример; допустим, что в двух половинах объема V, разделенных перегородкой, находятся равные числа (N) молекул *различных* сортов при одинаковых давлениях и температурах. Если перегородку удалить, то начнется взаимная диффузия молекул обоих сортов, в результате которого установится новое равновесное состояние с равномерным распределением молекул каждого сорта по всему объему V. Этот процесс, происходящий при отсутствии теплообмена и при постоянной температуре и давлении, несколько напоминает рассмотренное выше адиабатическое расширение в вакуум, так как молекулы каждого сорта, занимавшие ранее только часть объема ($V/2$), в новом состоянии занимают весь объем V, т. е. проникают в объем, где таких молекул не было. Это обстоятельство дало основание полагать, что изменение энтропии при диффузии молекул можно рассчитывать по формуле (2.55). Суммируя изменения энтропии каждой компоненты, получим

$$\Delta S = \Delta S_1 + \Delta S_2 = Nk \ln \frac{V}{V/2} + N k \ln \frac{V}{V/2} = 2Nk \ln 2. \qquad (2.58)$$

Однако этот рассчет приводит к парадоксу Гиббса: изменение энтропии оказывается *не зависящим* от различия между молекулами. Если это различие постепенно сводить к нулю, то величина ΔS изменяться не будет; формула (2.58) окажется применимой и в том предельном случае, когда по обе стороны перегородки находились одинаковые молекулы. Но диффузия одинаковых молекул не может изменять энтропию газа; такая диффузия происходит, например, в равновесном состоянии газа, когда энтропия сохраняется постоянной. Правильная формула для ΔS должна была бы содержать некоторую величину α, показывающую, насколько отличаются друг от друга перемешивающиеся молекулы; при $\alpha \to 0$ должно быть $\Delta S \to 0$.

Формула (2.58) была получена из (2.55) в предположении, что проникновение молекул одного сорта в объем, где имеются молекулы второго сорта, равносильно расширению в вакуум. При таком предположении изменение энтропии, а следовательно, и сама энтропия

какой-нибудь компоненты газа *не будут зависеть от присутствия молекул других сортов*. Очевидно, что правильная формула для изменения энтропии при диффузии может быть получена на основании более общего допущения о зависимости энтропии одной компоненты газа от всего состава газа и его общего состояния.

Согласно формуле (2.54), изменение энтропии, а следовательно, и сама энтропия пропорциональны *числу частиц*. Это означает, что каждая молекула газа имеет некоторую «удельную энтропию», и если газ переходит из одного состояния в другое, то удельная энтропия одной молекулы изменяется на величину

$$\Delta s = \frac{\Delta S}{N} = k \ln \frac{V_2}{V_1} \left(\frac{T_2}{T_1}\right)^{i/2}. \tag{2.59}$$

Следовательно, сама удельная энтропия отдельной молекулы (s) должна зависеть от общего состояния газа (от объема, в пределах которого может перемещаться рассматриваемая молекула, и от температуры газа). В конечном счете энтропия одной молекулы должна зависеть от наличия в данном объеме других молекул и их суммарной энергии. Допустим, что в газе имеется только одна молекула, отличающаяся от всех остальных; мы должны иметь возможность вычислить ее вклад в общую энтропию газа. Наконец, можно представить себе газ, состоящий только из различных молекул. Очевидно, что энтропия такого газа должна быть представлена как сумма удельных энтропий его молекул:

$$S = s_1 + s_2 + \ldots + s_N. \tag{2.60}$$

Если в газовой смеси имеется N_1 молекул одного сорта, N_2 молекул второго сорта и т. д., то можно группировать удельные энтропии одинаковых молекул и тогда

$$S = N_1 s_1 + N_2 s_2 + \ldots, \tag{2.61}$$

а полная энтропия газовой смеси будет представлена как сумма «парциальных энтропий» составных частей:

$$S = S_1 + S_2 + \ldots \tag{2.62}$$

Лишь в частном случае, когда все молекулы одинаковые, мы получим формулу

$$S = Ns; \quad S_2 - S_1 = N(s_2 - s_1). \tag{2.63}$$

Однако удельная энтропия отдельной молекулы должна зависеть не только от общего состояния газа, но и от массы, формы и размеров самой молекулы. В следующем параграфе будет показано, что число столкновений, испытываемых одной молекулой в единицу времени, также зависит от общего состояния газа и от характеристик данной молекулы; ввиду этого между удельной энтропией и числом столкновений должна существовать определенная связь.

В общем случае изменение энтропии газа может происходить при наличии всех трех процессов: 1) теплообмена (ΔS), 2) выравнивания температур и давления ($\Delta S'$) и 3) перемешивания (диффузии) различных молекул ($\Delta S''$). Следовательно, изменение энтропии должно сос-

тоять из трех частей:

$$S_2 - S_1 = \Delta S + \Delta S' + \Delta S''.$$

Если процесс протекает равновесно, т. е. в промежуточных состояниях газа температуры, давления и плотности (каждой компоненты) в пределах объема одинаковы, то изменение энтропии, вызываемое выравниванием температур, давлений ($\Delta S'$) и диффузией ($\Delta S''$), будет отсутствовать и тогда $S_2 - S_1 = \Delta S = S = \int_1^2 \frac{dQ}{T}$. Если же изучаемый процесс является неравновесным, то $\Delta S' \neq 0$, $\Delta S'' \neq 0$ и тогда

$$S_2 - S_1 > \int_1^2 \frac{dQ}{T}. \tag{2.64}$$

§ 12. ДИФФУЗИЯ, ТЕПЛОПРОВОДНОСТЬ И ВНУТРЕННЕЕ ТРЕНИЕ В ГАЗАХ; ЧИСЛО СТОЛКНОВЕНИЙ И ДЛИНА СВОБОДНОГО ПРОБЕГА МОЛЕКУЛ

Переход идеального газа из неравновесных состояний в равновесное происходит благодаря так называемым «явлениям переноса» — диффузии, теплопроводности и внутреннему трению. В частности, равновесное протекание газовых процессов — изобарического, изотермического и др. — возможно только при наличии в газе этих явлений; благодаря им происходит непрерывное выравнивание плотности, давления и температуры в пределах объема газа. Это выравнивание происходит как при наличии, так и при отсутствии внешнего воздействия на газ. В некоторых случаях явления переноса представляют самостоятельный интерес, например когда необходимо рассчитать количество теплоты, проходящее через прослойку газа, или среднюю скорость течения газа через узкие трубки и т. д.

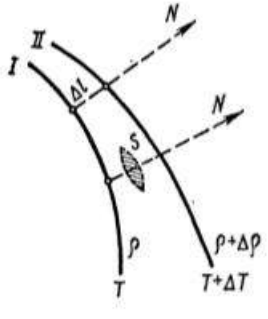


Рис. II.21

Рассмотрим сначала д и ф ф у з и ю; допустим, что плотность газа ρ в различных местах его объема различна. На рис. II.21 изображена условная «поверхность равной плотности» I, в окрестности каждой точки которой плотность имеет одно и то же значение ρ, а II есть такая же поверхность, соответствующая плотности $\rho + \Delta\rho$.

Выберем на поверхности I какую-либо точку и проведем из нее нормаль N к поверхности, направленную в сторону возрастания плотности. Обозначим через Δl расстояние между соседними «поверхностями равной плотности», измеренное вдоль этой нормали. Тогда отношение $\Delta\rho/\Delta l$ или, лучше, предел этого отношения

$$\lim_{\Delta l \to 0} \frac{\Delta\rho}{\Delta l} = \frac{d\rho}{dl} \tag{2.65}$$

будет характеризовать быстроту изменения плотности в направлении нормали к «поверхности равной плотности»; эта величина называется *градиентом плотности* в данном месте.

Выберем некоторую площадку S, ориентированную перпендикулярно N. Так как плотность газа по обе стороны этой площадки различна, то число молекул, переходящих через S в направлении убывания плотности, будет больше, чем в обратном направлении. Если умножить разность этих чисел на массу одной молекулы, то получим результирующий односторонний перенос массы в направлении убывания плотности.

Количество вещества $\mathrm{d}m$, которое диффундирует через площадку S за время $\mathrm{d}t$, пропорционально S, $\mathrm{d}t$ и градиенту плотности:

$$\mathrm{d}m = -D\frac{\mathrm{d}\rho}{\mathrm{d}l}S\,\mathrm{d}t, \tag{2.66}$$

где знак минус указывает, что диффузия происходит в направлении *убывания* плотности, а D — *коэффициент диффузии*, зависящий от вещества газа (формы и размеров его молекул и т. п.).

Аналогичные законы можно написать и для других явлений переноса. Рассмотрим теплопроводность в газе, т. е. односторонний перенос теплоты через какую-нибудь площадку, обусловленный наличием разности температур по обе стороны этой площадки. Допустим, что I на том же рис. II.21 есть *изотермическая поверхность*, проведенная через точки, в которых температура одинакова и равна T, а II есть такая же поверхность, проходящая через точки с температурами $T + \Delta T$. Тогда

$$\lim_{\Delta l \to 0}\frac{\Delta T}{\Delta l} = \frac{\mathrm{d}T}{\mathrm{d}l} \tag{2.67}$$

есть *градиент температуры*, показывающий, как быстро изменяется температура газа в направлении нормали к изотермической поверхности. Теплопроводность в газе объясняется тем, что частицы, проходящие через площадку S в одном направлении, переносят с собой большее количество энергии, чем частицы, движущиеся в обратном направлении. Очевидно, с той стороны, где температура газа выше и, следовательно, средняя кинетическая энергия молекул больше, молекулами газа будет перенесено большее количество теплоты, чем в обратном направлении. При этом предполагается, что числа частиц, ежесекундно проходящих через площадку в одном и обратном направлениях, равны. Результирующий перенос теплоты $\mathrm{d}Q$ через площадку S за время $\mathrm{d}t$ прямо пропорционален S, $\mathrm{d}t$ и градиенту температуры:

$$dQ = -\lambda\frac{\mathrm{d}T}{\mathrm{d}l}S\,dt, \tag{2.68}$$

где λ — *коэффициент теплопроводности*, зависящий от вещества газа; формула (2.68) выражает закон теплопроводности.

Рассмотрим теперь внутреннее трение, возникающее между соседними слоями газа при их относительном движении. Выделим два очень тонких слоя газа, движущихся со скоростями v и $v + \Delta v$, и обозначим через Δl расстояние между этими слоями, изме-

ренное в направлении, перпендикулярном их скорости движения (рис. II.22). По аналогии введем понятие *градиента скорости*

$$\lim_{\Delta l \to 0} \frac{\Delta v}{\Delta l} = \frac{dv}{dl}, \tag{2.69}$$

который показывает, как сильно изменяется скорость течения газа (т. е. скорость упорядоченного движения его молекул) в направлении, перпендикулярном этой скорости. Очевидно, молекулы газа, совершая кроме этого упорядоченного движения еще и беспорядочное тепловое движение, переходят из одного слоя в другой и обратно. Каждая молекула, имеющая массу m_0, при переходе из одного слоя в другой изменяет свой импульс на $m_0\Delta v$. Умножив $m_0\Delta v$ на число молекул, ежесекундно переходящих из одного слоя в другой, получим суммарное изменение импульса, происходящее в каждом слое в единицу времени; согласно второму закону механики, ежесекундное изменение импульса в системе есть приложенная к ней внешняя сила. Таким образом, перенос импульса от одного слоя к другому воспринимается как сила трения, действующая на данный слой со стороны соседних слоев. Эта сила пропорциональна площади S и градиенту скорости (от площади S зависит число молекул, ежесекундно поступающих и покидающих рассматриваемый слой, а от градиента скорости зависит изменение импульса каждой молекулы). Таким образом, з а к о н в н у т р е н н е г о т р е н и я можно записать в виде

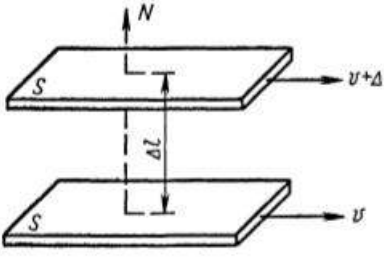


Рис. II.22

$$F = \eta \frac{dv}{dl} S, \tag{2.70}$$

где η — *коэффициент внутреннего трения*, зависящий от вещества газа.

Введенные выше градиенты плотности, теплопроводности и скорости рассматривались как скалярные величины. Однако их удобнее определять как векторные величины; они ориентируются в направлении возрастания соответствующей физической величины. Поэтому, например, диффузия происходит в направлении, противоположном вектору градиента плотности, и т. д.

У идеального газа коэффициенты диффузии, внутреннего трения и теплопроводности соответственно равны:

$$D = \frac{1}{3}\lambda\langle v\rangle;\ \eta = \frac{1}{3}\lambda\langle v\rangle\rho;\ \varkappa = \frac{1}{3}\lambda\langle v\rangle\rho c_V, \tag{2.71}$$

где λ — *средняя длина свободного пробега молекул* от одного столкновения до другого; $\langle v\rangle$ — средняя арифметическая скорость молекул; ρ — плотность газа; c_V — удельная теплоемкость газа при постоянном объеме. Эти формулы получены на основе предположения, что молекулы переходят через площадку S (рис. II.23), испытав по-

следнее столкновение в среднем на расстоянии λ от этой площадки, и переносят с собой массу, а также импульс и кинетическую энергию, которые они имеют на этом расстоянии.

Для расчета λ — среднего расстояния, пробегаемого молекулой от одного столкновения до другого, найдем сначала среднее число столкновений, испытываемых одной молекулой в единицу времени. Допустим, что молекулы газа можно рассматривать как шары, имеющие диаметр σ и движущиеся со средней скоростью $u = \langle v \rangle$. За время dt частица пройдет средний путь $u\,dt$ и столкнется со всеми молекулами, центры которых находятся в пределах объема цилиндра $\pi\sigma^2 u\,dt$ (рис. II.24). Действительно, если центр какой-нибудь молекулы лежит

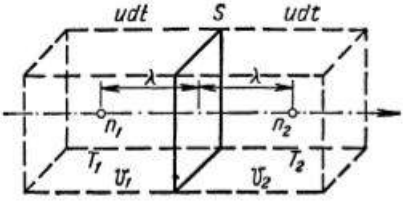


Рис. II.23

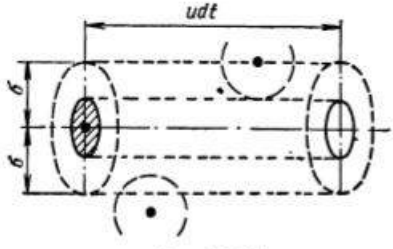


Рис. II.24

на поверхности этого цилиндра или ближе к его оси, то движущаяся молекула столкнется с ней; если же центр молекулы лежит за пределами указанного цилиндра, то столкновения не произойдет. Число столкновений, испытываемое молекулой за время dt, будет, таким образом, равно числу молекул в объеме $\pi\sigma^2 u\,dt$. Однако точный расчет с учетом относительного движения всех молекул газа дает поправочный множитель, равный $\sqrt{2}$. Тогда *число столкновений в единицу времени* (n — число частиц в единице объема)

$$\nu = \frac{\sqrt{2}\,\pi\sigma^2 n u\,dt}{dt} = \sqrt{2}\,\pi\sigma^2 n u, \qquad (2.72)$$

а длина свободного пробега равна отношению среднего пути, проходимого молекулой за единицу времени (u), к числу испытываемых за это время столкновений:

$$\lambda = \frac{u}{\nu} = \frac{1}{\sqrt{2}\,\pi\sigma^2 n}. \qquad (2.73)$$

Коэффициенты теплопроводности и внутреннего трения для каждого газа зависят только от его температуры, но не зависят от плотности газа. Действительно, произведение $\lambda\rho = \lambda n m_0 = \frac{m_0}{\sqrt{2}\,\pi\sigma^2}$ есть для данного газа постоянная величина, поэтому коэффициенты теплопроводности и внутреннего трения зависят только от средней скорости теплового движения, т. е. от температуры.

В изолированном газе выравнивание температуры, давления и плотности (или числа молекул в единице объема n) происходит благодаря столкновениям между молекулами, число которых в единицу

времени, согласно формуле (2.72), будет при этом выравнивании изменяться. Однако процесс выравнивания, т. е. переход газа из неравновесного состояния в равновесное, сопровождается также увеличением энтропии. Вследствие этого должна существовать связь между энтропией газа и ежесекундным числом столкновений между его молекулами.

Перепишем формулу (2.72) с учетом значения средней арифметической скорости $u = \langle v \rangle$ (см. равенство (2.16), в виде

$$\nu = \sqrt{2}\pi\sigma^2 \sqrt{\frac{8}{\pi}\frac{kT}{m_0}}\, n = anT^{1/2} = a\frac{N}{V}T^{1/2}, \tag{2.74}$$

где $a = \sqrt{2}\,\pi\sigma^2 \sqrt{\frac{8k}{\pi m_0}}$ — величина, учитывающая размеры и массу молекулы газа. Из выражения (2.74) для двух различных состояний газа следует, что

$$\frac{\nu_1}{\nu_2} = \frac{V_2}{V_1}\left(\frac{T_1}{T_2}\right)^{1/2}. \tag{2.75}$$

Воспользуемся теперь выражением (2.54) для изменения энтропии газа, в котором заменим отношение объемов на отношение чисел столкновений; получим

$$S_2 - S_1 = Nk \ln \frac{\nu_1}{\nu_2}\left(\frac{T_2}{T_1}\right)^{\beta}, \tag{2.76}$$

где $\beta = (i + 1)/2$ зависит от числа степеней свободы молекул данного газа; в частности, для одноатомного газа $i = 3$, $\beta = 2$.

Формула (2.76) связывает *изменение энтропии* идеального газа с *изменением числа столкновений* между его молекулами (в единицу времени).

В частности, если переход идеального газа в равновесное состояние происходит при постоянной температуре, то увеличение энтропии сопровождается уменьшением ежесекундного числа столкновений между молекулами этого газа.

Формула (2.76) показывает также, что удельная энтропия одной молекулы может быть выражена через число столкновений, которое данная молекула испытывает в единицу времени. Это число зависит от массы, формы и размеров данной молекулы и при данных условиях будет различным для различных молекул.

Глава 3

РЕАЛЬНЫЕ ГАЗЫ, ЖИДКОСТИ И ТВЕРДЫЕ ТЕЛА

§ 13. ОТСТУПЛЕНИЯ ОТ ЗАКОНОВ ИДЕАЛЬНЫХ ГАЗОВ. НАСЫЩЕННЫЕ ПАРЫ. КРИТИЧЕСКОЕ СОСТОЯНИЕ

Измерения показывают, что соотношения, полученные в предыдущей главе для идеального газа, в частности уравнение равновесного состояния

$$pV = \frac{m}{M}RT, \tag{3.1}$$

могут быть применимы и к реальным газам, но только при небольших давлениях и высоких температурах. Например, у азота при температуре 0° С произведение pV отличается от $\frac{m}{M}RT$ для интервала давлений от 1 до 100 атм не более чем на 0,5%, а для давлений, близких к 1000 атм это отличие доходит до 100%.

ДЕЙСТВИЕ МОЛЕКУЛЯРНЫХ СИЛ

Такое расхождение объясняется главным образом действием молекулярных сил. Если взять реальный газ при малом давлении (малой плотности, т. е. малом числе молекул в единице объема газа), то среднее расстояние между молекулами будет велико и силы взаимодействия между ними можно не учитывать; молекулы бóльшую часть времени пребывают на таких расстояниях друг от друга, на которых силы взаимодействия между ними практически равны нулю. Эти силы действуют только в течение коротких промежутков времени, когда происходит «столкновение» молекул и обмен энергиями между ними. Такой газ мало отличается от идеального; некоторое различие можно отметить разве только в продолжительности самих процессов столкновения молекул. Обозначим среднее время от одного столкновения до другого через τ, а продолжительность самого процесса столкновения — через τ_0. У молекул идеального газа (рассматриваемых как упругие шары весьма малых размеров) время взаимодействия при столкновении исчезающе мало по сравнению с временем свободного пробега τ. У реального газа молекулярные силы фактически действуют в течение всего времени пробега τ, однако при очень малой плотности газа время τ_0, в течение которого эти силы оказывают заметное действие на движение молекул, будет мало по сравнению с τ.

При больших давлениях (плотностях) газа молекулы его все время находятся под действием значительных молекулярных сил притяжения; силы отталкивания действуют только в процессе столкновений, когда молекулы газа на короткое время приближаются друг к другу на очень малые расстояния. Можно утверждать, что силы притяжения между молекулами «помогают» стенкам сосуда удерживать газ в определенном объеме, т. е. их действие эквивалентно действию некоторого внешнего давления. По мере увеличения плотности газа, т. е. уменьшения среднего расстояния между его молекулами, силы притяжения все более и более связывают между собой молекулы вещества, ограничивают их свободное перемещение в пределах объема сосуда (увеличивается время пребывания в состоянии, когда велика потенциальная энергия их взаимодействия между собой). При переходе к жидким и твердым состояниям эти силы благодаря малым расстояниям между молекулами имеют настолько большие значения, что они «цементируют» молекулы вещества в небольшом объеме и почти совершенно парализуют их тенденцию «разбежаться», столь характерную для газового состояния. Поэтому в жидком и твердом состояниях объем вещества определяется главным образом молекулярными силами и лишь в значительно меньшей степени — внешним давлением.

Кроме малой плотности реального газа, другим условием применимости уравнения (3.1) является достаточно высокая температура. При очень высокой температуре кинетическая энергия поступательного движения молекул значительно больше потенциальной энергии их взаимодействия между собой. Поэтому свойства газа при высоких температурах близки к свойствам идеального газа, у которого потенциальная энергия взаимодействия молекул отсутствует вовсе.

ОСНОВНЫЕ ХАРАКТЕРИСТИКИ ГАЗОВ

Важнейшими величинами, характеризующими свойства газа, являются:

1) *температурный коэффициент объемного расширения* — относительное изменение объема газа, приходящееся на один градус изменения температуры при постоянном давлении:

$$\beta_p = \frac{1}{V}\left(\frac{\Delta V}{\Delta T}\right)_{p=\text{const}}; \qquad (3.2)$$

2) *температурный коэффициент давления* — относительное изменение давления газа, приходящееся на один градус изменения температуры при постоянном объеме:

$$\beta_V = \frac{1}{p}\left(\frac{\Delta p}{\Delta T}\right)_{V=\text{const}}; \qquad (3.3)$$

3) *удельные или молярные теплоемкости* при постоянном объеме или давлении.

У идеальных газов, согласно закону Гей-Люссака, $\beta_p = \beta_V = 1/273{,}15 = 0{,}003661$ 1/К. У реальных газов эти коэффициенты не совпадают и, кроме того, имеют не одинаковые значения для различных веществ и интервалов температур; так, например, для воздуха: в интервалах температур 0—50° С $\beta_p = 0{,}003675$, $\beta_V = 0{,}003676$ 1/К; в интервалах температур 0—200° С $\beta_p = 0{,}003674$, $\beta_V = 0{,}003672$ 1/К.

Теплоемкость реальных газов заметно зависит от температуры. Например, молярная теплоемкость при постоянном объеме, которая у идеальных газов есть постоянная величина и равна $C_V = iR/2$, у реальных газов при высоких и низких температурах сильно отклоняется от этой величины. Например, у двухатомного идеального газа $C_V = 20{,}79$ Дж/(К·моль), у кислорода и азота эта величина при 0° С равна 20,6, а при 1200° С — 23,4 Дж/(К·моль).

ПАР

Наибольшее расхождение с идеальным газом обнаруживают пары (определение понятия «пар» дано ниже). Различие заключается прежде всего в том, что количество пара, которое может содержаться в данном объеме сосуда при определенной температуре, ограничено, тогда как идеального газа можно вместить любое количество. Допустим, что имеется совершенно пустой сосуд, температура стенок которого поддерживается постоянной. Будем вводить в этот сосуд воду или какую-

нибудь другую жидкость маленькими дозами. Молекулы первых порций жидкости, получив энергию от стенок сосуда, смогут преодолеть связывающие их силы и разлететься по всему объему сосуда. То же самое произойдет с последующими дозами жидкости, однако лишь до тех пор, пока в сосуд не будет введено определенное количество жидкости; последующие дозы ее уже не будут испаряться. В сосуде установится равновесное состояние, когда число молекул, испаряющихся из жидкости в единицу времени, равно числу молекул, возвращающихся в жидкость за то же время. Для того чтобы вызвать дальнейшее испарение жидкости, необходимо повысить температуру.

Если в сосуде при данной температуре имеется предельно возможное количество испарившегося вещества, то говорят, что в этом сосуде находится *насыщающий*, или *насыщенный, пар*. Чем выше температура,

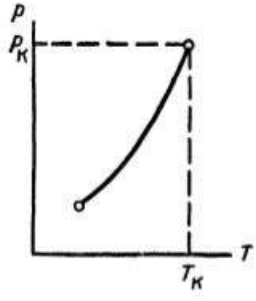


Рис. II.25

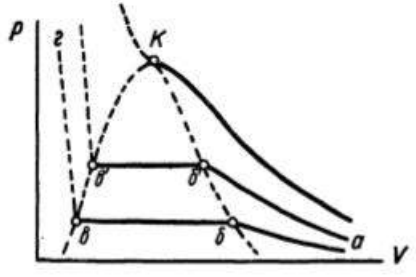


Рис. II.26

тем большее количество насыщенного пара может содержать данный объем, т. е. тем больше плотность и давление этого пара. На рис. II.25 показана связь между давлением насыщенных паров и их температурой; существенно, что кривая $p = p(T)$ заканчивается при некоторых значениях давления $p_к$ и температуры $T_к$, называемых *критическими*. При температуре выше критического значения насыщенный пар не получается и любое количество жидкости, введенное в сосуд, испаряется полностью. Очевидно, при этом средняя кинетическая энергия поступательного движения молекул велика по сравнению с работой, необходимой для разрушения связей между молекулами. Если температура вещества намного выше критической, то кинетическая энергия его молекул значительно превосходит потенциальную энергию их взаимодействия, поэтому состояние вещества будет близко к идеальному газу.

Рассмотрим изотермическое сжатие и расширение паров. Пока пар ненасыщенный, сжатие его сопровождается увеличением давления, но не по закону Бойля — Мариотта ($T = \text{const}$; $pV = \text{const}$), а по более сложному закону, выражаемому на рис. II.26 участком *а—б*. Как только при этом сжатии пар становится насыщенным, дальнейшее уменьшение объема больше не сопровождается увеличением давления: происходит конденсация паров в жидкость (участок *б—в*); эта конденсация заканчивается в точке *в*. Таким образом, в отличие

от газов насыщенный пар вследствие конденсации не поддается изотермическому сжатию. При таком сжатии конденсация паров постепенно уменьшает их количество в данном объеме сосуда: от 100% в точке *б* до 0% в точке *в*. В точке *в* вещество находится в жидком состоянии.

На участке *в*—*г* показано крутое возрастание давления, необходимое для изотермического сжатия жидкости.

При изотермическом расширении процесс идет в обратном направлении. Участок *г*—*в* соответствует постепенному уменьшению внешнего давления на жидкость (при постоянной температуре), вследствие чего жидкость несколько расширяется. В точке *в* жидкость начинает кипеть; дальнейшее увеличение объема уже не сопровождается понижением давления, так как освобождающийся объем занимается насыщенным паром. В точке *б* вся жидкость превращается в насыщенный пар. Расширение насыщенного пара ведет к уменьшению его давления, и пар делается все более и более ненасыщенным.

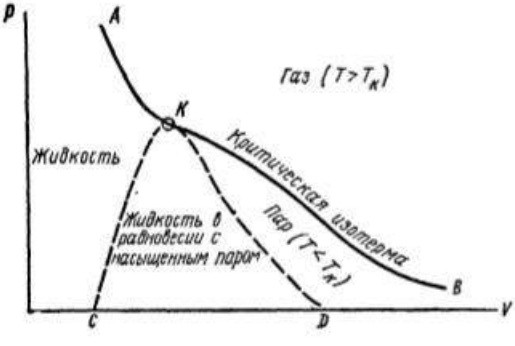


Рис. II.27

Если провести изотермическое сжатие того же количества пара при более высокой температуре, то состояние насыщения наступает при меньшем объеме и большем давлении (точка *б'*). Жидкое же состояние при тех же значениях температуры и давления (точка *в'*) характеризуется бóльшим объемом; при переходе от *в* к *в'* уменьшение объема жидкости от повышения давления меньше, чем увеличение объема от повышения температуры. Таким образом, с увеличением температуры вещества точки *б* и *в* сближаются, т. е. удельный объем насыщенного пара приближается к удельному объему жидкости. При некоторой температуре точки *б* и *в* совпадают, т. е. удельные объемы (или плотности) насыщенного пара и кипящей жидкости становятся равными. Состояние вещества (обозначенное на рис. II. 26 буквой *K*), при котором плотность жидкости и плотность насыщенного пара, находящегося в равновесии с ним, равны, называется *критическим состоянием вещества*. Критическое состояние для каждого вещества характеризуется определенными значениями давления $p_к$, температуры $T_к$ и удельного объема $V_к$.

На рис. II.27 изображены типичные кривые, разграничивающие различные состояния какого-нибудь вещества: кислорода, CO_2, H_2O и т. д.; по оси абсцисс отложен удельный объем этого вещества, т.е. величина, обратная его плотности. Кривая *AKB* есть изотерма, соответствующая критической температуре $T_к$. Точка *K* означает критическое состояние вещества. На кривой *KC* жидкость находится под дав-

лением, равным давлению насыщенного пара при данной температуре; чем ближе к точке K, тем выше температура. На кривой KD вещество находится в состоянии насыщенного пара. Участок критической изотермы KB разграничивает обе области, в которых существуют газообразные состояния вещества: 1) область $T < T_к$ и 2) область $T > T_к$. *Газообразные состояния при температурах ниже критических называются паром*; при температурах выше критических мы имеем реальный газ.

ПЕРЕХОДЫ В СИСТЕМЕ ЖИДКОСТЬ — ПАР — ГАЗ

Рассмотрим изохорические, изобарические и изотермические процессы, при которых вещество переходит из одного состояния в другое (рис. II.28). При изобарическом сжатии *1 → 2*, сопровождаемом охлаждением, вещество переходит из газообразного состояния «скачком» в жидкое, как только температура станет меньше критической. Такой же «скачкообразный» переход из газообразного состояния в жидкое имеет место и при изохорическом охлаждении *1 → 3*. Однако если изобарическое сжатие производить при давлении, меньшем критического (изобара *4 → 5*), то газ сначала переходит в ненасыщенный пар, затем становится насыщенным, а при дальнейшем сжатии (и охлаждении) постепенно конденсируется в жидкость. Изохорическое охлаждение *4 → 6* в отличие от *1 → 3* производится при удельном объеме, большем, чем критический удельный объем; охлаждаемый газ проходит через область ненасыщенного пара, превращается в насыщенный и затем частично конденсируется. При изотермическом процессе *7—8—9—10* в исходном состоянии *7* имеется ненасыщенный пар; путем сжатия этот пар переводится в насыщенное состояние *8*, затем происходит постепенная конденсация, которая заканчивается в точке *9*; на участке *9—10* производится изотермическое сжатие жидкости. Наконец, если изотермический процесс идет при температуре выше критической (участок *11—12*), то вещество все время остается в газообразном состоянии; таким образом, путем изотермического сжатия нельзя превратить газ в жидкость, если температура этого процесса выше критического значения. Для сжижения газа *необходимо охладить его до температуры ниже критической.*

Если изотермический процесс (кривая *7—8—9—10* на рис. II.28) производить с тщательно очищенным веществом, лишенным всяких посторонних примесей, и самый процесс вести достаточно медленно, то можно наблюдать так называемые *метастабильные состояния* (на рис. II.29 участки *2—5*, *3—7* и *9—10*). На участке *2—5* вещество находится в парообразном состоянии и имеет ту же температуру, что и насыщенный пар (точка *2*). Однако по мере удаления от точки *2* давление и плотность пара становятся больше, чем у насыщенного пара, поэтому участок *2—5* (за исключением точки *2*) характеризует состояние пересыщенного пара. Это состояние не является устойчивым; при встряхивании или введении в объем сосуда посторонних тел, способствующих конденсации, часть пересыщенного пара превращается

в жидкость, а остальная часть — в насыщенный пар; вещество из точки *5* переходит в точку *6*.

На участке *3—7* жидкость имеет одну и ту же температуру. Однако во всех точках этого участка, кроме точки *3*, внешнее давление на жидкость меньше, чем давление насыщенного пара, соответствующее этой температуре. В таком состоянии вещество называется *перегретой*

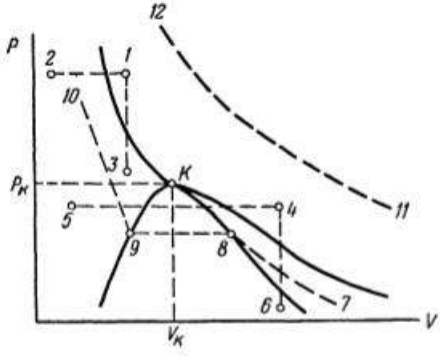


Рис. II.28

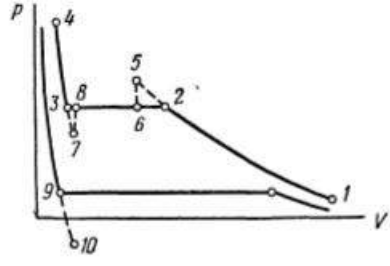


Рис. II.29

жидкостью, так как температура его оказывается выше температуры кипения. Действительно, при переходе $3 \rightarrow 7$ внешнее давление на жидкость уменьшается и температура кипения понижается, однако жидкость, не закипая, сохраняет свою высокую температуру. Если бы в жидкости имелись посторонние тела (пузырьки воздуха и др.), способствующие кипению, то при проведении процесса в направлении $4 \rightarrow 3$, как только внешнее давление на нее сделалось бы равным давлению насыщенного пара (в точке *3*), жидкость стала бы кипеть. При внесении таких тел в перегретую жидкость или при встряхивании ее часть жидкости быстро (почти со взрывом) испаряется; происходит переход $7 \rightarrow 8$.

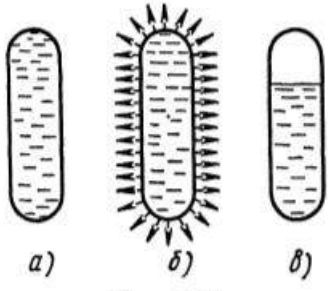


Рис. II.30

При достаточно низких температурах состояние перегретой жидкости может оказаться в области отрицательных внешних давлений, стремящихся увеличить объем жидкости, «растянуть» ее (участок *9—10* и точка *10* на рис. II.29). Такое состояние можно, например, получить, если сосуд, заполненный тщательно очищенной жидкостью, медленно охлаждать (рис. II.30, *а*). В этом случае силы, действующие на пограничные молекулы жидкости со стороны стенок сосуда, осуществляют «отрицательное внешнее давление», не позволяя жидкости уменьшить свой объем при охлаждении (рис. II.30, *б*). При встряхивании связь между частицами жидкости и стенок сосуда нарушается и часть жидкости испаряется (рис. II.30, *в*).

§ 14. УРАВНЕНИЕ ВАН-ДЕР-ВААЛЬСА. ВНУТРЕННЯЯ ЭНЕРГИЯ РЕАЛЬНОГО ГАЗА

При выводе уравнения состояния идеального газа, которое для одного моля имеет вид

$$pV = RT, \tag{3.4}$$

были сделаны два существенно важных предположения:

1) молекулы газа не взаимодействуют между собой; только во время столкновений на короткое время появляются силы отталкивания;

2) собственный объем молекул очень мал по сравнению с объемом сосуда, в котором находится газ.

Для реальных газов, а тем более для жидкостей, эти предположения не могут быть использованы. Были сделаны многочисленные попытки найти такое уравнение состояния для реального вещества, которое могло бы охватить если не все состояния вещества, то хотя бы газообразное, парообразное и жидкое. Из всех предложенных уравнений наибольшей известностью пользуется *уравнение Ван-дер-Ваальса*

$$\left(p + \frac{a}{V^2}\right)(V - b) = RT, \tag{3.5}$$

написанное *для одного моля вещества*. Здесь a и b — постоянные для данного вещества величины; R — по-прежнему универсальная газовая постоянная. Это уравнение отличается от выражения (3.4) двумя «поправками»: величиной a/V^2, учитывающей взаимодействие молекул, и величиной b, учитывающей собственный объем молекул. Выше упоминалось, что действие молекулярных сил притяжения, стремящихся связать молекулы вещества между собой, эквивалентно действию некоторого дополнительного давления, помогающего внешнему давлению удержать вещество в данном объеме. В первом приближении это добавочное давление, обусловленное действием молекулярных сил, можно считать пропорциональным квадрату плотности газа или обратно пропорциональным квадрату удельного объема (объема одного моля газа):

$$p_i = \frac{a}{V^2},$$

где постоянная a учитывает состав и структуру молекул газа. Поправка на объем — величина b, как показывают расчеты, должна равняться учетверенному собственному объему молекул. Величины a и b уравнения Ван-дер-Ваальса имеют различные значения для различных газов; их можно найти, измеряя объем, давление и температуру данного газа в различных состояниях.

Уравнение (3.5) написано для одного моля вещества. Для вещества, имеющего массу m, это уравнение запишется в следующем виде:

$$\left(p + \frac{m}{M^2}\frac{a}{V^2}\right)\left(V - \frac{m}{M}b\right) = \frac{m}{M}RT. \tag{3.6}$$

При постоянной температуре уравнение Ван-дер-Ваальса дает некоторую связь между объемом и давлением, графически представленную на рис. II.31 для четырех различных температур.

Эти кривые называются *изотермами Ван-дер-Ваальса*. При очень высоких температурах они имеют форму, близкую к гиперболе $pV =$

= const, и описывают газообразное состояние вещества (почти идеальный газ). По мере уменьшения температуры форма изотермы несколько изменяется и при некоторой температуре $T_к$ обнаруживает «точку перегиба» K. При еще меньших температурах изотерма Ван-дер-Ваальса приобретает сложную форму и может заходить даже в область отрицательных давлений.

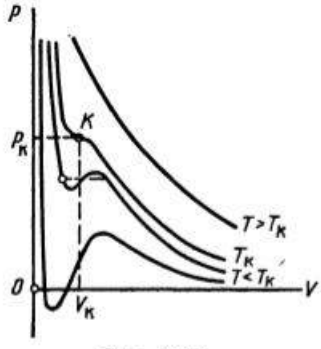


Рис. II.31

Рассмотрим изотермы, соответствующие низким температурам. Измерения показывают (рис. II.32, *а*), что изотермы реального вещества приближаются к изотермам Ван-дер-Ваальса на участках *1—2*, соответствующих жидким состояниям, и на участках *5—6*, соответствующих парообразным состояниям вещества; однако в средней части реальная изотерма идет не по кривой *2—3—4—5*, как этого требует уравнение Ван-дер-Ваальса, а по изобаре *2—5* (в точке *2* имеется только «кипящая жидкость», а в точке *5* — только насыщенный пар). Но если опыты провести с очень чистым веществом, а сжатие, расширение, подвод и отвод теплоты производить достаточно медленно, то можно обнаружить состояния (рис. II.32, *б*), соответствующие участкам *2—3* (перегретая жидкость)

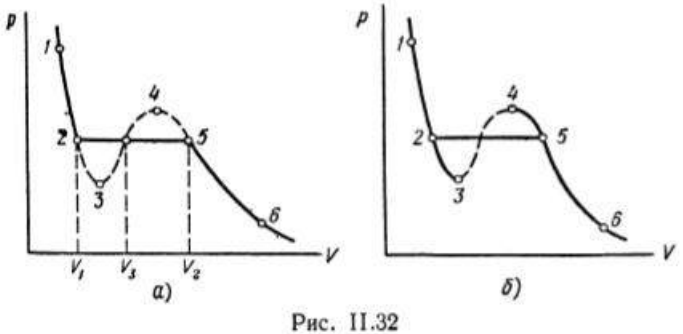


Рис. II.32

и *5—4* (пересыщенный пар). Таким образом, не реализуется только небольшая часть изотермы Ван-дер-Ваальса — участок *3—4*; заметим, что эта часть изотермы соответствует неустойчивым состояниям вещества: при сжатии давление не увеличивается, а уменьшается, т. е. вещество не только не оказывает «сопротивление» сжатию, но, наоборот, само «способствует» этому.

ПАРАМЕТРЫ КРИТИЧЕСКОГО СОСТОЯНИЯ

В хорошем соответствии с результатами измерений уравнение Ван-дер-Ваальса показывает сближение точек *2* и *5* по мере повышения температуры, что позволяет получить из этого уравнения

параметры критического состояния K ($p_к$, $V_к$ и $T_к$), при котором удельные объемы (плотности) кипящей жидкости и насыщенного пара совпадают. Действительно, уравнение Ван-дер-Ваальса является относительно объема алгебраическим уравнением третьей степени, поэтому при фиксированном давлении и температуре имеет три корня:

$$V^3-\left(\frac{RT}{p}+b\right)V^2-\frac{a}{p}V+\frac{ab}{p}=0,$$

или

$$(V-V_1)(V-V_2)(V-V_3)=0.$$

При температурах выше критических это уравнение имеет один действительный и два мнимых корня, т. е. каждому значению давления и температуры соответствует только одно действительное значение объема вещества. При температурах ниже критической это уравнение имеет три действительных корня; один корень дает удельный объем кипящей жидкости (точка *2* на рис. II.32); второй корень есть объем насыщенного пара (точка *5*), третий корень соответствует неустойчивому состоянию вещества. В критическом состоянии все три корня совпадают: $V_1 = V_2 = V_3 = V_к$. Решая уравнение Ван-дер-Ваальса для критического состояния ($p = p_к$, $T = T_к$), получаем:

$$T_к=\frac{8a}{27bR};\quad V_к=3b;\quad p_к=\frac{a}{27b^2}. \tag{3.7}$$

Пользуясь этими соотношениями, можно по известным a и b вычислить критические параметры вещества и, наоборот, по критическим параметрам найти постоянные уравнения Ван-дер-Ваальса. Так как величина b равна учетверенному собственному объему самих молекул вещества, то критический объем в 12 раз больше этого объема. Зная b или $V_к$ и представляя молекулы в виде шариков, можно приблизительно оценить объем и диаметр молекул.

Уравнение Ван-дер-Ваальса не является точным уравнением состояния реальных веществ. Если величины a и b считать постоянными, то обнаруживаются расхождения между измеренными и рассчитанными значениями параметров p, V и T; в частности, реальные изотермы не совпадают с изотермами Ван-дер-Ваальса. Для того чтобы получить хорошее совпадение, приходится для различных областей давления или температуры придавать величинам a и b различные значения. Поэтому уравнение Ван-дер-Ваальса используется не столько для точных расчетов, сколько для выяснения связей между величинами, характеризующими свойства вещества. Примером могут служить полученные из этого уравнения формулы (3.7), в которых обнаружилась связь между критическими параметрами вещества и постоянными a и b.

ПРИМЕНЕНИЯ УРАВНЕНИЯ ВАН-ДЕР-ВААЛЬСА

Другим применением уравнения Ван-дер-Ваальса является получение формулы для внутренней энергии реального газа. Эта энергия должна состоять из двух частей: кинетической энергии поступательного и вращательного движения молекул $E_к$ и потенциальной энергии

их взаимодействия $E_п$. Предполагая, что к реальному газу по-прежнему применима теорема о равномерном распределении энергии по степеням свободы, можно для расчета $E_к$ воспользоваться формулой, полученной для идеального газа (для одного моля):

$$E_к = \frac{iR}{2} T = C_V T. \tag{3.8}$$

Выражение для потенциальной энергии взаимодействия молекул найдем следующим образом. При расширении газа от объема V_1 до V_2 силы внутреннего или молекулярного давления $p_i = a/V^2$ совершают работу, которая равна

$$A = \int_{V_1}^{V_2} p_i \, dV = \int_{V_1}^{V_2} \frac{a}{V^2} \, dV = -\frac{a}{V_2} - \left(-\frac{a}{V_1}\right) = \frac{a}{V_1} - \frac{a}{V_2}.$$

Так как работа внутренних сил равна изменению потенциальной энергии системы, то можно считать, что a/V есть искомое выражение для потенциальной энергии одного моля газа. Эта энергия имеет отрицательный знак, так как молекулярные силы, создающие внутреннее давление p_i, являются силами притяжения (см. ч. I, § 12). Окончательно *внутренняя энергия одного моля реального газа*

$$U = C_V T - \frac{a}{V}. \tag{3.9}$$

Заметим, что в этой формуле объем газа определяет среднее расстояние между молекулами, от которого зависит величина сил, действующих между молекулами, а следовательно, и взаимная потенциальная энергия молекул.

При помощи этой формулы можно получить ряд важных результатов. Рассмотрим изменение температуры реального и идеального газов при адиабатических процессах, протекающих различным образом. При *равновесном адиабатическом расширении* от объема V_1 до V_2 охлаждение реального газа будет несколько больше, чем у идеального. При $Q = 0$ и одинаковой внешней работе A имеем $U_1 - U_2 = A$. Тогда для идеального газа (рассматривается один моль газа)

$$T_1 - T_2 = \frac{A}{C_V},$$

а из формулы (3.9) для реального газа получаем

$$T_1 - T_2 = \frac{A + \left(\frac{a}{V_1} - \frac{a}{V_2}\right)}{C_V}.$$

Различие между охлаждением реального и идеального газов зависит от величины a.

Рассмотрим другое адиабатическое расширение, при котором изолированный от теплообмена газ внешней работы не совершает (так называемое «расширение газа в вакуум»). Такое расширение можно осуществить, соединяя газ с областью, где имеется вакуум (рис. II.33; будем, например, удалять перегородку в направлении, перпендику-

лярном направлению силы давления). На основании первого закона термодинамики $Q = (U_2 - U_1) + A$ для такого расширения имеем $Q = 0$, $A = 0$ и, следовательно, $U_2 = U_1$, т. е. *при адиабатическом расширении без совершения внешней работы внутренняя энергия газа не изменяется.* Этот результат в равной мере относится и к идеальному, и к реальному газам. Однако если газ идеальный, то равенство $U_1 = U_2$ означает $T_1 = T_2$, т. е. при адиабатическом расширении идеального газа в вакуум его температура не изменяется. Если же расширяющийся газ — реальный, то

$$U_1 = C_V T_1 - \frac{a}{V_1};\ U_2 = C_V T_2 - \frac{a}{V_2}. \qquad (3.10)$$

Приравнивая $U_1 = U_2$, получаем

$$T_1 - T_2 = \frac{a}{C_V}\left(\frac{1}{V_1} - \frac{1}{V_2}\right).$$

Так как $V_2 > V_1$, то $T_1 > T_2$, т. е. *реальный газ при адиабатическом расширении в вакуум охлаждается.* Это охлаждение объясняется тем,

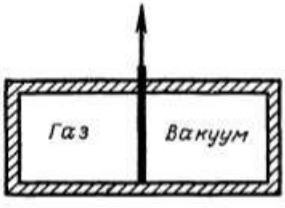


Рис. II.33



Рис. II.34

что часть кинетической энергии теплового движения молекул переходит при этом расширении в потенциальную энергию их взаимодействия.

Рассмотрим еще один пример адиабатического расширения, при котором внешняя работа может варьироваться в широких пределах. Такой процесс можно осуществить при помощи установки, схема которой показана на рис. II.34. Газ при помощи внешней силы, действующей на поршень *I*, «продавливается» через пористую перегородку и, расширяясь, совершает на другой стороне перегородки внешнюю работу перемещения поршня *II*. Давления p_1 и p_2 справа и слева от перегородки поддерживаются постоянными, а весь процесс расширения происходит адиабатически (установка имеет хорошую теплоизоляцию).

Допустим, что слева от перегородки находится газ при давлении p_1, объеме V_1 и температуре T_1, а справа газ отсутствует, т. е. поршень *II* придвинут вплотную к перегородке. Далее допустим, что после продавливания всего газа через перегородку в правой части установки газ имеет давление p_2, объем V_2 и температуру T_2. На основании первого закона термодинамики для этого процесса имеем $Q = (U_2 - U_1) + A = 0$. Внешняя работа, совершаемая газом, состоит из положительной работы при движении поршня *II*:

$$A_2 = p_2 S_2 l_2 = p_2 V_2$$

(S и l — площадь и ход поршня) и отрицательной, совершаемой газом при движении поршня I:

$$A_1 = p_1 S_1 l_1 = p_1 V_1$$

(для внешних сил знаки работ будут обратными). Следовательно,

$$(U_2 - U_1) + (A_2 - A_1) = 0.$$

Подставив в эту формулу выражения (3.10) для внутренней энергии реального газа и рассчитанные из уравнения Ван-дер-Ваальса значения произведений p_1V_1 и p_2V_2, имеем

$$pV = RT - \frac{a}{V} + pb + \frac{ab}{V^2}.$$

Опуская алгебраические преобразования, напишем окончательный результат для разности температур газа после такого расширения:

$$(C_V + R)(T_1 - T_2) = 2a\left(\frac{1}{V_1} - \frac{1}{V_2}\right) - b(p_1 - p_2) - ab\left(\frac{1}{V_1^2} - \frac{1}{V_2^2}\right). \quad (3.11)$$

Знак разности $T_1 - T_2$ зависит от соотношения между положительными и отрицательными величинами в правой части этого выражения.

Рассмотрим два предельных случая; допустим, у газа постоянная a очень мала, а b велика, т. е. силы взаимодействия между молекулами очень слабы, но размеры самих молекул велики. Тогда

$$(C_V + R)(T_1 - T_2) \approx -b(p_1 - p_2).$$

Так как $p_2 < p_1$, то из этого соотношения получается $T_2 > T_1$, т. е. при указанном выше расширении такой газ нагревается. Допустим, что у другого газа постоянная b очень мала, но a велика, т. е. размеры молекул малы, а силы взаимодействия между ними значительные; тогда

$$(C_V + R)(T_1 - T_2) \approx 2a\left(\frac{1}{V_1} - \frac{1}{V_2}\right).$$

Так как $V_2 > V_1$, то $T_2 < T_1$, т. е. такой газ охлаждается. Изменение температуры при расширении реального газа через пористую перегородку называется *эффектом Джоуля — Томсона*. Это изменение достигает нескольких десятых долей градуса. Если газ охлаждается, то, многократно пропуская через пористую перегородку, можно понизить его температуру на десятки градусов. Этим пользуются в холодильной технике для получения низких температур.

Заметим, что величина и знак разности температуры $T_1 - T_2$ зависят не только от значения постоянных a и b, но и от значений объемов и давлений до и после расширения, а следовательно, и от значения начальной температуры газа T_1; в частности, можно выбрать такие начальные значения давления и температуры, чтобы при заданных a и b изменение температуры ($T_1 - T_2$) было равно нулю.

§ 15. ЖИДКОСТИ; ИХ ОСНОВНЫЕ СВОЙСТВА. МОЛЕКУЛЯРНОЕ ДАВЛЕНИЕ И ПОВЕРХНОСТНОЕ НАТЯЖЕНИЕ

Для превращения вещества из газообразного состояния в жидкое следует: 1) *сжимать газ*, т. е. уменьшать среднее расстояние между его молекулами, и 2) *охлаждать газ*, т. е. доводить до температуры ниже критической, а также отводить теплоту конденсации. Уменьшение расстояний между молекулами необходимо, чтобы силы взаимодействия сделались достаточно большими и могли связать молекулы между собой. Охлаждение газа необходимо, чтобы уменьшить среднюю кинетическую энергию молекул по сравнению с потенциальной энергией их взаимодействия. Тогда молекулы вещества не смогут преодолеть действующие между ними силы и разлететься, как это имеет место в газообразном состоянии; они вынуждены будут совершать беспорядочные движения по соседству друг с другом. Однако в отличие от твердого состояния молекулы жидкости имеют еще некоторую свободу передвижения; не отходя на большие расстояния друг от друга, молекулы переходят от «области действия сил» одних соседних молекул к области действия сил других молекул. Этим объясняется существование в жидкостях диффузии, а также малая скорость этой диффузии по сравнению с газами.

ПЕРЕХОД ОТ ГАЗА К ЖИДКОСТИ

В § 13 указывалось, что возможны два способа превращения вещества из газообразного состояния в жидкое:

1) газ путем охлаждения и сжатия сначала доводится до состояния насыщенного пара и затем конденсируется: число связанных между собой молекул в жидкой фазе постепенно увеличивается за счет числа свободных молекул в парообразной фазе.

Этот процесс осуществляется, если внешнее давление меньше критического давления;

2) газ путем охлаждения и сжатия переводится в жидкое состояние «сразу», по всему объему, минуя парообразное состояние, т. е. без разделения вещества на жидкость и насыщенный пар.

Этот процесс осуществляется при давлениях выше критического (см. изобару *1—2* или изохору *1—3* на рис. II.28).

На практике преимущественно осуществляется первый способ. Для этого способа характерно образование сначала небольшого количества жидкой фазы, масса которой в дальнейшем увеличивается за счет пара до тех пор, пока не исчезнет вся парообразная фаза. Образование первых порций жидкости из пара возможно следующим образом. К стенкам сосуда, где производится охлаждение и сжатие пара, может прилипнуть некоторое количество молекул пара; отдав свою кинетическую энергию холодным стенкам сосуда и одновременно сблизившись между собой до малых расстояний, эти молекулы окажутся связанными и между собой и со стенками сосуда, образовав на этих стенках тонкую жидкую «пленку». Толщина этой пленки увеличивается за счет других молекул, летящих к стенке и теряющих здесь свою кинетическую энергию. Так как при конденсации должна

отводиться теплота («теплота конденсации», равная «теплоте парообразования»), то стенки сосуда, через которые происходит этот отвод теплоты, должны все время иметь температуру, несколько меньшую, чем температура пара.

Вследствие наличия теплового движения молекул в самой жидкой фазе происходит испарение: некоторые молекулы, случайно получив достаточную кинетическую энергию, преодолевают силы, связывающие их с остальными молекулами, и переходят снова в пар. Со временем наступает равновесное состояние, когда число молекул, вылетающих из жидкости в единицу времени, становится равным числу молекул, переходящих из пара в жидкость за то же время. Если в таком состоянии произвести некоторое сжатие, то увеличится плотность пара и, следовательно, возрастет число молекул, переходящих из пара в жидкость — до наступления нового равновесия. Таким образом, постепенно сжимая пар, можно перевести его полностью в жидкое состояние.

Кроме стенок сосуда, первоначальное образование жидкой фазы может происходить и на поверхности мельчайших «пылинок», которые в реальных условиях всегда имеются в газах и парах. Эти пылинки являются «центрами конденсации», к которым прилипают молекулы пара, образуя сначала маленькие, а затем и большие капельки жидкости (туман, дождь, облака).

Центрами конденсации могут служить также мельчайшие капельки других жидкостей, заряженные частицы (ионы), сложные молекулы, обладающие несимметричным электрическим полем; например, молекулы серной кислоты вызывают конденсацию водяных паров из воздуха, поэтому над поверхностью крепкой серной кислоты образуется «дымок» из мельчайших капелек воды. Не исключена возможность образования центров конденсации из молекул самого пара, при их случайных тепловых столкновениях, при которых могли бы образоваться небольшие «молекулярные группы» — некоторое количество связанных между собой молекул. При высоких температурах вероятность возникновения таких групп меньше, чем вероятность их разрушения во время столкновений со свободными молекулами, но при низких температурах такие «группы» могли бы существовать долго, и за счет прилипающих к ним свободных молекул превратиться сначала в микроскопические, а затем и в более крупные капельки. Однако вероятность и интенсивность такого процесса конденсации при малых давлениях пара невелика; лишь в критическом состоянии, когда плотности жидкости и ее насыщенного пара равны, наблюдается интенсивное образование таких случайных «скоплений молекул», переходящих при охлаждении в мельчайшие капельки жидкости.

Это явление можно наблюдать, если прозрачную жидкость в стеклянном сосуде путем осторожного нагревания, охлаждения или сжатия провести через критическое состояние; в момент прохождения вещества через критическое состояние наблюдается сильное помутнение, называемое «критической опалесценцией».

Соединение молекул в более или менее прочно связанные группы происходит также и в жидком состоянии вещества, особенно при низ-

ких температурах, вблизи точки затвердевания. Взаимное расположение молекул в этих группах, т. е. структура этих соединений, зависит от характера силового поля молекул; в этом отношении они напоминают элементарные кристаллики твердых тел. Исследования жидкостей при помощи рентгеновских лучей показали, что внутреннее строение их имеет много общего со строением твердых тел, состоящих из большого числа очень маленьких кристалликов с беспорядочным расположением осей. Однако «элементарные кристаллики», которые непрерывно образуются и разрушаются внутри жидкости, имеют весьма малые размеры и в оптический микроскоп не видны. В отличие от твердых тел они могут перемещаться в пределах объема жидкости.

СВОЙСТВА ЖИДКОСТЕЙ

Важными свойствами жидкостей, которые определяются их молекулярным составом и строением, являются *вязкость и текучесть*.

Вязкость жидкостей есть свойство, аналогичное внутреннему трению в газах, и относится к «явлениям переноса», происходящим в жидкой среде. Вязкость различных жидкостей оценивается *коэффициентом вязкости* η, от которого зависит сила трения между двумя соседними слоями жидкости при их относительном движении:

$$F = \eta \frac{dv}{dl} S, \tag{3.12}$$

где dv/dl — градиент скорости; S — поверхность трения. Коэффициент внутреннего трения газов с уменьшением температуры уменьшается (пропорционально скорости молекул, т. е. квадратному корню от температуры), тогда как коэффициент вязкости жидкостей при охлаждении возрастает приблизительно по экспоненциальному закону:

$$\eta = \eta_0 e^{-\Delta W/kT}, \tag{3.13}$$

где k — постоянная Больцмана; ΔW — энергия перехода молекул из одного положения равновесия в другое. Большое значение имеет сильное возрастание вязкости жидкостей при охлаждении до температур, близких к температуре затвердевания, что будет рассматриваться в следующем параграфе.

Текучесть жидкостей оценивается величиной, обратной коэффициенту вязкости. Текучесть и вязкость обусловлены той свободой движения молекул в объеме жидкости, которая еще допускается силами сцепления между ними. В отличие от твердых тел молекулы жидкости не связаны между собой «жестко»; каждая молекула, совершая беспорядочное тепловое движение, одновременно изменяет свое расположение относительно «соседей» и с течением времени перемещается в объеме сосуда.

Несколько упрощая, можно представить себе, что молекулы жидкости совершают колебательное движение вокруг временных положений равновесия. Побывав в окрестности одного положения равновесия некоторое малое время τ, молекула затем «перескакивает» в другое место и совершает колебательное движение вокруг нового положения равновесия в течение такого же короткого времени и т. д. Время τ

пребывания молекулы «в одном месте» иногда образно называют временем «оседлой жизни» молекулы. В течение этого времени молекула находится под действием упругих сил, связывающих и удерживающих ее в окрестности данной точки. Если время действия внешних сил, приложенных к жидкости, меньше, чем τ, то для того, чтобы вызвать деформацию жидкости, т. е. сместить одни молекулы жидкости относительно других, необходимо преодолеть упругие силы, связывающие между собой молекулы жидкости (при сближении молекул появляются силы отталкивания, при удалении — силы притяжения). Поэтому *жидкости по отношению к кратковременно действующим силам ведут себя как упругое твердое тело.* Если же время действия внешних сил значительно больше τ, то молекулы успевают сместиться друг относительно друга (во время их «перескакивания» из одних положений равновесия в другие); в этом случае деформация жидкости происходит за счет теплового движения молекул, без преодоления упругих сил, которые препятствовали бы этой деформации.

Можно утверждать, что чем больше время τ пребывания молекул жидкости «в одном месте», т. е. время их «оседлой жизни», тем больше вязкость и меньше текучесть жидкости. В пределе, когда τ делается равным бесконечности, т. е. когда молекулы вещества оказываются «привязанными» к определенным местам объема и могут совершать в этих местах только колебательные движения, то в этом случае вещество находится в твердом состоянии.

Рассмотрим действие молекулярных сил, приложенных к какой-нибудь молекуле, находящейся внутри жидкости. На эту молекулу действуют главным образом соседние молекулы, расположенные в ее ближайшей окрестности, так как молекулярные силы очень быстро убывают с увеличением расстояния. Равнодействующая всех сил $\mathbf{f}_i$, приложенных к этой молекуле, $\mathbf{R} = \Sigma \mathbf{f}_i$ близка к нулю; она может быть в точности равна нулю, если окружающие молекулы будут расположены относительно данной молекулы совершенно симметрично и на равных расстояниях. Однако неизбежное изменение в относительном расположении соседних молекул вследствие их теплового движения (а также движения самой рассматриваемой молекулы) делает эту сумму $\Sigma \mathbf{f}_i$ отличной от нуля. Таким образом, каждая молекула жидкости все время движется под действием равнодействующей силы $\mathbf{R}$, быстро меняющейся со временем по величине и направлению. Эта равнодействующая направлена в ту сторону, где имеется большая концентрация молекул или ближе расположены соседние молекулы. Следовательно, молекулярные силы взаимного притяжения способствуют концентрации молекул. Однако по мере сближения молекул силы взаимодействия между ними сначала возрастают, а затем уменьшаются до нуля (см. ч. II, § 2); на очень малых расстояниях молекулы начинают отталкивать друг друга. В жидком состоянии молекулы вещества находятся в среднем на расстояниях, когда начинают действовать силы отталкивания. Этим объясняется слабая сжимаемость жидкостей: для того чтобы преодолеть силы отталкивания между молекулами и вызвать заметное изменение объема жидкости, необходимо применять очень высокие давления.

На молекулы, находящиеся в поверхностном слое жидкости, действуют соседние молекулы, расположенные под этим слоем и глубже. Поэтому равнодействующая сил $\mathbf{f}_i$, приложенных к каждой молекуле поверхностного слоя, $\mathbf{R}_i = \Sigma \mathbf{f}_i$ не равна нулю и направлена внутрь жидкости (рис. II.35). Выберем на поверхности жидкости некоторую круговую площадку S; молекулы, охваченные этой площадкой, связаны между собой внутренними силами взаимодействия и поэтому весь этот «мономолекулярный диск» можно рассматривать как некоторое тело. Будем интересоваться внешними силами, действующими на этот «диск» со стороны других молекул жидкости. Во-первых, к каждой молекуле площадки S приложена равнодействующая сила $\mathbf{R}_i = \Sigma \mathbf{f}_i$, направленная внутрь жидкости перпендикулярно площадке (рис. II.35). Если все эти силы $\mathbf{R}_i$ сложить и разделить на величину площади S, то мы получим давление, с которым поверхностный слой действует на остальную массу жидкости:

$$p_{\text{п}} = \frac{\Sigma R_i}{S}. \qquad (3.14)$$

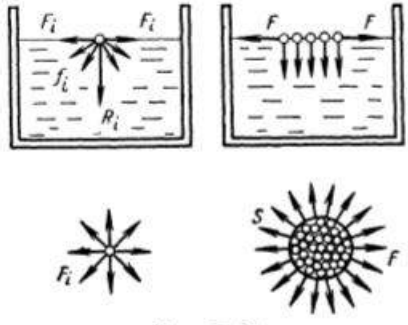


Рис. II.35

Кроме того, на молекулы поверхностного слоя жидкости действуют силы $\mathbf{F}_i$, лежащие в плоскости, касательной к поверхности жидкости. Для всех молекул, лежащих внутри площадки S, все эти силы $\mathbf{F}_i$ взаимно уравновешиваются. Однако для молекул, расположенных вдоль контура площадки S, силы $\mathbf{F}_i$, направленные наружу, являются внешними силами; они перпендикулярны периметру и касательны к поверхности жидкости. Эти внешние силы, растягивающие пленку, называют *силами поверхностного натяжения*. Сила поверхностного натяжения, которая приходится на единицу длины l контура площадки, называется *коэффициентом поверхностного натяжения* или просто *поверхностным натяжением* данной жидкости:

$$\alpha = \frac{F}{l} \qquad (3.15)$$

(α выражается в Н/м или дин/см). У воды при 0° С $\alpha = 0{,}075$ Н/м, у ртути — 0,47 Н/м, у расплавленной меди — 1,12 Н/м.

В этих рассуждениях мы игнорировали воздействие со стороны молекул вещества, находящихся *над поверхностью* рассматриваемой жидкости. Очевидно, давление $p_{\text{п}}$ и коэффициент поверхностного натяжения α должны зависеть также от той среды, которая граничит с данной жидкостью. Если, например, над поверхностью жидкости находится достаточно разреженный газ, то можно пренебречь влиянием молекул второй среды и тогда $p_{\text{п}}$ и α определяются только свойствами данной жидкости (ее плотностью и характером силового поля ее моле-

кул). При наличии над поверхностью жидкости плотного газа (пара) другой жидкости или твердого тела $p_п$ и α зависят также и от свойств этих веществ и, кроме того, от характера взаимодействия между данной жидкостью и соприкасающимися телами. Например, твердые тела могут *смачиваться* или не смачиваться жидкостью; соприкасающиеся жидкости могут *смешиваться* (растворяться, как, например, вода и спирт) или не смешиваться (весьма слабо растворяться, как, например, вода и масло). Обычно в справочных таблицах указываются значения $p_п$ и α в случае, когда над поверхностью жидкости находится ее насыщенный пар.

Коэффициент поверхностного натяжения жидкостей с повышением температуры уменьшается и делается равным нулю при критической температуре, когда исчезает различие между жидкостью и ее насыщенным паром. При наличии примесей или загрязнений в поверхностном слое α заметно уменьшается.

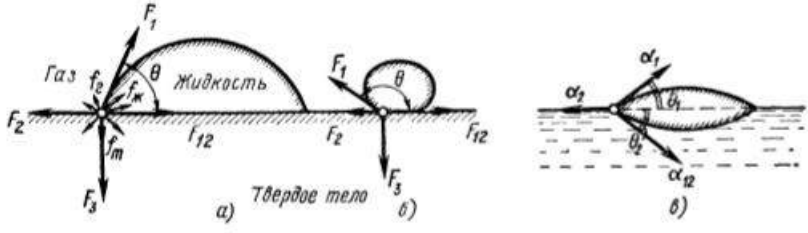


Рис. II.36

Рассмотрим действие молекулярных сил на границе раздела трех сред: жидкости, твердого тела и газа. Силы, действующие на каждую пограничную молекулу со стороны молекул жидкости $\mathbf{f}_ж$, твердого тела $\mathbf{f}_т$ и газа $\mathbf{f}_г$, изображены на рис. II.36.

Рассмотрим проекции этих сил на три направления: касательное к поверхности капли (F_1), на плоскую поверхность тела, на которой находится капля (F_2 и F_{12}), и на направление, перпендикулярное этой плоскости (F_3). При равновесии $F_2 = F_{12} + F_1 \cos\theta$, $F_3 = F_1 \sin\theta$. Образовавшийся при этом угол θ между поверхностью твердого тела и касательной к поверхности жидкости называется *краевым углом.* Если $F_2 > F_{12}$ (рис. II.36, *а*), то говорят, что жидкость смачивает тело: краевой угол меньше $\pi/2$. Если же $F_2 < F_{12}$ (рис. II.36, *б*), то краевой угол будет больше $\pi/2$ и говорят, что жидкость не смачивает твердого тела. Например, стекло смачивается водой, но не смачивается ртутью; ртуть смачивает чистую поверхность железа и т. д. При $\theta = 0$ имеет место полное смачивание; при $\theta = \pi$ — полное или абсолютное несмачивание.

Аналогичные рассуждения позволяют объяснить растекание и форму капель одной жидкости на поверхности другой. Вместо сил F_1, F_2 и F_{12}, действующих по касательным направлениям на одну молекулу, можно взять суммарные силы, действующие на единицу длины контура, т. е. соответствующие коэффициенты поверхностного натяжения: α_1 — первой жидкости, α_2 — второй жидкости и α_{12} — поверх-

ностного натяжения на границе между первой и второй жидкостями (рис. II.36, *в*). Для равновесного существования капли необходимо, чтобы

$$\alpha_2 = \alpha_1 \cos\theta_1 + \alpha_{12} \cos\theta_2. \qquad (3.16)$$

Если α_2 больше, чем правая часть этого выражения, то капля растекается; при этом углы θ_1 и θ_2 уменьшаются и капля, растекаясь, превращается в тонкую пленку на поверхности второй жидкости. В некоторых случаях первые капли, растекаясь по поверхности другой жидкости в виде пленки или растворяясь в ней, уменьшают коэффициент поверхностного натяжения α_2 этой жидкости, так что для следующих капель уже может соблюдаться условие (3.16); например, так образуются жировые капли на поверхности воды.

Силы притяжения между молекулами жидкости создают, как и у реальных газов, некоторое дополнительное давление, которое вместе с внешним давлением производит сжатие жидкости. Для приблизительного определения этого давления (называемого *молекулярным давлением*) можно воспользоваться уравнением Ван-дер-Ваальса, поскольку оно применимо к жидкому состоянию. В этом уравнении дополнительное давление, обусловленное молекулярными силами притяжения, оценивается величиной $p_i = a/V^2$ (V — объем одного моля). Расчет показывает, что для некоторых жидкостей величина p_i превышает давление атмосферы в десятки тысяч раз (у воды $p_i \approx 1{,}7 \cdot 10^9$ Па), поэтому все жидкости уже сильно сжаты внутренними молекулярными силами; для того чтобы вызвать дополнительное уменьшение их объема, необходимо приложить очень большое внешнее давление.

К молекулам, находящимся на поверхности жидкости, можно приложить внешние растягивающие силы, например со стороны молекул стенок сосуда, как это показано на рис. 2.30. Эти силы создадут некоторое отрицательное внешнее давление, растягивающее жидкость. Опыты показали, что к воде можно приложить отрицательное давление до $1{,}5 \cdot 10^7$ Па, что, однако, значительно меньше p_i.

§ 16. КАПИЛЛЯРНЫЕ ЯВЛЕНИЯ

Если поверхность жидкости не плоская, а искривленная, то силы поверхностного натяжения создают дополнительное давление на жидкость, которое прибавляется в давлению p_n или вычитается из него. Допустим, что поверхность жидкости имеет форму сферы радиуса R; выделим на этой поверхности площадку S, опирающуюся на круглое основание S_0 радиуса $r = R \cos\theta$ (рис. II.37, *а*).

Силы поверхностного натяжения, действующие по периметру этой площадки, дадут равнодействующую, перпендикулярную S_0 и равную

$$F_0 = \alpha \cdot 2\pi r \cos\theta \qquad (3.17)$$

(составляющие F' дадут в сумме нуль). Разделив эту силу на площадь основания $S_0 = \pi r^2$, получим дополнительное давление на жидкость от сил поверхностного натяжения, *обусловленное кривизной поверх-*

ности,

$$\Delta p = \frac{F_0}{S_0} = \frac{\alpha \cdot 2\pi r \cos\theta}{\pi r^2} = \frac{2\alpha}{r}\cos\theta = \frac{2\alpha}{R}. \qquad (3.18)$$

Мы рассматривали простой случай, когда поверхность жидкости сферическая. Для более общей формы поверхности получено

$$\Delta p = \alpha\left(\frac{1}{R_1} + \frac{1}{R_2}\right) \qquad (3.19)$$

(формула Лапласа), где R_1 и R_2 — главные радиусы кривизны поверхности в данном месте. (Напомним, что через нормаль к поверхности можно провести множество рассекающих плоскостей; линии пересечения этих плоскостей с поверхностью будут иметь в окрестности точки, в которой проведена нормаль, какие-то радиусы кривизны.

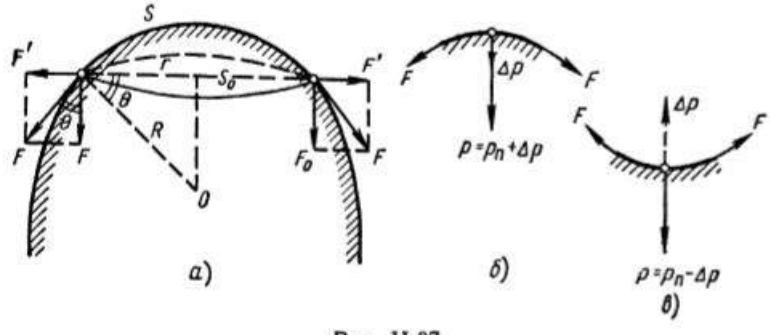


Рис. II.37

Среди множества этих радиусов кривизны выделяются два: R_1 — минимальный и R_2 — максимальный: они лежат во взаимно перпендикулярных плоскостях и называются *главными радиусами кривизны* поверхности в данной ее точке.)

Если поверхность жидкости выпуклая (рис. II.37, *б*, *в*), то дополнительное давление Δp должно быть прибавлено к давлению p_n для плоской поверхности; если же поверхность жидкости вогнутая, то давление Δp будет направлено в противоположную сторону и должно вычитаться из p_n. Таким образом, суммарное давление, обусловленное молекулярными силами, действующими на поверхностный слой жидкости, в общем случае равно

$$p = p_n \pm \Delta p = p_n \pm \alpha\left(\frac{1}{R_1} + \frac{1}{R_2}\right). \qquad (3.20)$$

Дополнительное давление Δp, обусловленное кривизной поверхности жидкости, иногда называется *капиллярным* или *лапласовским* давлением.

Рассмотрим применение формулы (3.20) в капиллярных явлениях.

Поднятие (при смачивании) или опускание (при несмачивании) жидкости в капиллярах (рис. II.38, *а*) можно объяснить тем, что давление на плоской поверхности p_n больше, чем давление на вогнутой

поверхности ($p_1 = p_п - \Delta p$), и меньше, чем давление на выпуклой поверхности ($p_2 = p_п + \Delta p$). Приравняем разность ($p_п - p_1$, $p_2 - p_п$) давлений, т. е. лапласовское давление Δp, гидростатическому давлению столба жидкости высотой h, которое равно ρgh (ρ — плотность жидкости, g — ускорение свободного падения); получим

$$\alpha\left(\frac{1}{R_1}+\frac{1}{R_2}\right)=\rho gh. \tag{3.21}$$

Если капилляры имеют круглое сечение, то $R_1 = R_2 = R$. Радиус кривизны *мениска* в капилляре, очевидно, не совпадает с радиусом самого капилляра r; для сферического мениска, согласно рис. II.37, $r = R\cos\theta$. При этих условиях

$$h=\frac{2\alpha}{\rho gR}=\frac{2\alpha}{\rho gr}\cos\theta. \tag{3.22}$$

Угол θ — краевой угол. Если диаметр капилляра очень мал, то форма мениска близка к сферической и тогда при полном смачивании ($\theta = 0$) можно положить $R \approx r$.

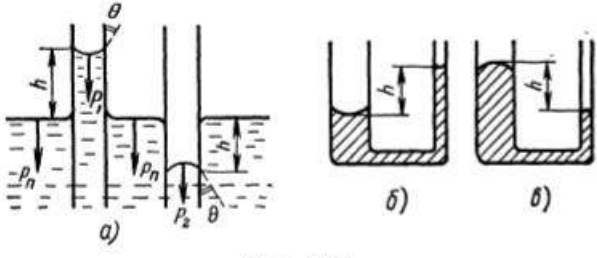


Рис. II.38

В сообщающихся капиллярах уровни жидкости будут различными, как это показано на рис. II.38, *б* (смачивание) и *в* (несмачивание).

При точных расчетах необходимо не только учитывать сложную форму мениска и краевые углы, но также и разность давлений воздуха над поверхностью жидкости в капилляре и вне его. Вследствие земного тяготения давление газов (в частности, давление атмосферного воздуха) уменьшается с высотой. Для малых высот h это убывание можно рассчитать по той же формуле $\Delta p = \rho_г gh$ ($\rho_г$ — плотность газа), по которой рассчитывается гидростатическое давление в жидкостях. Учитывая уменьшение атмосферного давления с высотой, получим для круглых капилляров более точную формулу

$$\frac{2\alpha}{R}+\rho_г gh=\rho gh;\quad h=\frac{2\alpha}{R(\rho-\rho_г)g}. \tag{3.23}$$

Эта формула должна применяться в тех случаях, когда над поверхностью жидкости находится ее пар, имеющий большую плотность (при высоких давлениях).

Формулы (3.22) и (3.23) используются для расчета коэффициента поверхностного натяжения жидкостей по высоте поднятия их уровня

в капиллярах. Помещая капилляры в закрытых сосудах, можно определить α при различных температурах и давлениях вплоть до критического состояния. Применяя тугоплавкие кварцевые капилляры, можно определить коэффициенты поверхностного натяжения также и расплавленных металлов.

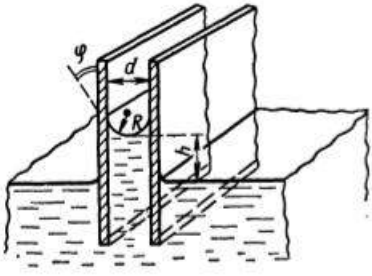


Рис. II.39

Если в жидкость погрузить две близко расположенные плоские пластинки (рис. II.39), то уровень жидкости между ними поднимется (при смачивании) или опустится (при несмачивании) на величину h, зависящую от расстояния d между ними.

Допустим, что в горизонтально расположенном капилляре имеется жидкость с разными температурами на концах (рис. II.40). Так как поверхностное натяжение с повышением температуры уменьшается, то при $T > T'$ $\alpha < \alpha'$. Разность молекулярных давлений на вогнутых поверхностях жидкости будет равна

$$p_1 - p_1' = (p_n - \Delta p) - (p_n - \Delta p') = \Delta p' - \Delta p = \frac{2\alpha'}{R'} - \frac{2\alpha}{R}.$$

Пренебрегая изменениями радиуса капилляра от температуры и полагая $R' = R = r$, получим

$$p_1 - p_1' = \frac{2}{r}(\alpha' - \alpha). \qquad (3.24)$$

Так как $\alpha' > \alpha$, то $p_1 > p_1'$. Если эта разность давлений достаточна для преодоления сил вязкости, то жидкость будет перемещаться к холодному концу капилляра; несмачивающая жидкость (выпуклый мениск) перемещается к теплому концу.

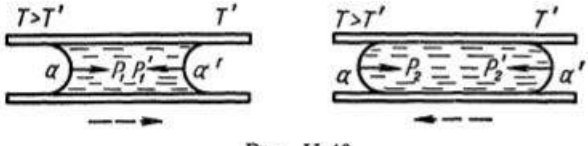


Рис. II.40

Рассмотрим вытекание смачивающей жидкости из капилляров (рис. II.41). Так как на конце капилляра образуется выпуклый мениск и $p_2 > p_1$, то для вытекания жидкости из капилляра необходимо, чтобы сумма p_1 и гидростатического давления ρgh была больше p_2. Из условия равновесия $p_1 + \rho gh = p_2$ можно получить максимальное значение h, при котором жидкость вытекать не будет:

$$\left.\begin{aligned} \rho gh = p_2 - p_1 = \left(p_n + \frac{2\alpha}{R}\right) - \left(p_n - \frac{2\alpha}{R}\right) = \frac{4\alpha}{R}, \\ h = \frac{4\alpha}{\rho g R}. \end{aligned}\right\} \qquad (3.25)$$

Действие лапласовского давления и сил поверхностного натяжения проявляется также в процессе образования капель жидкости, причем отделение каждой капли происходит не путем разрыва жидкости по сечению шейки, а путем постепенного уменьшения радиуса шейки

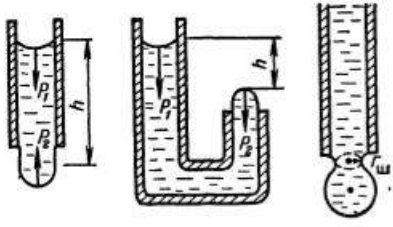


Рис. II.41

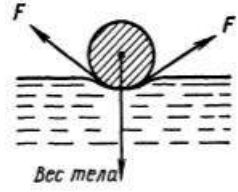


Рис. II.42

$r_ш$ до нуля. Плавание небольших тел на поверхности жидкости объясняется лапласовским давлением со стороны искривленной поверхности жидкости (рис. II.42).

Для того чтобы сплющить каплю несмачивающей жидкости между двумя пластинками (рис. II.43), необходимо приложить внешние силы F, которые преодолевали бы разность давлений Δp выпуклой

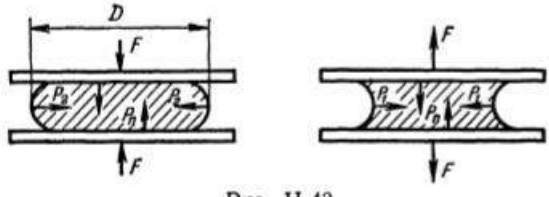


Рис. II.43

и плоской поверхностей жидкости; их можно рассчитать, пользуясь формулой Лапласа:

$$F = \Delta p S = \alpha \left(\frac{1}{R_1} + \frac{1}{R_2} \right) \frac{\pi D^2}{4}, \qquad (3.26)$$

где R_1 — наибольший радиус кривизны; его можно принять равным $D/2$, а R_2 — наименьший радиус, принимаемый равным половине расстояния между пластинками. Если жидкость смачивает пластинки, то они будут притягиваться друг к другу; при малой толщине пленки эти силы могут быть значительными, поэтому смоченные стеклянные пластинки предпочитают отделять друг от друга скользящим движением вдоль плоскости пленки.

§ 17. ИСПАРЕНИЕ И КИПЕНИЕ ЖИДКОСТЕЙ; КОНДЕНСАЦИЯ ПАРОВ

Рассмотрим явления, в которых происходит изменение свободной поверхности жидкости. Прежде всего заметим, что под действием одних только молекулярных сил притяжения жидкость принимает форму, имеющую при данном объеме минимальную поверхность, т. е. форму

шара. Действительно, силы, приложенные к молекулам в поверхностном слое жидкости, имеют равнодействующую, отличную от нуля и направленную внутрь жидкости; поэтому число молекул, которые «вынуждены» оставаться в поверхностном слое жидкости, должно быть минимальным, следовательно, вся поверхность жидкости (при данном объеме) должна иметь наименьшее возможное значение. *При отсутствии внешних сил форма жидкости будет шарообразной.* Наличие различных внешних сил, в частности земного тяготения, искажает форму жидкости. Для больших масс поверхностные молекулярные силы оказываются незначительными по сравнению с объемными силами тяготения, поэтому форма жидкости определяется главным образом силой тяжести: жидкость принимает форму сосуда, в который она налита, а свободная поверхность перпендикулярна направлению силы тяжести. По мере уменьшения массы жидкости значение поверхностных сил возрастает, форма капель, находящихся на поверхности твердых и жидких тел (см. рис. II.36), определяется уже главным образом поверхностными силами. Очень маленькие капельки жидкости (ртуть, вода) имеют почти правильную шарообразную форму.

ИЗМЕНЕНИЕ ЭНЕРГИИ ПОВЕРХНОСТНОГО СЛОЯ

При увеличении свободной поверхности жидкости необходимо перенести соответствующее количество молекул из глубины жидкости на поверхность. При этом должны быть преодолены силы, действующие со стороны остальных молекул: внутри жидкости они в среднем уравновешены, но на поверхности дают равнодействующую R_i (см. рис. II.35). Работа, которая при этом совершается молекулярными силами, имеет отрицательный знак. Напомним, что работа внутренних сил равна изменению потенциальной энергии системы. Поэтому, если внутренние молекулярные силы совершают отрицательную работу, то это означает увеличение потенциальной энергии жидкости. Наоборот, при уменьшении свободной поверхности жидкости, когда молекулы переходят из поверхностного слоя жидкости в ее глубину, молекулярные силы взаимодействия совершают положительную работу, вследствие чего потенциальная энергия жидкости должна уменьшаться.



Рис. II.44

Из суммарной потенциальной энергии взаимодействия всех молекул жидкости можно выделить ту потенциальную энергию, которая приходится на долю только поверхностного слоя жидкости. Обозначим эту энергию через $U_{пов}$; она изменяется при увеличении или уменьшении поверхности жидкости. Для расчета этой энергии рассмотрим простой опыт с мыльным пузырем (рис. II.44). Наружная (выпуклая) и внутренняя (вогнутая) поверхности сжимают воздух в пузыре с давлением $2\alpha/r_1 + 2\alpha/r_2 \approx 4\alpha/r$ (так как $r_1 \approx r_2$), поэтому давление p воздуха в пузыре должно быть больше наружного p_0 на

$$\Delta p = \frac{4\alpha}{r}. \qquad (3.27)$$

Допустим, что, впуская воздух в пузырь, мы увеличим его радиус на Δr; если Δr мало, то $\Delta p \approx \text{const}$, и тогда работа расширения пузыря будет равна произведению Δp на изменение объема:

$$\Delta A = \Delta p \, \Delta V = \Delta p \cdot 4\pi r^2 \Delta r = \alpha \cdot 16\pi r \, \Delta r.$$

Найдем отношение этой работы к изменению поверхности мыльной пленки (наружной и внутренней) $\Delta S = 2\,[4\pi\,(r + \Delta r)^2 - 4\pi r^2] \approx \approx 16\pi r \Delta r$:

$$\Delta A = \alpha \, \Delta S; \quad \frac{\Delta A}{\Delta S} = \alpha. \tag{3.28}$$

Таким образом, *коэффициент поверхностного натяжения численно равен работе, совершаемой при изменении поверхности жидкости на единицу площади*; α выражается в Н/м или Дж/м^2.

Заметим, что при перемещении молекул жидкости из глубины на поверхность преодолеваются силы, препятствующие этому движению, поэтому средняя кинетическая энергия переносимых на поверхность молекул уменьшается, и жидкость несколько охлаждается. Наоборот, при уменьшении поверхности жидкости молекулярные силы, совершающие положительную работу, вызывают увеличение средней кинетической энергии молекул, переходящих с поверхности в глубину, и жидкость нагревается.

Таким образом, увеличение или уменьшение поверхностного слоя жидкости на ΔS сопровождается не только совершением внешней работы $\Delta A = \alpha \Delta S$, но и переходом потенциальной энергии взаимодействия молекул в кинетическую энергию их теплового движения и обратно. Обозначим изменение энергии теплового движения молекул жидкости в этих процессах через ΔQ. Тогда

$$\Delta U_{\text{пов}} = \Delta A + \Delta Q. \tag{3.29}$$

Если, например, происходит сокращение поверхности жидкости и ее энергия уменьшается на $\Delta U_{\text{пов}}$, то часть $\Delta U_{\text{пов}}$ расходуется на совершение внешней работы, а остальная часть переходит в энергию теплового движения молекул. При расширении поверхности жидкости увеличение ее потенциальной энергии на $\Delta U_{\text{пов}}$ частично происходит за счет работы внешних сил ΔA, частично — за счет кинетической энергии теплового движения молекул.

Внешняя работа $\Delta A = \alpha \Delta S$ может быть совершена при различных тепловых режимах, в которых находится жидкость; в частности, изменение поверхностного слоя жидкости можно произвести адиабатически (тогда температура жидкости, а следовательно, и α будут изменяться) или изотермически ($\alpha = \text{const}$). Та часть полной потенциальной энергии поверхностного слоя жидкости $U_{\text{пов}}$, которая превращается в механическую работу при изотермическом сокращении поверхности, называется *свободной энергией* этого слоя. Иногда коэффициент поверхностного натяжения α определяют как свободную энергию одного квадратного сантиметра поверхностного слоя жидкости.

Весьма важными поверхностными процессами, происходящими на границе между жидкостью и газом или паром, являются испарение, кипение и конденсация. Рассмотрим равновесие между жидкостью и ее насыщенным паром в случае, если поверхность жидкости не плоская, а искривленная. Допустим, что в сосуде имеется насыщенный пар и жидкость, в которую погружены два капилляра: один — смачиваемый, другой — не смачиваемый этой жидкостью (рис. II.45). В равновесном состоянии между жидкостью и ее насыщенным паром давление этого пара убывает с высотой подобно тому, как уменьшается с высотой давление воздуха над поверхностью Земли. Разность давлений пара на плоской поверхности жидкости p_0 и на высоте h равна «гидростатическому» давлению столба пара высотой h, т. е. $p_0 - p = \rho_п g h$, где $\rho_п$ — плотность пара; g — ускорение силы тяжести.

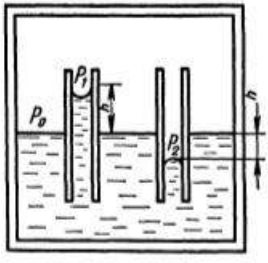


Рис. II.45

Вогнутый мениск находится в равновесии с насыщенным паром, имеющим давление $p_1 < p_0$. Так как высота поднятия жидкости в капилляре (круглого сечения) $h = \frac{2\alpha}{(\rho_ж - \rho_п) gR}$, где $\rho_ж$ и $\rho_п$ — плотности жидкости и пара, а R — радиус кривизны мениска, то давление насыщенного пара над вогнутой поверхностью жидкости

$$p_1 = p_0 - \rho_п g h = p_0 - \frac{2\alpha}{R} \frac{\rho_п}{(\rho_ж - \rho_п)}. \qquad (3.30)$$

Для выпуклого мениска давление насыщенного пара больше, чем для плоской поверхности:

$$p_2 = p_0 + \frac{2\alpha}{R} \frac{\rho_п}{(\rho_ж - \rho_п)}. \qquad (3.31)$$

Различие между давлением насыщенного пара над плоской и искривленной поверхностями жидкости может быть объяснено следующим образом. Равновесие между насыщенным паром и жидкостью означает равенство между числами молекул, ежесекундно переходящих из жидкости в пар и обратно. При вогнутой поверхности переходу молекулы из жидкости в пар препятствует большее число «соседних» (находящихся поблизости) молекул, чем при плоской поверхности (рис. II.46), поэтому при данной температуре ежесекундное число молекул, которые могли бы преодолеть притяжение к вогнутому поверхностному слою жидкости и перейти в пар, меньше чем у плоской поверхности. Наоборот, переход молекул из выпуклой поверхности жидкости в пар потребует меньшей затраты энергии, так как число близко расположенных молекул, препятствующих этому переходу, будет меньше, чем при плоской поверхности, следовательно, число молекул, ежесекундно покидающих выпуклую поверхность, при одинаковой температуре больше, чем это же число для плоской поверхности.

Допустим, что над плоской поверхностью жидкости в атмосфере ее пара находятся капельки этой жидкости. Так как пар является насыщенным для плоской поверхности, т. е. имеет давление p_0, то для капелек с выпуклой поверхностью этот же пар оказывается ненасыщенным; согласно формуле (3.31), выпуклые поверхности находятся в равновесии с паром большего давления, чем плоские поверхности. Вследствие этого капельки жидкости испаряются. Рассмотрим другой пример, когда плоской поверхности жидкости нет и в атмосфере пара имеются только капельки двух размеров: радиусов r_1 и r_2. Допустим, что пар находится в равновесии с капельками малого радиуса и имеет некоторое давление p_2. Тогда, согласно формуле (3.31), этот пар для капелек большего радиуса будет пересыщенным, так как для них давление насыщенного пара должно быть меньше, чем для капелек малого радиуса (чем больше R, которое для сферических капель равно их радиусу r, тем меньше давление насыщенного пара p_2). Вследствие

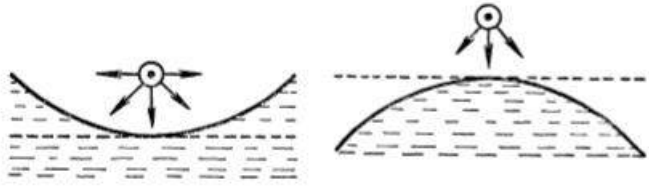


Рис. II.46

этого пар конденсируется на поверхности больших капель; при этом уменьшается давление пара, что вызывает испарение маленьких капель. Большие капли жидкости растут за счет маленьких капелек.

Это обстоятельство должно быть учтено, когда рассматривается конденсация паров на поверхности различных «инородных тел» — центров конденсации (пылинки, мелкие кристаллики твердых тел, микроскопические капельки жидкостей и т. д.). При этом имеют значение не только размеры этих тел, но и их формы. Наиболее активную роль при конденсации играют тела больших размеров и с мало искривленной поверхностью. Если эти тела сферической формы, то можно применить формулу (3.31): чем меньше радиус «центра конденсаций», тем большее давление пара необходимо, чтобы на их поверхности началась конденсация.

Между прочим, если воздух хорошо очищен, то в нем можно получить пар, давление которого в несколько раз превосходит «обычное» давление насыщенного пара при той же температуре.

Тела, на которых конденсируются пары, могут образовать и вогнутые поверхности жидкости, и тогда для конденсаций паров требуется тем меньшее давление пара, чем больше кривизна (или меньше радиус кривизны) этих поверхностей. Например, если в атмосферу пара, насыщенного по отношению к плоской поверхности жидкости, ввести смачиваемые капилляры (например, тела с капиллярными парами), то в этих капиллярах происходит конденсация паров, так как для

вогнутых поверхностей жидкости внутри капилляров этот пар уже пересыщенный.

Таким образом, *давление насыщенного пара есть определенная функция от температуры только для плоской поверхности жидкости*; для искривленных поверхностей давление насыщенного пара, согласно формулам (3.30) и (3.31), зависит еще от кривизны поверхности.

КИПЕНИЕ

Кипение жидкостей также связано с поверхностными явлениями: при кипении происходит испарение жидкости внутрь воздушных пузырьков, которые имеются как в объеме самой жидкости, так и на границе со стенками сосуда. Рассмотрим «механизм» кипения; на рис. II.47 показаны различные стадии развития воздушных пузырьков, «прикрепившихся» к стенке сосуда. По мере испарения жидкости внутрь этих пузырьков давление пара в них повышается, внешнее и гидростатическое давления преодолеваются и пузырек начинает

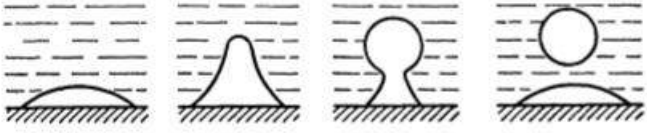


Рис. II.47

расти вверх. Увеличение размеров пузырька происходит особенно быстро, когда давление насыщенного пара внутри пузырька становится равным внешнему давлению (или немного больше). При этом поверхностные силы, деформируя пузырек, отделяют от него некоторую часть, которая архимедовой силой поднимается вверх и освобождает содержащийся в ней пар на поверхности жидкости. Оставшаяся часть пузырька продолжает играть роль «резервуара» (для накапливания пара) и «генератора» новых пузырьков пара.

Таким образом, для кипения необходимы следующие условия: 1) наличие воздушных пузырьков (на границе с твердыми телами), являющихся генераторами пузырьков пара, и 2) жидкость должна быть нагрета до температуры, при которой давление насыщенного пара равно внешнему давлению.

Количество теплоты, уносимое пузырьками пара от стенок сосуда в единицу времени, равно количеству теплоты, подводимой к этим стенкам извне. Если усилить подвод теплоты извне, то отделение пузырьков происходит более интенсивно; небольшое повышение температуры жидкости увеличит давление пара внутри пузырьков, а это ускорит процесс роста и отделения пузырьков пара. Этим объясняется, почему при усилении или ослаблении подвода теплоты к кипящей жидкости температура жидкости не изменяется, а изменяется только скорость парообразования. Если жидкость освобождена от воздушных пузырь-

ков, то ее можно нагреть до температур, превышающих температуру кипения при данном внешнем давлении; например, вода, тщательно очищенная от воздуха, при атмосферном давлении может быть перегрета до 105—107° С; вода, полученная в виде больших капель, плавающих внутри специальных масел, не закипает даже при температуре 178° С. Наоборот, введение в жидкость твердых тел, содержащих на своей поверхности большое количество пузырьков воздуха, способствует кипению и препятствует перегреву жидкости.

УДЕЛЬНАЯ ТЕПЛОТА ПАРООБРАЗОВАНИЯ

Для превращения 1 кг вещества из жидкого состояния в парообразное (при кипении или испарении) необходимо сообщить некоторое количество теплоты, называемое *удельной теплотой испарения* или *парообразования*. При обратном переходе из парообразного состояния в жидкое эта теплота должна отниматься от вещества. Удельная теплота испарения (обозначается через r) состоит из двух частей: одна часть (r_1) расходуется на преодоление молекулярных сил, удерживающих молекулы в объеме жидкости; другая часть (r_2) тратится на совершение работы расширения против внешнего давления на жидкость.

Обе эти части, а следовательно, и их сумма $r = r_1 + r_2$ зависят от плотности (или давления) пара над поверхностью жидкости. Например, вблизи критического состояния, когда плотность насыщенного пара приближается к плотности самой жидкости, силы, действующие на молекулу в сторону пара и в сторону жидкости, мало отличаются друг от друга и работа перехода молекул из жидкости в пар очень мала; очевидно, в критическом состоянии $r_1 = 0$. Вторая часть теплоты парообразования r_2 обусловлена тем, что объем 1 кг вещества в жидком состоянии значительно меньше, чем в парообразном (при тех же температуре и давлении). Для расширения вещества от $V_ж$ до $V_п$ при постоянном давлении p потребуется работа

$$A = p(V_п - V_ж) = r_2.$$

В критическом состоянии разность между удельными объемами жидкости $V_ж$ и пара $V_п$ становится равной нулю и поэтому $r_2 = 0$. Уменьшение удельной теплоты парообразования по мере увеличения плотности пара или его температуры можно иллюстрировать следующими данными для воды:

t, °С	0	100	200	300	350	374
r, (Дж/г)	2490	2260	1940	1380	878	0

Зависимость между удельной теплотой испарения, температурой и изменением удельного объема вещества ($V_п - V_ж$) может быть найдена в общем виде, если воспользоваться вторым законом термодинамики. Рассмотрим круговой процесс, состоящий из: 1) изотермически-изобарного превращения единицы массы кипящей жидкости в насыщенный пар (участок *1—2*, рис. II.48); 2) изменения параметров насы-

щенного пара с p_1, V_2, T_1 до p_2, V_3, T_2 (участок *2—3*); 3) изотермически-изобарного превращения насыщенного пара в жидкость (участок *3—4*) и 4) изменения параметров от p_2, V_4, T_2 до p_1, V_1, T_1 (участок *4—1*).

Если разности давлений $p_1 - p_2 = \Delta p$ и температур $T_1 - T_2 = \Delta T$ очень малы, то можно пренебречь подводом и отводом тепла на участках *2—3* и *4—1*, т. е. считать эти процессы почти адиабатическими. Тогда рассматриваемый круговой процесс складывается из двух изотерм и двух адиабат, т. е. является циклом Карно, для которого

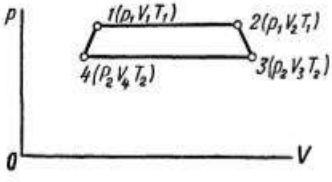


Рис. II.48

$$\frac{Q_1 - Q_2}{Q_1} = \frac{T_1 - T_2}{T_1}.$$

Здесь Q_1 — удельная теплота парообразования r_1 на участке *1—2*. Разность $Q_1 - Q_2$ есть результирующая внешняя работа, совершаемая при этом круговом процессе; она измеряется площадью *1—2—3—4* и при малых Δp и ΔT с достаточной точностью равна

$$Q_1 - Q_2 = \Delta A \approx \Delta p\,(V_п - \Delta_ж),$$

где $V_п$ — удельный объем насыщенного пара в состоянии *2* или *3* ($V_2 \approx V_3$), а $V_ж$ — удельный объем жидкости в состоянии *1* или *4* ($V_1 \approx V_4$). Таким образом, мы получаем приближенное равенство

$$\frac{\Delta p\,(V_п - V_ж)}{r_1} \approx \frac{\Delta T}{T_1}.$$

Если довести Δp и ΔT до бесконечно малых значений, то это равенство сделается точным; перепишем его в виде

$$\frac{dp}{dT} = \frac{r}{T\,(V_п - V_ж)}. \qquad (3.32)$$

Это соотношение, называемое *формулой Клапейрона—Клаузиуса*, связывает между собой удельную теплоту парообразования, температуру насыщенных паров, изменение удельного объема вещества при переходе из жидкого состояния в парообразное и быстроту изменения давления насыщенных паров с температурой dp/dT; Так как $V_п$ больше $V_ж$, то dp/dT всегда положительно, т. е. с увеличением температуры давление насыщенных паров всегда возрастает (см. ч. II, § 13, рис. II.25).

§ 18. КРИСТАЛЛИЧЕСКИЕ И АМОРФНЫЕ ТЕЛА. КРИСТАЛЛИЗАЦИЯ, ПЛАВЛЕНИЕ И ИСПАРЕНИЕ ТВЕРДЫХ ТЕЛ

Для превращения вещества из жидкого состояния в твердое необходимо только охлаждение, т. е. отвод от вещества некоторого количества теплоты. При этом переходе происходит незначительное уменьшение объема тел, но у некоторых веществ (вода, висмут) наблюдается даже увеличение объема. Следовательно, переход вещества из жидкого в твердое состояние означает не столько сближение молекул, сколько

дальнейшее связывание их между собой, ограничение свободы их теплового движения в пределах объема вещества. Для этой цели от молекул вещества и должно быть отнято некоторое количество энергии.

Известны два различных процесса перехода вещества из жидкого состояния в твердое:

1) *затвердевание вследствие кристаллизации вещества.* В этом случае в жидкости сначала появляются мельчайшие кристаллики, содержащие небольшое число молекул, правильно расположенных друг относительно друга и прочно связанных между собой. Затем, по мере отвода тепла, эти кристаллики начинают расти за счет прилипающих к ним молекул жидкой фазы до полного исчезновения этой фазы;

2) *затвердевание вследствие постепенного увеличения вязкости жидкости.* При этом у некоторых веществ кристаллизация совсем не происходит; такие вещества называются *аморфными*; к ним относятся воск, сургуч, смолы. У других веществ (способных к кристаллизации) от быстрого и сильного возрастания вязкости при охлаждении затвердевание наступает до кристаллизации; процесс кристаллизации весьма медленно может происходить после затвердевания. При повышении вязкости сильно затрудняется перемещение молекул в объеме жидкости, необходимое для формирования и роста кристаллов. Вещества, затвердевшие раньше, чем успеет произойти кристаллизация, называются *стеклообразными* (типичным представителем таких веществ является обычное стекло). Известно, что в твердых стеклообразных телах со временем наблюдается очень медленно протекающий процесс кристаллизации («расстекловывание»); этот процесс заметно ускоряется при повышении температуры; так, например, в стеклах, содержащих свинец, при температуре выше 400° С происходит образование большого числа мелких кристалликов.

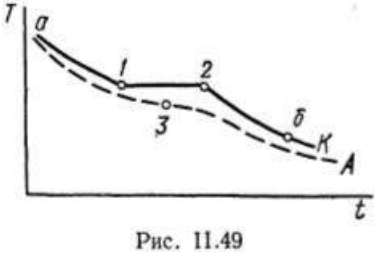


Рис. 11.49

Иногда стеклообразные тела называют «переохлажденной жидкостью», и так как внутри их нет кристалликов, то их также относят к разряду аморфных тел. К аморфным телам относят еще и твердые органические вещества и их производные, молекулы которых представляют собой нитеобразные соединения из огромного числа атомов (от сотен до миллионов), как, например, каучуки и резина, дерево и бумага, кожа, пластмассы, шерстяные, шелковые и хлопковые волокна и т. д.

Различие между указанными выше процессами затвердевания обнаруживается, если вычертить кривую изменения температуры со временем (рис. 11.49). У кристаллических тел температура уменьшается по кривой *К* в направлении *а—1—2—б*; на участке *а—1* вещество находится в жидкой фазе и по мере охлаждения в его объеме происходит формирование зародышей будущих кристаллов. Кристаллизация начинается в точке *1* и происходит при постоянной температуре, которая называется *температурой затвердевания* $T_{\text{затв}}$. На участке *1—2*

одновременно существуют обе фазы — жидкая и твердая; по мере охлаждения содержание твердой фазы увеличивается за счет исчезновения жидкой фазы.

Постоянство температуры при кристаллизации объясняется следующим образом. Образование и рост кристаллов возможны, если от молекул жидкости отнимается некоторое количество энергии. В кристалле жестко связанная молекула совершает только колебательное движение, тогда как при той же температуре в жидкости молекула совершает еще и поступательное движение. Поэтому при кристаллизации от вещества необходимо отвести теплоту, соответствующую поступательному движению молекул. Если ускорить отвод теплоты, то в жидкости увеличивается содержание молекул, обладающих малыми энергиями, и процесс кристаллизации происходит быстрее; молекулы, потерявшие излишек своей энергии, присоединяются к кристаллам, а в жидкости остаются молекулы, обладающие пока еще большой энергией; температура в системе остается постоянной. Если же отвод теплоты от вещества производить медленнее, то процесс кристаллизации замедляется, но температура по-прежнему сохраняется постоянной. Как только процесс кристаллизации заканчивается (точка *2*), дальнейший отвод теплоты сопровождается понижением температуры образовавшегося твердого тела.

При нагревании кристаллических тел процесс протекает в обратном направлении *б—2—1—а*. В точке *2* достигается такая температура $T_{пл}$ (называемая *температурой плавления*), при которой становится возможным отрыв молекул от поверхности нагревшихся кристаллов. В дальнейшем вся подводимая теплота передается отрывающимся молекулам; эти молекулы, получив дополнительную энергию, могут (при той же температуре) совершать уже не только колебательное движение, но еще и поступательное движение в пределах объема вещества. Так постепенно образуется жидкая фаза за счет разрушения твердой фазы. Усиление или ослабление подвода теплоты вызывает только ускорение или замедление процесса плавления; температура при этом не изменяется, так как подводимая теплота целиком «захватывается» отрывающимися от кристаллов молекулами. Если прекратить подвод теплоты, то останавливается и плавление (точнее в веществе устанавливается равновесное состояние, когда числа молекул, переходящих в единицу времени из твердой фазы в жидкую и обратно, равны). После исчезновения твердой фазы дальнейший подвод тепла сопровождается повышением температуры жидкости (участок *1—а*). Заметим, что температура плавления равна температуре затвердевания только при достаточно медленном протекании процессов; при очень быстром подводе или отводе теплоты ввиду конечной скорости роста и разрушения кристаллов возможны небольшой перегрев при плавлении, а также переохлаждение при затвердевании.

У аморфных тел изменение температуры со временем (кривая *А* на рис. II.49) не имеет «ступеньки» с постоянной температурой, а только точку перегиба *3*. Понижение температуры жидкой фазы сопровождается плавным увеличением вязкости, и вещество «незаметно» переходит в твердое состояние. При обратном процессе, когда под-

водится теплота и повышается температура, происходит такое же плавное размягчение вещества до состояния сначала очень вязкой, а затем менее вязкой жидкости. У аморфных тел нельзя указать такую определенную температуру, выше которой можно было бы констатировать жидкое состояние, а ниже — твердое состояние. Выделяется только температура, соответствующая точке перегиба *3* кривой, показывающей изменение температуры со временем. Эту температуру условно называют *температурой размягчения* аморфных тел.

Кристаллические тела характеризуются определенной *удельной теплотой плавления*: количеством теплоты, сообщаемым единице массы вещества для его превращения из твердого состояния в жидкое при постоянной температуре плавления. Эта теплота расходуется для разрушения связей, удерживающих молекулы в кристалле, и для сообщения им дополнительной энергии поступательного движения. При кристаллизации такое же количество теплоты должно быть отнято от жидкого вещества при постоянной температуре затвердевания. У аморфных тел теплота плавления отсутствует, подвод теплоты сопровождается повышением температуры, постепенным увеличением энергии теплового движения молекул, плавным уменьшением времени τ пребывания молекул «в одном месте». В § 3 ч.II указывалось, что с уменьшением времени τ «оседлой жизни» молекул возрастает текучесть жидкости и уменьшается ее вязкость. При плавлении кристаллических тел это время «скачком» изменяется от бесконечности для твердого состояния до некоторого очень малого значения для жидкого состояния; у аморфных тел время «оседлой жизни» молекул изменяется непрерывно.

КРИСТАЛЛИЗАЦИЯ

Для образования кристаллов при затвердевании необходимо существование «зародышей», или «центров кристаллизации», которые, постепенно обрастая молекулами, превращаются сначала в маленькие, а затем и в большие кристаллы. Наблюдения показывают, что соединение молекул в кристаллы очень часто происходит на стенках сосуда, а также вокруг различных инородных тел, имеющихся внутри жидкости: пылинок — мельчайших частиц твердых тел, капелек другой жидкости и даже пузырьков газа. Однако зародыши кристаллов могут возникнуть и из молекул данной жидкости в результате случайного соединения их в более или менее прочные «группы», с определенным расположением друг относительно друга, зависящим от характера их силовых полей. Внутри таких зародышей кристаллов молекулы совершают только колебательное движение, не отрываясь друг от друга.

Для того чтобы эти зародыши были достаточно прочными и не разрушались при взаимодействии с соседними молекулами жидкости, необходимы два условия:

1) молекулы, входящие в состав этих зародышей, должны обладать малыми энергиями, чтобы при колебаниях они не отходили очень далеко друг от друга, так как связывающие их силы взаимодействия очень быстро убывают с расстоянием;

2) число молекул, связанных друг с другом в зародыше, должно быть большим (около тысячи и более), чтобы взаимодействие с соседними молекулами не могло бы вызвать в этой системе больших «деформаций» и разрушения.

Со временем такие зародыши должны присоединять к себе и связывать силами притяжения те молекулы жидкости, которые обладают малыми энергиями, и, наоборот, отражать без особых последствий для себя те молекулы, которые обладают большими энергиями. Заметим, что второе условие, в частности, соблюдается для присутствующих в жидкости мельчайших инородных тел: прилипающие к ним молекулы жидкости благодаря массе и большой поверхности этих тел образуют сразу достаточно прочные и быстро растущие зародыши кристаллизации. Зародыши же, которые формируются без такой «подкладки», из молекул чистой жидкости, могут быть разрушены раньше, чем достигнут необходимой степени «прочности». Между прочим кристаллизация вокруг чрезвычайно маленьких инородных тел иногда используется для тщательной очистки жидкости от этих тел: добиваются, чтобы кристаллы, выросшие на этих частицах, достигли больших размеров, затем их удаляют фильтрованием вместе с находящимися внутри кристалла инородными частицами.

Таким образом, можно предполагать, что не все зародыши кристаллов, возникающие из молекул данной жидкости, продолжают существовать и дают начало кристаллам. Наряду с процессом образования таких зародышей может идти и процесс их разрушения. На наличие таких зародышей еще в жидкой фазе (вблизи температуры затвердевания) указывают результаты изучения жидкостей при помощи рентгеновских лучей. Так как рост зародышей происходит за счет молекул, обладающих малыми энергиями, то скорость увеличения массы этих зародышей, т. е. скорость образования кристалла или скорость кристаллизации, зависит от температуры. Очевидно, с понижением температуры вероятность образования зародышей и скорость их роста должны увеличиваться. Однако необходимо учесть, что с понижением температуры жидкости сильно возрастает ее вязкость, препятствующая перемещению молекул в пределах объема жидкости и, следовательно, задерживающая как образование, так и рост зародышей кристаллов.

Выше уже указывалось, что возрастающая вязкость может сделать вещество твердым до наступления кристаллизации (стеклообразные тела); в других случаях вязкость может приостановить начавшуюся кристаллизацию, так что после затвердевания получается аморфное тело, в объеме которого оказывается «вкрапленным» некоторое число маленьких кристалликов («сферолитов»). Если же все существующие при затвердевании зародыши успеют вырасти в более или менее крупные кристаллы, то твердое тело приобретает *поликристаллическую структуру*, т. е. состоит из большого числа микроскопических кристаллов, беспорядочно ориентированных в объеме тела. При этом в зависимости от числа зародышей в единице объема, скорости кристаллизации, а также от быстроты охлаждения может быть получена либо мелкозернистая, либо крупнозернистая структура. При некоторых предосторожностях возможно получить кристаллы очень больших

размеров; каждый отдельный такой кристалл называется *монокристаллом*. Быстрота охлаждения определяет время, в течение которого происходит рост кристалла. Скорость кристаллизации зависит от состава и строения молекул вещества: «компактные» молекулы малых размеров, имеющие симметричное расположение заряженных частиц и поэтому создающие вокруг себя слабое электрическое поле, в жидкости более подвижны и для них скорость кристаллизации большая. Наоборот, «громоздкие» молекулы, содержащие большое число атомов, имеющие большие размеры или создающие вокруг себя сильное электрическое поле, менее подвижны и для них скорость кристаллизации маленькая.

Если жидкость тщательно очищена от посторонних тел, которые могли бы служить центрами кристаллизации, и если ввиду сложности состава и структуры молекул или большой вязкости жидкости образование собственных зародышей кристаллизации происходит очень медленно, то при осторожном охлаждении можно понизить температуру жидкости ниже той, при которой происходило бы затвердевание при наличии центров кристаллизации. Такая *переохлажденная жидкость* (а в случае растворов — *перенасыщенный раствор*) есть метастабильное состояние вещества; при встряхивании или при введении посторонних тел — центров кристаллизации — в жидкости происходит кристаллизация, переводящая вещество в стабильное для нее (при данной температуре) твердое состояние. Таким образом, наличие или отсутствие центров кристаллизации влияет на температуру, при которой жидкость переходит в твердое состояние. Если в жидкости имеется достаточно большое число центров кристаллизации, то температура затвердевания равна температуре плавления; при отсутствии центров кристаллизации затвердевание происходит при более низких температурах. Вследствие этого при характеристике тепловых свойств веществ обычно приводят не температуру затвердевания, а более определенную — температуру плавления.

ТЕМПЕРАТУРА ПЛАВЛЕНИЯ

Температура плавления зависит от внешнего давления. Если вещество при плавлении расширяется (так ведет себя большинство веществ), то увеличение внешнего давления вызывает повышение температуры плавления. Если же вещество при плавлении сжимается (вода, висмут и др.), то увеличение внешнего давления сопровождается понижением температуры плавления. Эту зависимость можно получить теоретически, если рассмотреть круговой процесс, содержащий переходы из твердого состояния в жидкое и обратно, и воспользоваться вторым законом термодинамики. При этом придется повторить рассуждения, которые были приведены в § 17 ч.II в связи с выводом формулы Клапейрона—Клаузиуса. Допустим, участок *1—2* на рис. II.48 изображает процесс плавления 1 кг вещества, происходящего при внешнем давлении p_1 и температуре плавления T_1, а участок *3—4* есть процесс затвердевания этого же вещества при внешнем давлении p_2 и температуре T_2, бесконечно близких к p_1 и T_1 (ввиду этого остальные части кругового

процесса *2—3* и *4—1* не имеют значения). Тогда Q есть удельная теплота плавления λ; $V_1 \approx V_4$ — удельный объем твердой фазы $V_т$, а $V_2 \approx \approx V_3$ — удельный объем жидкой фазы $V_ж$ и уравнение Клапейрона—Клаузиуса запишется так:

$$\frac{dp}{dT} = \frac{\lambda}{T(V_ж - V_т)}. \qquad (3.33)$$

Если $V_ж > V_т$, то dp/dT положительно (dp и dT имеют одинаковые знаки), следовательно, с увеличением внешнего давления температура плавления также увеличивается. Если же $V_ж < V_т$, то dp/dT отрицательно (dp и dT имеют разные знаки); в этом случае увеличение внешнего давления сопровождается уменьшением температуры плавления.

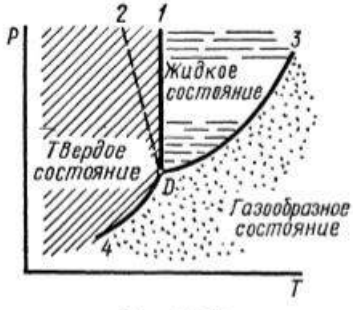


Рис. 11.50

На рис. 11.50 в координатах p и T сплошная линия *1 — D* показывает связь между температурой плавления и внешним давлением для веществ с $dp/dT > 0$, а пунктирная линия *2 — D* — для веществ с $dp/dT < 0$. Слева от этих линий вещество находится при температурах, которые меньше температуры плавления, т. е. в твердой фазе; справа от этих линий вещество находится при более высокой температуре, т. е. в жидком состоянии. Вдоль самих линий *1 — D* или *2 — D* в каждой их точке твердая и жидкая фазы существуют одновременно.

ТРОЙНАЯ ТОЧКА

Точка D соответствует самому низкому давлению, при котором еще возможно превращение вещества из твердого состояния в жидкое и обратно, т. е. равновесное состояние между твердой и жидкой фазами; при меньших давлениях жидкое состояние не существует. На этом же чертеже нанесена кривая *3 — D* зависимости давления насыщенного пара от температуры, о которой упоминалось в § 13 ч.II (см. рис. II.25) и в § 17 (см. формулу (3.32)). Слева от этой кривой вещество находится при температурах, которые меньше температуры кипения, т. е. в жидком состоянии; точки, расположенные справа от этой кривой, соответствуют более высоким температурам и, следовательно, состояниям ненасыщенного пара. Вдоль самой кривой *3 — D* жидкость и насыщенный пар существуют одновременно, т. е. каждая точка этой кривой соответствует равновесному состоянию между кипящей жидкостью и ее насыщенным паром. Точка D этой кривой соответствует самой низкой температуре и самому низкому давлению, при которых еще может существовать жидкость в равновесии со своим насыщенным паром. Точка D называется *тройной точкой* данного вещества. В этой точке одновременно существуют в равновесии друг с другом все три фазы или агрегатные состояния вещества: твердое, жидкое и газообразное (насыщенный пар). Для каждого вещества имеются определенные значе-

ния давления и температуры тройной точки. При меньших давлениях могут существовать только твердое и газообразное состояния. Кривая $4 — D$ показывает зависимость между давлением и температурой насыщенного пара, находящегося в равновесии с твердой фазой.

Непосредственное испарение твердых тел (называемое «возгонкой», или сублимацией) происходит при температурах, меньших, чем температура тройной точки.

Для многих веществ (металлы и др.) нормальная температура (15° С) значительно ниже температуры их тройных точек. Согласно кривой $4—D$ эти вещества при нормальной температуре имеют очень маленькое давление насыщенных паров. Очевидно, чем больше давление насыщенного пара над поверхностью твердого тела, тем интенсивнее может протекать процесс испарения этих тел. Например, легко испаряется йод, который в своей тройной точке (114° С) имеет давление насыщенных паров, равное 90 мм рт. ст. Высыхание мокрого белья на морозе объясняется тем, что находящаяся в нем вода сначала замерзает, а образующийся лед затем испаряется, так как на морозе температура льда меньше, чем температура тройной точки воды (+0,00748° С), а давление насыщенных водяных паров достигает нескольких миллиметров ртутного столба (при —10° С это давление равно 2 мм рт. ст.).

Заметим, что удельная теплота возгонки, т. е. теплота, необходимая для превращения 1 кг вещества из твердого состояния непосредственно в газообразное, оказывается равной сумме удельных теплот плавления и парообразования.

Температура тройной точки воды используется как *реперная температура* термодинамической шкалы Кельвина. Допустим, что идеальное рабочее тело, совершая идеальный цикл Карно, получает теплоту Q_1 при температуре T_1 и отдает теплоту Q_2 при температуре T_2. Измеряя Q_1 и Q_2, можно, согласно § 7 ч. II, найти отношение температур $T_1/T_2 = Q_1/Q_2$. Для того чтобы построить температурную шкалу, необходимо придать T_1 и T_2 определенные («реперные») значения. В качестве таких реперных температур можно взять температуры кипения воды и таяния льда при нормальных условиях, приняв их разность за 100° С. Тогда по измеренному отношению T_1/T_2 и выбранной разности $T_1 — T_2$ можно определить всю температурную шкалу. Очевидно, реперные точки должны воспроизводиться с очень большой точностью, так как они определяют величину градуса. Поэтому В. Томсон (Кельвин) и независимо от него Д. И. Менделеев предложили создать шкалу с одной реперной температурой, в качестве которой выбрана очень точно воспроизводимая (с ошибкой, не превышающей 10^{-4} градуса) тройная точка воды. В Международной системе единиц физических величин (СИ) дано следующее определение градуса: к е л ь в и н — единица температуры по термодинамической температурной шкале, в которой для температуры тройной точки воды установлено значение 273,16° (точно).

Второй реперной точкой является абсолютный нуль, так что кельвин есть 1/273,16 температурного интервала между абсолютным нулем и тройной точкой воды. Так как тройная точка лежит выше

температуры таяния льда на 0,01° С, то связь между температурами по термодинамической (абсолютной) шкале T и по шкале Цельсия t выражается формулой $T = t + 273{,}15$.

§ 19. ТИПЫ КРИСТАЛЛИЧЕСКИХ РЕШЕТОК. ТЕПЛОВЫЕ И УПРУГИЕ СВОЙСТВА ТВЕРДЫХ ТЕЛ

Кристаллизация вещества осуществляется благодаря силам притяжения, которые действуют между его частицами; однако возрастающие на малых расстояниях силы отталкивания препятствуют дальнейшему сближению частиц. Вследствие этого молекулы вещества можно уподобить (в первом приближении) твердым частицам, в частности шарикам определенного радиуса, которые силами притяжения могут быть доведены только до «соприкосновения». Эти силы всегда стремятся расположить частицы вещества как можно ближе друг к другу. При наличии беспорядочного теплового движения всякая возможность максимального сближения соседних молекул реализуется. Благодаря этому кристалл, представляющий собой систему связанных между собой частиц веществ, должен характеризоваться максимально плотной «упаковкой» частиц. Эта идея, называемая *принципом наиплотнейшей упаковки,* позволяет объяснить существующие в природе формы кристаллов с различным расположением частиц в их объеме. При этом частицам различных веществ приписываются различные формы и размеры (радиусы).

При плотной упаковке шарообразных частиц каждая из них оказывается в окружении некоторого числа близлежащих («соседних») частиц, расположенных на одинаковых расстояниях от нее. Это число называется *координационным числом* и имеет значения 12, 8, 6, 4 и 2. Если кристалл составлен из частиц (шаров) двух размеров, то координационное число зависит от отношения их радиусов $r_1/r_2 = a$:

$$\underset{12}{a \approx 1} \qquad \underset{8}{1 > a > 0{,}73} \qquad \underset{6}{0{,}73 > a > 0{,}41} \qquad \underset{4}{0{,}41 > a > 0{,}22} \qquad \underset{2}{a < 0{,}22}$$

Упорядоченное расположение частиц в объеме кристалла образует так называемую *пространственную кристаллическую решетку.* Геометрически эту решетку можно получить, если провести три системы плоскостей, пересекающихся между собой под углами α, β и γ; в каждой системе плоскости параллельны между собой и отстоят друг от друга на равных расстояниях a, b и c. Такие плоскости разбивают объем кристалла на элементарные ячейки (рис. II.51), как бы кирпичики, из которых состоит весь кристалл. В зависимости от соотношения между размерами ребер ячейки a, b и c (т. е. расстояний между плоскостями, образующими эти ячейки), а также от углов α, β и γ различают семь кристаллографических систем:

1) $a = b = c$; $\alpha = \beta = \gamma = 90°$ — правильная или кубическая система,

2) $a = b \neq c$; $\alpha = \beta = 90°$; $\gamma = 120°$ — гексагональная система,

3) $a = b \neq c$; $\alpha = \beta = \gamma = 90°$ — тетрагональная система,

4) $a = b = c$; $\alpha = \beta = \gamma \neq 90°$ — тригональная система,

5) $a \neq b \neq c$; $\alpha = \beta = \gamma = 90°$ — ромбическая система,
6) $a \neq b \neq c$; $\alpha = \gamma = 90° \neq \beta$ — моноклинная система,
7) $a \neq b \neq c$; $\alpha \neq \beta \neq \gamma$ и отличны от 90° — триклинная система.

Точки пересечения плоскостей, образующих пространственную кристаллическую решетку (т. е. вершины элементарных ячеек кристалла), называются *узлами* этой решетки. В этих узлах располагаются частицы вещества; в зависимости от того, какие частицы находятся

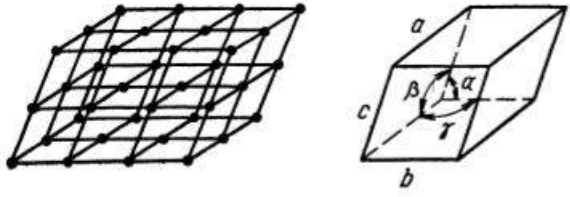


Рис. II.51

в узлах решетки и каков характер сил взаимодействия между ними, различают:

1) а т о м н ы е р е ш е т к и; в узлах этих решеток находятся нейтральные атомы данного вещества. Типичную атомную решетку дает кристалл алмаза, составленный из нейтральных атомов углерода (рис. II.52, *а*). Связь между одинаковыми атомами в молекуле или в кристаллической решетке называется *гомеополярной*;

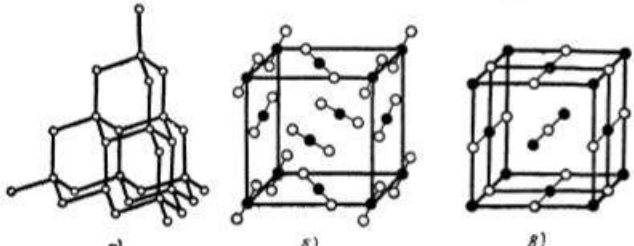


Рис. II.52

2) м о л е к у л я р н ы е р е ш е т к и, в узлах которых расположены нейтральные молекулы вещества; например, кристаллы льда, CO_2 (рис. II.52, *б*) и др.;

3) и о н н ы е р е ш е т к и; в их узлах находятся электрические заряженные частицы — ионы. В решетках типа NaCl (рис. II.52, *в*) соседние ионы имеют противоположные знаки (Na^+ и Cl^-) и притягиваются друг к другу, что обеспечивает прочность кристалла;

4) м е т а л л и ч е с к и е р е ш е т к и, составленные из положительно заряженных ионов металла, находящихся в «атмосфере» электронного газа. Электроны и положительные ионы металла взаимно связывают друг друга; электроны не могут покинуть пределы металла благодаря притяжению к положительно заряженным атомам металла,

а кристаллическая решетка, составленная из этих атомов, не «разлетается» благодаря «цементирующему» действию электронного газа.

Так как силы, обеспечивающие устойчивость и прочность кристаллических решеток металлов, имеют несколько иной, специфический характер, то эти решетки выделяют в особую группу.

ТЕПЛОЕМКОСТЬ КРИСТАЛЛА

Частицы (атомы, молекулы ионы), образующие кристалл, продолжают участвовать в тепловом движении, совершая колебания вокруг своих положений равновесия — узлов пространственной решетки. Так как эти колебания являются беспорядочными, т. е. с одинаковой вероятностью происходят в любом из трех пространственных направлений, то каждой частице следует приписать шесть степеней свободы колебательного движения. Допустим, что кристаллическая решетка составлена из атомов, которые совершают только колебательные движения (молекулы могли бы совершать еще и вращательное движение). Если на каждую степень свободы в среднем приходится одинаковая энергия, равная $\frac{1}{2}\,kT$, то теплоемкость одного моля кристаллического вещества при постоянном объеме можно было бы определить по формуле

$$G_V = N_A \frac{i}{2}\,kT = \frac{iR}{2}.$$

Измерения показали, что у многих простых веществ при обычных температурах *молярная теплоемкость близка к* $3R$ (закон Дюлонга и Пти), что соответствует $i = 6$, например:

Серебро, алюминий	$3{,}1R$
Золото, железо	$3{,}22R$
Олово	$3{,}25R$
Медь	$2{,}99R$

В ионных решетках двухатомных соединений типа NaCl колебательные движения в узлах совершаются отдельными ионами Na и Cl. Один моль такого вещества содержит $2N_A$ атомов, поэтому молярная теплоемкость кристаллической соли должна равняться $6R$; для кристаллических трехатомных соединений эта теплоемкость должна быть равна $9R$, для четырехатомных — $12R$.

Измерения дают следующие значения:

CuO	$5{,}7R$
NaCl	$6{,}1R$
$BaCl_2$	$9{,}4R$
KNO_3	$12{,}2R$

Расхождения между расчетными и измеренными значениями теплоемкостей увеличиваются при переходе к низким температурам. Например, для меди:

t, °C	—240	—186	—39	+50	+180
C_V	$0{,}273R$	$1{,}675R$	$2{,}82R$	$2{,}98R$	$3{,}08R$

В более точной квантовой теории теплоемкостей утверждается, что на каждую степень свободы колебательного движения приходится различное количество энергии в зависимости от температуры и частоты колебаний. Эта теория дает удовлетворительное согласие с экспериментальными данными.

Заметим, что в твердых телах вследствие слабого изменения объема при изобарическом нагревании теплоемкости при постоянных объеме и давлении мало отличаются друг от друга, но это различие возрастает с повышением температуры.

Разность теплоемкостей $C_p - C_V$ равна внешней работе, совершаемой при изобарическом нагревании одного моля данного вещества на один градус. Очевидно, эта работа тем больше, чем больше коэффициент линейного, а следовательно, и объемного расширения тела.

Кристаллы благодаря определенному расположению частиц в их пространственной решетке являются анизотропными телами, т. е. их физические свойства различны в различных направлениях. Если в кристалле в двух каких-нибудь направлениях провести прямые, на которых лежат частицы кристалла, то расстояния между частицами по этим направлениям оказываются различными; тогда все те свойства вещества, которые зависят от расстояний между частицами, будут различаться по этим направлениям.

Аморфные тела являются изотропными, т. е. их физические свойства одинаковы по всем направлениям. Следует отметить, что поликристаллические тела, состоящие из большого числа беспорядочно ориентированных мелких кристалликов, в целом также оказываются изотропными телами, так как в них ни одно направление не выделяется среди других. Например, куски металла не обнаруживают анизотропии, хотя и являются кристаллическими телами; беспорядочная ориентировка анизотропных кристаллических зерен создает общую изотропию тела. Любой порядок, наведенный в ориентировках кристаллических зерен путем механических или других воздействий, сделает этот кусок металла в той или иной степени анизотропным.

Анизотропия кристаллов обнаруживается при измерении следующих величин, характеризующих свойства вещества:

1) коэффициента линейного теплового расширения;
2) коэффициента теплопроводности;
3) модуля продольной упругости; скорости распространения звука;
4) диэлектрической проницаемости (для диэлектриков), скорости распространения света (для прозрачных тел).

ТЕПЛОВОЕ РАСШИРЕНИЕ КРИСТАЛЛОВ

Тепловое расширение кристаллических тел при нагревании объясняется несимметричностью сил взаимодействия между частицами относительно положений равновесия, т. е. узлов кристаллической решетки. Если бы зависимость сил притяжения и сил отталкивания от расстояния между частицами была одинаковой, то увеличение или уменьшение амплитуды колебаний не могло бы отразиться на взаимном расположении узлов кристаллической решетки. Однако при колеба-

ниях частиц вокруг этих узлов силы отталкивания между ними при сближении растут быстрее, чем силы притяжения при удалении. Вследствие этого увеличение амплитуды колебаний частиц при повышении температуры сопровождается смещением положений равновесия, т. е. увеличением расстояний между узлами решетки. Заметим кстати, что в различных направлениях внутри кристалла это увеличение может быть различным (в зависимости от средних расстояний между частицами в рассматриваемом направлении).

Коэффициент линейного расширения

$$\alpha = \frac{\Delta l}{l\,\Delta t}, \tag{3.34}$$

где Δl — изменение длины l кристалла в данном направлении при изменении температуры на Δt. В кристаллах α имеет различные значения в различных направлениях, поэтому *коэффициент объемного расширения* β приходится определять отдельно:

$$\beta = \frac{\Delta V}{V\,\Delta t}, \tag{3.35}$$

где ΔV — изменение объема V кристалла при изменении его температуры на Δt. Для изотропных тел (аморфных или поликристаллических) α одинаково по всем направлениям, поэтому $\beta = 3\alpha$.

Теплопроводность кристаллов объясняется тем, что увеличение амплитуды колебаний частиц в каком-нибудь месте благодаря силам взаимодействия вызывает увеличение амплитуды колебаний соседних частиц. *Благодаря связи между частицами теплота стремится к равномерному распределению по объему кристалла.* Теплопроводность вещества оценивается *коэффициентом теплопроводности* λ в формуле для расчета количества теплоты $\mathrm{d}Q$, которое проходит через площадку S за время $\mathrm{d}t$:

$$\mathrm{d}Q = \lambda \frac{\mathrm{d}T}{\mathrm{d}l}\, S\, \mathrm{d}t \tag{3.36}$$

($\mathrm{d}T/\mathrm{d}l$ — градиент температуры); в анизотропном теле λ различен в различных направлениях. Например, у кристаллического кварца для двух взаимно перпендикулярных осей коэффициенты теплопроводности отличаются почти в два раза. В металлах теплота передается не только колебаниями атомов в кристаллической решетке, но и благодаря электронному газу, участвующему в тепловом движении; этим объясняется высокая теплопроводность металлов.

УПРУГОСТЬ И ПЛАСТИЧНОСТЬ

Важным свойством твердых тел является упругость, проявляющаяся при деформации этих тел внешними силами. Основные виды деформации: продольное (линейное) и всестороннее (объемное) растяжение и сжатие, сдвиг и кручение. Если при устранении внешних сил деформированное тело в точности возвращается к первоначальной форме и размерам, то деформация называется *упругой*, а само тело — упругим телом.

При деформациях твердых тел развиваются внутренние силы, противодействующие внешним силам и стремящиеся вернуть тело к перво-

начальной форме и размерам. Противодействие изменению объема (всестороннему сжатию и растяжению) называется *объемной упругостью*; противодействие изменению формы называется *упругостью формы*. Тело называется *пластичным*, если оно обладает слабой упругостью формы. При малых деформирующих силах твердые тела являются упругими; при очень больших сжимающих давлениях тела обнаруживают пластичность (текучесть).

Упругость и пластичность твердых тел зависят от температуры; пластичные при обыкновенной температуре тела (воск, свинец и др.) делаются упругими при низких температурах; упругие при нормальной температуре тела становятся пластичными при высоких температурах.

Способность тел к изменению формы (пластичность) различна по отношению к силам, действующим продолжительное или короткое время. Многие тела, например металлы (медь, алюминий и т. д.), под действием слабых сил, приложенных в течение очень короткого времени («ударные силы»), ведут себя как упругие тела, полностью восстанавливающие свою форму после устранения деформирующих сил; эти же тела под действием больших сил, действующих продолжительное время, ведут себя как пластичные тела и, подобно жидкостям, иногда «текут». В частности, цинк по отношению к «ударным» силам хрупок, а по отношению к долгодействующим силам является пластичным.

Упругость и пластичность твердых тел обусловлены сложным комплексом. Во-первых, под действием внешних сил происходит деформация кристаллической решетки, т. е. некоторое изменение расстояний между частицами; вследствие этого возбуждаются упругие силы, противодействующие внешним силам. Кроме того, если тело поликристаллическое, то при деформации возможны смещения (а также поворот) отдельных зерен друг относительно друга; внутренние силы развиваются на границах между зернами. Если такое тело многократно подвергалось деформациям, то может произойти частичное размельчение его зерен и, следовательно, изменение упругих свойств всего тела. Крупнозернистые металлы обычно более пластичны, чем мелкозернистые. Прочность кристалла зависит также от наличия посторонних примесей внутри него, а также от наличия трещин на его наружной поверхности.

ЖИДКИЕ КРИСТАЛЛЫ

В последнее время интенсивно изучаются и получают широкое применение ж и д к и е к р и с т а л л ы. Приведем краткие сведения о них.

Физические свойства вещества определяются его *составом* (из каких частиц — атомов, молекул, ионов и т. д. — состоит данное вещество), *структурой* (относительным расположением частиц в объеме данного тела) и *состоянием* (температурой, внешним давлением и т. п.). Частицы некоторых веществ могут быть «сферически симметричными», обладающими одинаковыми свойствами по всем направле-

ниям; другие частицы оказываются «несферическими» (или анизомерными), в частности стержнеобразными. От этой особенности частиц и от характера взаимодействия между ними зависит их относительное расположение в объеме рассматриваемого вещества. В твердых кристаллах частицы образуют в пространстве прочную трехмерную периодическую структуру, которую называют *дальним позиционным порядком*. В жидкостях такая периодичность осуществляется лишь для небольших групп частиц, т. е. имеется *ближний позиционный порядок*. Если же вещество состоит из достаточно жестких несферических частиц, то кроме «позиционного» появляется порядок, связанный с ориентацией частиц. Если во всем объеме вещества существует примерно одинаковая ориентация частиц, то говорят, что в этом объеме имеется *дальний ориентационный порядок*.

Если вещество состоит из физически симметричных частиц, то переход его из твердого состояния в жидкое означает исчезновение «дальнего позиционного порядка»; такой переход называют *позиционным плавлением*. У веществ, состоящих из анизомерных частиц, имеет место также *ориентационное плавление*, соответствующее исчезновению «дальнего ориентационного порядка». При этом возможны три случая:

1) температуры позиционного и ориентационного плавления *совпадают*. Такое совпадение присуще веществам, частицы которых по форме и взаимодействию являются сферически симметричными; для них существуют три отчетливо фиксируемых агрегатных состояния — твердое, жидкое и газообразное;

2) температура позиционного плавления *выше* температуры ориентационного плавления. Фаза вещества, существующая в интервале между этими температурами, называется пластическим кристаллом. Примером пластических кристаллов могут служить галоиды аммония: NH_4Cl и NH_4Br;

3) температура позиционного плавления *ниже* температуры ориентационного плавления. Промежуточная фаза вещества (*мезофаза*), существующая между этими температурами, называется термотропным жидким кристаллом. Таким свойством обладают вещества, состоящие из вытянутых (продолговатых, стержнеобразных) молекул.

Впервые жидкие кристаллы были исследованы в 1882 г. австрийским ботаником Ф. Рейнитцером, который заметил, что синтезированный им холестерилбензоат, по-видимому, имеет две точки плавления. При температуре 145° С твердое тело превращается в мутную жидкость, которая при 179° С становится прозрачной. В настоящее время известны десятки тысяч органических жидкокристаллических веществ. По механическим свойствам они напоминают вязкие жидкости; однако важнейшей особенностью жидких кристаллов является анизотропия их оптических, электрических и других свойств.

Термотропные жидкие кристаллы делятся на три основные группы:

1) *нематические* (от греч. «нема» — нить), молекулы которых позиционно не упорядочены и ориентация их длинных осей в среднем одинакова по всему объему;

2) *холестерические,* имеющие винтообразную структуру. У кристаллов этой группы направления ориентации длинных осей молекул образуют в пространстве спиральную поверхность; шаг этой спирали в сотни раз превышает длину самих молекул. Большое число веществ этой группы содержит холестерин;

3) *смектические* (от греч. «смегма» — мыло), имеющие позиционную структуру слоевого характера и обладающие дальним ориентационным порядком. Эта группа жидких кристаллов наиболее разнообразна.

Обычно жидкий кристалл с изменением температуры испытывает несколько фазовых переходов, превращаясь из твердого состояния в одну из смектических фаз, затем в нематическую или холестерическую и, наконец, в изотропную жидкость. Жидкокристаллические свойства обнаружены также и в ряде коллоидных растворов (например, водный раствор вируса табачной мозаики) и некоторых полимеров. Жидкие кристаллы этого класса называются *лиотропными.* Для лиотропных кристаллов в отличие от термотропных основным параметром, изменение которого вызывает фазовые переходы, является уже не температура, а концентрация.

Жидкие кристаллы суть анизотропные вещества, в которых упорядоченно расположенные молекулы образуют не твердое тело с его «упругостью формы», а жидкое тело, легко изменяющее свою форму под действием относительно слабых сил (например, под действием гравитационных сил, которые почти не изменяют форму твердых тел). Жидкокристаллические вещества обычно помещают между двумя тонкими полимерными пленками, и полученную таким образом «слоистую пленку» используют:

1) для измерения температур (до сотых долей градуса) и изучения температурного поля на поверхности твердых тел. Пленка прикладывается к исследуемому телу. Незначительная разница температур — порядка десятых долей градуса — вызывает хорошо заметные изменения цвета жидкого кристалла. Время, необходимое для изменения окраски, — порядка секунды, время сохранения ее (инертность процесса измерения) — несколько секунд;

2) в качестве датчиков для измерения давлений. В зависимости от давления на пленку изменяется окраска жидкокристаллического вещества; это обстоятельство позволяет, например, определить поле давлений в звуковых и ультразвуковых волнах;

3) в оптических, радиотехнических и электронных устройствах; в них пленка работает как вещество, преобразующее энергию электрических полей в световую энергию. Они могут быть использованы и в устройствах по передаче и воспроизведению различной информации.

Особое значение имеет жидкокристаллическое состояние в биологии. Многие структурно упорядоченные части клеток и тканей живых организмов и растений являются жидкими кристаллами (например, мышцы в определенных пределах могут многократно растягиваться и сжиматься без остаточных деформаций и разрушения).

Часть III
ЭЛЕКТРИЧЕСТВО И МАГНЕТИЗМ

Глава 1
ЭЛЕКТРОСТАТИКА

§ 1. ЭЛЕКТРИЧЕСКИЕ ЗАРЯДЫ. ПРОВОДНИКИ И ДИЭЛЕКТРИКИ

В этом разделе курса рассматриваются физические явления, в которых участвуют электрические заряды. Существует два рода электрических зарядов, которые условно названы *положительными* и *отрицательными* зарядами. Тела, имеющие электрические заряды одного знака, отталкиваются друг от друга; тела с зарядами противоположных знаков притягиваются.

Установлено, что электричество «атомистично», т. е. как положительные, так и отрицательные заряды состоят из целого числа одинаковых по величине элементарных (наименьших, неделимых) электрических зарядов. *Элементарный электрический заряд* равен $e = 1{,}602 \cdot 10^{-19}$ Кл; такой величины отрицательный заряд имеет, например, электрон, а положительный заряд — протон. Если в атоме или молекуле вещества имеется N_1 положительных и N_2 отрицательных элементарных зарядов, то полный заряд такой частицы будет равен $e(N_1 - N_2)$. Тело (или какая-нибудь его часть) электрически заряжено, если в пределах его объема число элементарных положительных зарядов больше или меньше, чем число отрицательных зарядов; для электрически нейтральных тел эти числа равны. Атомы или молекулы, имеющие избыток электрических зарядов одного знака, называются ионами (положительными или отрицательными).

В явлениях, где участвуют электрически заряженные тела и частицы, соблюдается закон сохранения электрических зарядов: *в замкнутой системе, какие бы процессы ни протекали в ней, алгебраическая сумма положительных и отрицательных зарядов с течением времени не изменяется.* Это означает, что внутри замкнутого объема изменение суммарного электрического заряда можно осуществить только путем внесения зарядов извне или извлечения их за пределы рассматриваемого объема.

Тела, в которых заряженные частицы (электроны, ионы) могут свободно перемещаться в пределах их объема, являются проводниками; к ним относятся металлы, электролиты, ионизированный газ. В диэлектриках заряженные частицы не имеют такой свободы перемещения и могут только несколько смещаться относительно определенных положений равновесия. В проводниках заря-

женные частицы участвуют в беспорядочном тепловом движении атомов и молекул, в диэлектриках — совершают беспорядочные колебания вокруг положений равновесия.

Тела называются электрически *однородными,* если их электрические свойства одинаковы в пределах всего объема; тела называются электрически *изотропными,* если их электрические свойства одинаковы по всем направлениям.

Измерения показывают, что, например, у кристаллических тел удельное сопротивление (у проводников) или диэлектрическая проницаемость (у диэлектриков) различны в различных направлениях. Такие тела называются *анизотропными.* В той или иной степени большинство тел неоднородны и анизотропны.

Большое применение получили полупроводники — вещества, электрические свойства которых в зависимости от их состава, строения и состояния изменяются в очень широких пределах. В одних условиях (при низких температурах) они имеют большое удельное сопротивление, в других (высокие температуры) — малое. Некоторые полупроводники (селен) заметно уменьшают свое электрическое сопротивление под действием света. Свойства и применения полупроводников изложены отдельно в § 16 ч. III.

Удельные сопротивления ρ (Ом·м) различных веществ лежат в широких пределах:

у проводников — 10^{-8} — 10^{-6};

у полупроводников — 10^{-6} — 10^{-3};

у диэлектриков — 10^{3} — 10^{16}.

Заметим, что деление веществ на проводники, полупроводники и диэлектрики производится не по значению их удельных сопротивлений, а по совокупности их электрических свойств и различию в их внутренней структуре.

§ 2. ЭЛЕКТРИЧЕСКОЕ ПОЛЕ; ИНДУКЦИЯ И НАПРЯЖЕННОСТЬ ПОЛЯ. ЗАКОН КУЛОНА

Каждое заряженное тело окружено электрическим полем, которое теоретически простирается до бесконечности. Электрические поля нескольких заряженных тел образуют общее электрическое поле, в котором нельзя отличить поле одного заряда от поля других зарядов. В современной физике электрическое поле рассматривается как особая форма объективной реальности — материи, обладающей специфическими физическими свойствами. Одним из этих свойств, легко наблюдаемым и измеряемым, является следующее: на электрические заряды, помещенные в электрическом поле, действуют силы, пропорциональные величине этих зарядов.

Допустим, что электрическое поле в вакууме образовано зарядом q, сосредоточенным в некоторой точке O. Проведем из точки O сферу радиуса r и выберем на этой сфере элементарную площадку dS (рис. III.1), выделяющую телесный угол $d\omega = dS/r^2$. Так как поле покоящегося точечного заряда предполагается сферически симметричным, а телесный угол вокруг точки равен 4π, то $d\omega/4\pi$ будет пока-

зывать, какая часть электрического поля приходится на долю телесного угла $d\omega$. Допустим, что Ψ — некоторая величина, при помощи которой можно охарактеризовать *все поле* данного электрического заряда. Очевидно, Ψ должно быть пропорционально величине заряда q, создающего данное поле; для простоты можно полагать, что Ψ численно равно q. Тогда та часть поля, которая охватывается телесным углом $d\omega$, будет характеризоваться величиной

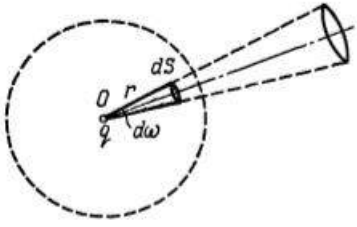


Рис. III 1

$$d\Psi = q\frac{d\omega}{4\pi} = \frac{q}{4\pi r^2}dS. \qquad (1.1)$$

В этой формуле выделим величину

$$D = \frac{q}{4\pi r^2}. \qquad (1.2)$$

Она является характеристикой электрического поля в пределах площадки dS, т. е. на расстоянии r от точечного заряда q. Условимся полагать **D** вектором, направленным по радиусу от точки O, если заряд q положительный, и к точке O, если заряд отрицательный. Это условие можно записать в следующем виде:

$$\mathbf{D} = \frac{q}{4\pi r^2}\frac{\mathbf{r}}{r}. \qquad (1.3)$$

Вектор **D** для точечного заряда определяется исходя из сферической симметрии поля. Для нахождения вектора **D** суммарного электрического поля системы зарядов пользуются принципом суперпозиции электрических полей: *в каждой точке поля любой системы зарядов вектор* **D** *равен сумме векторов* $\mathbf{D}_1$, $\mathbf{D}_2$, ... *точечных зарядов, из которых состоит данная система* (рис. III.2):

$$\mathbf{D} = \mathbf{D}_1 + \mathbf{D}_2 + \ldots \qquad (1.4)$$

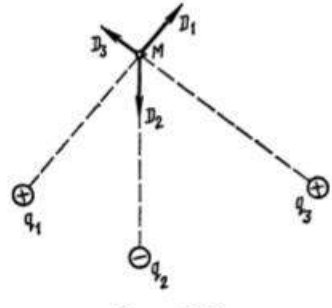


Рис. III.2

Важно отметить, что если известны заряды, создающие данное поле, то для определения вектора **D** в каждой точке поля нет необходимости производить измерения; этот вектор может быть *рассчитан* по заданному расположению электрических зарядов. Кроме того, значение вектора $\mathbf{D}_i$ для каждого отдельного заряда q_i, а также для выделенной нами системы зарядов не зависит от наличия или отсутствия среды, заполняющей поле, или других тел, содержащих заряды. В любом случае можно выделить ту или иную систему зарядов и для их электрического поля рассчитать вектор **D**. Во многих таких расчетах учитывают все заряды, включая те, которые появляются на поверхности или в объеме проводников и диэлектриков, помещенных в данное

электрическое поле. Однако очень часто выделяют поле «избыточных зарядов» от поля, создаваемого «связанными зарядами» в различных средах (диэлектриках).

Вектор $\mathbf{D}(x, y, z)$, определяемый формулами (1.3) и (1.4) и дающий геометрическую характеристику электрического поля данного заряда или данной системы зарядов, называется *вектором электрической индукции* или электрического смещения *. Соотношение (1.4), выражающее принцип суперпозиции для вектора индукции, утверждает, что *индукция общего поля системы зарядов в данной точке равна векторной сумме индукций*, которые создаются каждым зарядом в отдельности при отсутствии остальных. Скалярная величина $d\Psi$ называется *потоком электрической индукции* (смещения) через площадку dS и будет использована в § 3 при расчете электрических полей.

Для опытного изучения электрического поля можно воспользоваться «пробным» (точечным и для определенности положительным) зарядом q_0, который должен быть настолько мал, чтобы его собственное электрическое поле не искажало изучаемого поля (не вызывало бы перемещения зарядов на телах, с которыми связано это поле). Помещая заряд q_0 в ту или иную точку поля, измеряют действующую на него силу F_0. Отношение F_0/q_0, обозначаемое E, не зависит от величины q_0 и называется *напряженностью электрического поля* в данной точке:

$$E = \frac{F_0}{q_0}. \tag{1.5}$$

Для вычисления E необходимо предварительно измерять величину пробного заряда q_0. Измерение электрических зарядов производится на основе з а к о н а К у л о н а: *два точечных заряда q_1 и q_2, находящиеся в однородной и изотропной среде* (границы которой находятся очень далеко от этих зарядов) *на расстоянии r друг от друга, взаимодействуют с силой*

$$F = k\frac{q_1 q_2}{\varepsilon r^2}, \tag{1.6}$$

где k — коэффициент, зависящий от выбора единиц измерений для силы, заряда и расстояния, а ε — безразмерная величина, характеризующая электрические свойства диэлектрика и показывающая, во сколько раз сила взаимодействия этих зарядов в данном диэлектрике меньше, чем в вакууме (для вакуума полагают $\varepsilon = 1$). В векторной записи закон Кулона имеет вид

$$\mathbf{F} = k\frac{q_1 q_2}{\varepsilon r^2}\frac{\mathbf{r}}{r}. \tag{1.7}$$

В *абсолютной электростатической системе единиц* (СГС), в которой сила выражается в динах, расстояние — в сантиметрах, а коэффициент k полагается безразмерной величиной, равной единице, получим

$$[q] = \text{см}^{3/2} \cdot \text{г}^{1/2}\text{с}^{-1}.$$

---

* ГОСТ разрешает вектор **D** называть как вектором электрической индукции, так и вектором электрического смещения. В настоящем курсе предпочтение отдано первому термину.

В Международной системе (СИ) сила выражается в ньютонах, расстояние — в метрах, а заряд — в кулонах; тогда коэффициент k в формуле (1.6) оказывается равным

$$k = 9 \cdot 10^9 \frac{\text{ньютон} \cdot \text{метр}^2}{\text{кулон}^2} \left(\text{или } \frac{\text{Н} \cdot \text{м}^2}{\text{Кл}^2}\right).$$

Постоянную величину k принято заменять на

$$k = \frac{1}{4\pi\varepsilon_0}.$$

Тогда закон Кулона (1.7) записывается в виде

$$F = \frac{q_1 q_2}{4\pi\varepsilon_0 \varepsilon r^2} \frac{\mathbf{r}}{r}, \tag{1.8}$$

где

$$\varepsilon_0 = \frac{1}{4\pi k} = \frac{1}{36\pi \cdot 10^9} \frac{\text{Кл}^2}{\text{Н} \cdot \text{м}^2}. \tag{1.9}$$

Найдем формулу для расчета напряженности электрического поля в различных точках вокруг точечного заряда q. Для этого предположим, что в формуле (1.6) один из взаимодействующих зарядов пробный, т. е. $q_1 = q$; $q_2 = q_0$. Тогда, согласно определению напряженности поля (1.5), получаем

$$E = \frac{F}{q_0} = k \frac{q}{\varepsilon r^2} = \frac{q}{4\pi\varepsilon_0 \varepsilon r^2}, \tag{1.10}$$

или в векторном виде

$$\mathbf{E} = k \frac{q}{\varepsilon r^2} \frac{\mathbf{r}}{r} = \frac{q}{4\pi\varepsilon_0 \varepsilon r^2} \frac{\mathbf{r}}{r}. \tag{1.11}$$

При определении напряженности поля в вакууме следует положить $\varepsilon = 1$.

Таким образом, каждая точка электрического поля характеризуется двумя векторными величинами: **D** и **E**. Найдем связь между ними; согласно формулам (1.2) и (1.10), имеем

$$E = \frac{4\pi k}{\varepsilon} D. \tag{1.12}$$

Тогда (в векторной записи)

$$\mathbf{E} = \frac{1}{\varepsilon_0 \varepsilon} \mathbf{D}; \quad \mathbf{D} = \varepsilon_0 \varepsilon \mathbf{E}. \tag{1.13}$$

Величина ε_0 называется *электрической постоянной* (иногда диэлектрической проницаемостью вакуума). Произведение $\varepsilon_0\varepsilon$ называется *абсолютной диэлектрической проницаемостью данной* среды, а безразмерная величина ε, приводимая в справочных таблицах, называется *относительной диэлектрической проницаемостью* данной среды (диэлектрика) по отношению к вакууму.

Формулы (1.13) соответствуют Международной системе единиц (СИ). Величины D и E имеют различные размерности: индукция (смещение) поля D выражается в кулонах на квадратный метр (Кл/м^2),

а напряженность поля E — в ньютонах на кулон (Н/Кл) или (так как 1 Дж = 1 В·Кл) — в вольтах на метр (В/м).

В абсолютной электростатической системе СГС электрическая индукция (или смещение) D в поле точечного заряда определяется не по формулам (1.2) или (1.3), а без коэффициента 4π:

$$D = \frac{q}{r^2}; \quad \mathbf{D} = \frac{q}{r^2}\frac{\mathbf{r}}{r}. \tag{1.14}$$

Напряженность поля, согласно формулам (1.10) и (1.11), определяется по формуле, не содержащей коэффициента k:

$$E = \frac{q}{\varepsilon r^2}; \quad \mathbf{E} = \frac{q}{\varepsilon r^2}\frac{\mathbf{r}}{r}. \tag{1.15}$$

Поэтому связь между векторами **E** и **D** не содержит постоянной ε_0:

$$\mathbf{E} = \frac{1}{\varepsilon}\mathbf{D}. \tag{1.16}$$

Так как ε есть безразмерная величина, то в системе СГС размерности **E** и **D** совпадают ($г^{1/2} \cdot см^{-1/2} \cdot с^{-1}$).

Выясним, чем обусловлена необходимость введения двух векторов **D** и **E**, характеризующих каждую точку электрического поля. Различие между ними заключается в том, что индукцию **D** в каждой точке поля можно *вычислить* по известному расположению электрических зарядов на основании определения (1.2) и принципа суперпозиции (1.4). Напряженность же **E**, по определению (1.5), может быть найдена только *путем измерения* с применением «пробного» заряда; найти этот вектор на основании формулы $\mathbf{D} = \varepsilon\varepsilon_0\mathbf{E}$ невозможно, если диэлектрическая проницаемость среды ε неизвестна. Заметим, что для многих веществ ε не есть постоянная величина и зависит от интенсивности внешнего электрического поля, действующего на среду, а также от быстроты его изменения (например, от частоты колебаний); кроме того, у неоднородных и анизотропных сред диэлектрическая проницаемость различна в различных местах и по различным направлениям, взятым в этой среде. Таким образом, в соотношении $\mathbf{D} = \varepsilon\varepsilon_0\mathbf{E}$ из трех величин одна(**D**) может быть вычислена по расположению электрических зарядов, создающих данное поле, вторая (**E**) измерена, а третья ($\varepsilon = D/\varepsilon_0 E$) найдена как результат экспериментального изучения данной среды. Лишь в весьма частных случаях, когда электрическое поле существует в вакууме или в безграничной однородной и изотропной среде, возможен расчет **E** по значениям **D**.

В некоторых случаях диэлектрик заполняет только часть электрического поля взаимодействующих зарядов. На рис. III.3 показаны три таких примера. В § 6 будет показано, что на поверхности диэлектрика в таких случаях появляются некомпенсированные заряды, которые создают добавочное электрическое поле; это поле накладывается на поле рассматриваемых зарядов q_1 и q_2 и должно быть учтено при расчете **D** в каждой интересующей нас точке.

Допустим, что в вакууме заряды q_1, q_2, ... действуют на пробный заряд q_0 сначала каждый в отдельности (при отсутствии других)

с силами $\mathbf{F}_1$, $\mathbf{F}_2$, ..., а затем одновременно (при том же их взаимном расположении) с силой $\mathbf{F}$. Измерения показывают, что в вакууме соблюдается равенство

$$\mathbf{F} = \mathbf{F}_1 + \mathbf{F}_2 + \ldots \qquad (1.17)$$

Подставим в эту формулу выражение для каждой из сил; согласно равенствам (1.5) и (1.13), $\mathbf{F} = q_0\mathbf{E} = \frac{q_0}{\varepsilon_0}\mathbf{D}$ и поэтому

$$\mathbf{F}_1 = \frac{q_0}{\varepsilon_0}\mathbf{D}_1; \quad \mathbf{F}_2 = \frac{q_0}{\varepsilon_0}\mathbf{D}_2; \ \ldots; \ \mathbf{F} = \frac{q_0}{\varepsilon_0}\mathbf{D};$$

после подстановки и сокращения получим формулу (1.4). Таким образом, из опытного результата (1.17) получается принцип суперпозиции для вектора $\mathbf{D}$ и, наоборот, из принципа суперпозиции (1.4) можно вывести соотношение (1.17).

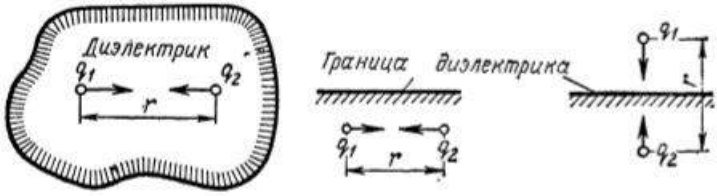


Рис. III.3

Однако если такой расчет произвести не для вакуума, а для различных диэлектрических сред, то силы $\mathbf{F}$ будут выражены в зависимости от относительной диэлектрической проницаемости ε: $\mathbf{F} = q_0\mathbf{E} = (q_0/\varepsilon_0\varepsilon)\,\mathbf{D}$. Величина ε, характеризующая электрическое состояние среды, может изменяться в зависимости от интенсивности поля, действующего внутри среды. Поэтому если в среде находится только один заряд q_1, то для расчета силы $\mathbf{F}_1$ следует полагать $\varepsilon = \varepsilon_1$; для заряда q_2 и силы $\mathbf{F}_2$ может оказаться, что $\varepsilon = \varepsilon_2$ и т. д.; для расчета силы $\mathbf{F}$ при одновременном существовании всех зарядов диэлектрическую проницаемость примем равной ε. Тогда

$$\mathbf{F}_1 = \frac{q_0}{\varepsilon_0\varepsilon_1}\mathbf{D}_1; \ \mathbf{F}_2 = \frac{q_0}{\varepsilon_0\varepsilon_2}\mathbf{D}_2; \ \ldots; \ \mathbf{F} = \frac{q_0}{\varepsilon_0\varepsilon}\mathbf{D}.$$

Воспользуемся теперь принципом суперпозиции (1.4) и выразим векторы $\mathbf{D}$ через векторы сил:

$$\frac{\varepsilon_0\varepsilon}{q_0}\mathbf{F} = \frac{\varepsilon_0\varepsilon_1}{q_0}\mathbf{F}_1 + \frac{\varepsilon_0\varepsilon_2}{q_0}\mathbf{F}_2 + \ldots$$

Очевидно, что соотношение (1.17) может выполняться только в том случае, если относительная диэлектрическая проницаемость среды — величина постоянная, т. е. $\varepsilon_1 = \varepsilon_2 = \ldots = \varepsilon$. При этих условиях будет соблюдаться и равенство

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 + \ldots, \qquad (1.18)$$

где $\mathbf{E}_1$, $\mathbf{E}_2$, ... — напряженности полей, которые создаются в данной точке среды каждым зарядом в отдельности при отсутствии других. Среда, в которой ε не зависит от $\mathbf{E}$, называется линейной.

Заметим, что направления векторов $\mathbf{E}$ и $\mathbf{D}$ совпадают только в изотропной среде. Допустим, что среда неизотропная, причем относительная диэлектрическая проницаемость в направлении OX равна ε_x, а в направлении OY равна ε_y. Разложим вектор $\mathbf{D}$ в точке O (рис. III.4) на составляющие по координатным осям $\mathbf{D}_x$ и $\mathbf{D}_y$. Тогда составляющие вектора напряженности $\mathbf{E}_x$ и $\mathbf{E}_y$ будут равны:

$$\mathbf{E}_x = \frac{1}{\varepsilon_0 \varepsilon_x} \mathbf{D}_x; \quad \mathbf{E}_y = \frac{1}{\varepsilon_0 \varepsilon_y} \mathbf{D}_y,$$

следовательно, $\operatorname{tg} \beta = E_y/E_x \neq \operatorname{tg} \alpha = D_y/D_x$. В общем случае следует учесть значение диэлектрической проницаемости ε_z и в направлении третьей оси (OZ).

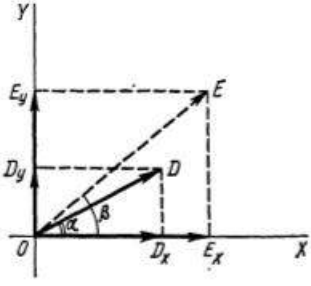


Рис. III.4

Ввиду этого при расчете силы $\mathbf{F}$, действующей на заряд в электрическом поле, удобнее пользоваться не вектором индукции

$$\mathbf{F} = \frac{q}{\varepsilon_0 \varepsilon} \mathbf{D},$$

а вектором напряженности

$$\mathbf{F} = q\mathbf{E}, \tag{1.19}$$

так как во всех случаях, согласно определению (1.5), сила F по направлению либо совпадает с направлением E, если заряд q положительный, либо противоположна ему, если заряд q отрицательный. В анизотропной среде компоненты силы F_x, F_y, F_z пропорциональны компонентам вектора напряженности: $F_x = qE_x$ и т. д.

Электрическое поле называется *однородным,* если во всех его точках напряженность одинакова и по величине и по направлению, т. е. $\mathbf{E} = \text{const}$. В большинстве случаев электрические поля различных заряженных тел неоднородные, т. е. $E = E(x, y, z)$.

Очевидно, что достаточно малые участки неоднородного поля можно с некоторым приближением считать однородными. Поэтому для изучения электрических полей должны быть взяты «пробные заряды» малых размеров, чтобы в пределах их объема поля можно было считать однородными.

Силы электрического взаимодействия заряженных элементарных частиц (электронов, протонов и т. д.) значительно превосходят силы гравитационного взаимодействия между ними. Так, для двух протонов, массы и электрические заряды которых равны соответственно $m = 1{,}67 \cdot 10^{-27}$ кг; $e = 1{,}6 \cdot 10^{-19}$ Кл, получим

$$\frac{F_{эл}}{F_{грав}} = \frac{ke^2}{Gm^2} = \frac{9 \cdot 10^9 (1{,}6 \cdot 10^{-19})^2}{6{,}67 \cdot 10^{-11} (1{,}67 \cdot 10^{-27})^2} \approx 1{,}23 \cdot 10^{36}.$$

Однако в том случае, когда между зарядами и массами тел нет той определенной связи, которая имеется у протонов, электронов и других заряженных элементарных частиц, это отношение может быть любым. Например, электрическое взаимодействие между такими космическими телами, как Солнце и планеты, значительно слабее их гравитационного взаимодействия.

§ 3. ЭЛЕКТРИЧЕСКОЕ ПОЛЕ СИСТЕМЫ ЗАРЯДОВ; ТЕОРЕМА ОСТРОГРАДСКОГО—ГАУССА

Простейшей системой зарядов является *электрический диполь* — два равных по величине разноименных точечных заряда, расположенных на расстоянии l друг от друга. В каждой точке окружающего диполь пространства суммарный вектор индукции $\mathbf{D}$ определится путем векторного сложения индукций $\mathbf{D}_{(+)}$ и $\mathbf{D}_{(-)}$, создаваемых каждым зарядом в отдельности:

$$\mathbf{D} = \mathbf{D}_{(+)} + \mathbf{D}_{(-)}.$$

Величину и направление вектора $\mathbf{D}$ в точке A с координатами x и y найдем по значениям проекций векторов $\mathbf{D}_{(+)}$ и $\mathbf{D}_{(-)}$ на координатные оси (рис. III.5, *а*):

$$\left.\begin{aligned} D_x &= D_{(+)}\cos\alpha + D_{(-)}\cos\beta = \frac{q}{4\pi\varepsilon_0 r_1^2}\frac{l/2+x}{r_1} + \frac{q}{4\pi\varepsilon_0 r_2^2}\frac{l/2-x}{r_2}; \\ D_y &= D_{(+)}\sin\alpha - D_{(-)}\sin\beta = \frac{q}{4\pi\varepsilon_0 r_1^2}\frac{y}{r_1} - \frac{q}{4\pi\varepsilon_0 r_2^2}\frac{y}{r_2}. \end{aligned}\right\} \quad (1.20)$$

Тогда $D = \sqrt{D_x^2 + D_y^2}$; угол γ между вектором $\mathbf{D}$ и положительным направлением оси OX определится из условия $\operatorname{tg}\gamma = D_x/D_y$. Расчет

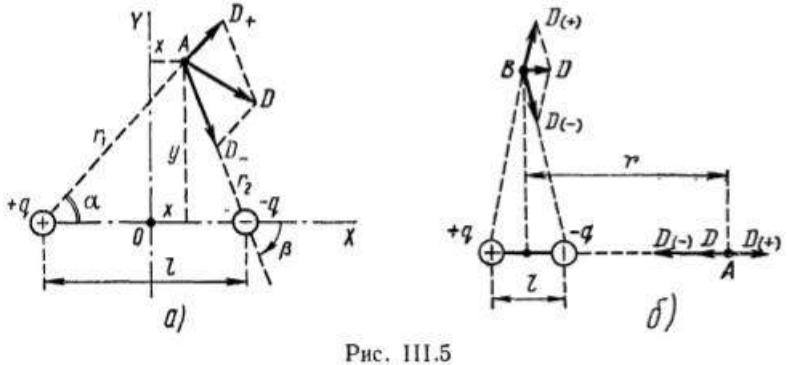


Рис. III.5

упрощается, если точка, в которой определяется D_1, расположена либо на линии, соединяющей заряды (точка A на оси OX, рис. III.5, *б*), либо на перпендикуляре, проведенном через центр O диполя (точка B на оси $OУ$, рис. III.5, *б*). В первом случае $y = 0$, $D_y = 0$, $r_1 = x + l/2$, $r_2 = x - l/2$. Если расстояние x от точки A до центра ди-

поля, которое обозначим через r, значительно больше l, то

$$D \approx \frac{2ql}{4\pi\varepsilon_0 r^3}.$$

Во втором случае для точки B имеем: $r_1 = r_2 = \sqrt{r^2 + l^2/4}$ (где $r = OB$), $D_x = 0$ и тогда при $l \ll r$

$$D = \frac{ql}{4\pi\varepsilon_0 \sqrt{(r^2 + l^2/4)^3}} \approx \frac{ql}{4\pi\varepsilon_0 r^3}.$$

Таким образом, электрическое поле диполя на больших расстояниях ($r \gg l$) убывает обратно пропорционально кубу этого расстояния.

Произведение ql называется *электрическим моментом* диполя и выражается в Кл·м. Расстояние l (называемое *плечом диполя*) рассматривается как вектор, направление которого ориентировано от отрицательного заряда к положительному. Таким образом, электрический момент диполя есть вектор

$$\mathbf{p} = q\mathbf{l}.$$

Выше было показано, что величина и направление вектора индукции в какой-нибудь точке электрического поля вокруг диполя зависят от расположения этой точки относительно электрического момента диполя. Допустим, что в некотором малом объеме ΔV имеется множество электрических диполей, различным образом ориентированных. В некоторой точке, расположенной достаточно далеко от этих диполей, суммарный вектор индукции $\mathbf{D}$ будет определяться векторной суммой электрических моментов всех диполей в объеме ΔV:

$$\mathbf{P} = \sum \mathbf{p}_i = \sum q_i \mathbf{l}_i,$$

которая при хаотическом расположении диполей может оказаться очень малой величиной. Можно показать, что у систем из равного числа симметрично расположенных положительных и отрицательных зарядов индукция поля убывает на больших расстояниях от этой системы обратно пропорционально более высоким степеням расстояния. Отсутствие заметного электрического поля вокруг нейтральных тел, содержащих большое число близко расположенных друг к другу положительных и отрицательных зарядов, также объясняется суперпозицией (векторным сложением) индукций этих зарядов.

Для расчета электрического поля зарядов, распределенных вдоль линии, поверхности или по объему, мысленно разбивают их на элементарные почти точечные заряды и затем производят векторное суммирование индукций от этих зарядов. Рассчитаем индукцию поля в точке A, расположенной на расстоянии a от равномерно заряженной нити (стержня, проволоки). Обозначим через τ заряд, приходящийся на единицу длины нити. От элементарного участка проводника длиной dl, на котором имеется «почти точечный» заряд $dq = \tau dl$, индукция в точке A (рис. III.6) будет равна

$$dD = \frac{dq}{4\pi r^2} = \frac{\tau\, dl}{4\pi r^2} = \frac{\tau}{4\pi a} d\alpha \qquad (1.21)$$

$$\left(\cos\alpha = \frac{a}{r};\quad d\alpha = \frac{dl\cos\alpha}{r}\right).$$

Векторное сложение индукций dD заменим скалярным сложением составляющих по координатным осям: $dD_x = dD\cos\alpha$; $dD_y = dD\sin\alpha$. Сначала рассчитаем D_x и D_y для участка нити $OB = l_1$:

$$D_x = \int_0^{\alpha_1} dD\cos\alpha = \frac{\tau}{4\pi a}\sin\alpha_1;$$

$$D_y = \int_0^{\alpha_1} dD\sin\alpha = \frac{\tau}{4\pi a}(1-\cos\alpha_1). \qquad (1.22)$$

Для участка нити $OC = l_2$ получаются те же формулы с заменой угла α_1 на α_2. Для расчета индукции, создаваемой в точке A всей нитью

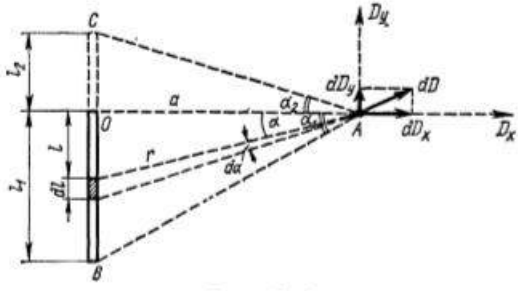


Рис. III.6

BC, необходимо складывать составляющие по оси OX и вычитать составляющие по оси OY, создаваемые участками OB и OC; получим:

$$D_x = \frac{\tau}{4\pi a}(\sin\alpha_1 + \sin\alpha_2);$$

$$D_y = \frac{\tau}{4\pi a}(\cos\alpha_2 - \cos\alpha_1). \qquad (1.23)$$

Если нить очень длинная: $l_1 \gg a$; $l_2 \gg a$, то $\alpha_1 \approx \alpha_2 \approx \pi/2$, $D_y \approx 0$, а $D_x \approx D$, следовательно,

$$D \approx \frac{\tau}{2\pi a}. \qquad (1.24)$$

При этих условиях в каждой точке электрического поля индукция D перпендикулярна нити и по величине обратно пропорциональна расстоянию до нее.

Этими результатами можно воспользоваться для расчета индукции поля, создаваемого равномерно заряженной плоскостью; для этой цели разобьем плоскость на множество прямолинейных полос бесконечно малой толщины dl (рис. III.7). Обозначим через σ заряд, находящийся на единице площади; тогда на единицу длины каждой полосы будет приходиться заряд $\tau = \sigma \cdot 1 \cdot dl = \sigma dl$. Индукцию поля в точке A, создаваемую одной полосой, можно рассчитать по формулам (1.23) и затем произвести суммирование индукций от всех полос.

Рассмотрим частный пример, когда расстояние b от точки A до плоскости очень мало по сравнению с расстояниями от этой точки до границ плоскости (т. е. когда плоскость «безгранично велика»). Тогда для каждой полосы может быть использована формула (1.24); индукция, создаваемая одной полосой в точке A, будет равна $dD = \sigma dl/2\pi a$. Ориентируем ось OX вдоль перпендикуляра, опущенного из точки A

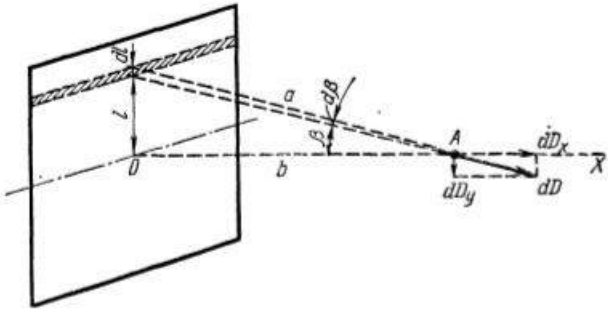


Рис. III.7

на плоскость; разложим dD на составляющие dD_x и dD_y и проинтегрируем для всех значений угла β от $+\pi/2$ до $-\pi/2$. Так как $\cos\beta = b/a$, $d\beta = (dl\cos\beta)/a$, то

$$dD_x = dD\cos\beta = \frac{\sigma}{2\pi}d\beta; \quad dD_y = dD\sin\beta = \frac{\sigma}{2\pi}\operatorname{tg}\beta\, d\beta.$$

Для составляющих по оси OY (т. е. параллельных плоскости) расчет дает нуль. Тогда $D = D_x$, следовательно,

$$D = \int_{-\pi/2}^{+\pi/2} \frac{\sigma}{2\pi}d\beta = \frac{\sigma}{2}, \qquad (1.25)$$

т. е. в каждой точке электрического поля вблизи равномерно заряженной плоскости индукция D перпендикулярна плоскости, не зависит от расстояния до нее и равна $\sigma/2$.

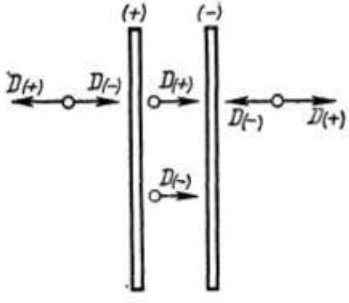


Рис. III.8

Две параллельные разноименно заряженные плоскости (плоский конденсатор) создают общее поле, индукция которого D в каждой точке пространства между ними будет равна сумме $D_{(+)}$ и $D_{(-)}$, создаваемых каждой плоскостью, а за пределами — их разности (рис. III.8). Если поверхностная плотность зарядов σ на плоскостях одинаковая, то в значительной части пространства между плоскостями (т. е. далеко от их краев) можно, согласно формуле (1.25), положить $D_{(+)} = D_{(-)} = \sigma/2$; следовательно, индукция поля

$$D = \sigma \qquad (1.26)$$

и направлена от положительно заряженной плоскости к отрицательной (по нормали к ним). Вблизи границ пластинок поле будет слабее; в остальном пространстве вокруг плоскостей поле практически отсутствует.

Расчет электрических полей распределенных зарядов в некоторых случаях заметно облегчается применением теоремы Остроградского — Гаусса. Допустим, что имеется система точечных зарядов q_1, q_2, ..., создающих электрическое поле; в каждой точке поля индукция D может быть рассчитана по формулам (1.3) и (1.4). Охватим эти заряды какой-нибудь замкнутой поверхностью S, которую разделим на маленькие, почти плоские, участки ΔS; единичные нормали $\mathbf{n}$ к этим участкам условимся ориентировать наружу (рис. III.9).

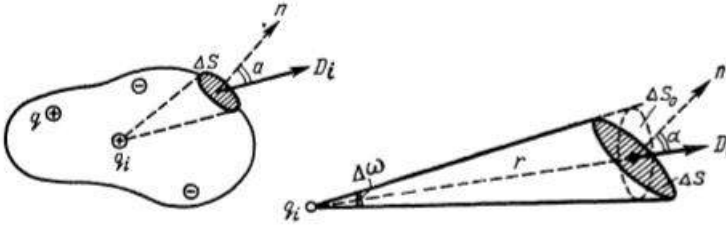


Рис. III.9

Допустим, что один из зарядов q_i создает в пределах площадки ΔS поле с вектором индукции $\mathbf{D}_i$. Произведение $\Delta\Psi_i = \mathbf{D}_i\mathbf{n}\Delta S = D_i\Delta S\cos\alpha$ — поток электрической индукции D_i через площадку ΔS. Произведение $\Delta S\cos\alpha = \Delta S_0$ — проекция площадки ΔS на сферическую поверхность радиуса r, поэтому отношение $\Delta S_0/r^2 = \Delta S\cos\alpha/r^2$ равно телесному углу $\Delta\omega$.

Воспользуемся формулой (1.3) для индукции; тогда

$$\Delta\Psi_i = D_i\,\Delta S\cos\alpha = \frac{q_i}{4\pi r_i^2}\,\Delta S_0 = \frac{q_i}{4\pi}\,\Delta\omega.$$

Найдем для заряда q_i полный поток вектора индукции создаваемого им электрического поля через всю замкнутую поверхность S; так как телесный угол вокруг точки равен 4π, то

$$\Psi_i = \sum\frac{q_i}{4\pi}\,\Delta\omega = \frac{q_i}{4\pi}\sum\Delta\omega = q_i.$$

Этот результат не зависит от формы и размеров выбранной нами замкнутой поверхности S и от местонахождения заряда q_i внутри охватываемого объема. Если заряд q_i положительный, то вектор индукции D_i во всех точках поверхности S направлен наружу, поэтому углы α будут острые и Ψ_i является положительной величиной; у отрицательных зарядов D_i направлены внутрь поверхности, углы будут тупые и Ψ_i получится отрицательным. Таким образом, знаки Ψ_i и q_i совпадают.

При наличии внутри замкнутой поверхности нескольких зарядов вектор индукции $\mathbf{D}$ поля в пределах ΔS равен сумме векторов индук-

ций $\mathbf{D}_i$ от каждого заряда: $\mathbf{D} = \mathbf{D}_1 + \mathbf{D}_2 + \ldots$ Следовательно, поток индукции через ΔS и через всю поверхность S будет состоять из суммы потоков индукции, создаваемых отдельными зарядами: $\Psi = \Sigma\Psi_i = \Sigma q_i$. Обозначая точную сумму потоков индукции через замкнутую поверхность интегралом, получим

$$\Psi = \oint D \cos\alpha \, dS = \sum q_i. \qquad (1.27)$$

Это соотношение выражает теорему Остроградского — Гаусса:

полный поток электрической индукции поля системы зарядов через охватывающую их замкнутую поверхность любой формы и размеров равен алгебраической сумме зарядов, находящихся внутри этой поверхности.

Заряды, расположенные за пределами выбранной нами замкнутой поверхности, в эту сумму не входят, так как для обеих частей поверхности (ABC и $AB'C$, рис. III.10) потоки индукции будут иметь равные и противоположные по знаку значения и их сумма будет равна нулю.

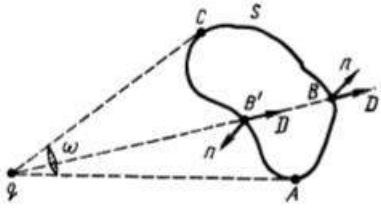


Рис. III.10

В самом общем случае электрические заряды могут быть «размазаны» по объему с некоторой плотностью $\rho = dq/dV$, причем эта плотность может быть различной в различных местах, т. е. $\rho = \rho(x, y, z)$. Тогда, выбрав некоторый объем V произвольной формы и размеров, можно записать теорему Остроградского — Гаусса в виде

$$\oint D_n \, dS = \int_V \rho \, dV, \qquad (1.28)$$

где S — замкнутая поверхность, охватывающая объем V, а D_n — проекция вектора индукции в данной точке этой поверхности на направление нормали к площадке dS.

ПРИМЕНЕНИЕ ТЕОРЕМЫ ОСТРОГРАДСКОГО—ГАУССА

Рассмотрим применения этой теоремы для расчета электрических полей вокруг и внутри некоторых заряженных тел.

1. Поле шара, равномерно заряженного по объему. Обозначим через ρ избыточный электрический заряд, содержащийся в единице объема тела (иногда называемый «объемной плотностью заряда»). Для расчета вектора $\mathbf{D}$ внутри этого шара в качестве замкнутой поверхности возьмем сферу радиуса r (рис. III.11, *а*). Из симметрии поля следует, что вектор $\mathbf{D}$ вдоль этой поверхности везде одинаков по величине и направлен по радиусу ($\cos\alpha = 1$). Тогда с одной стороны, по определению, $\Psi = \Sigma D \Delta S \cos\alpha = D\Sigma\Delta S = D \cdot 4\pi r^2$, а с другой стороны, по теореме Остроградского — Гаусса,

$$\Psi = \sum q_i = \frac{4}{3}\pi r^3 \rho.$$

Приравнивая эти выражения, получаем

$$D = \frac{1}{3}\rho r \quad (\text{для } r < R),$$

т. е. внутри равномерно заряженного (по объему) шара индукция электрического поля равна нулю в центре и возрастает пропорционально расстоянию от центра шара. Для сферы, взятой за пределами шара, по определению, $\Psi = D \cdot 4\pi r^2$, а по теореме Остроградского—Гаусса, $\Psi = \Sigma q_i = q$, где q — заряд всего шара. Из этих соотношений получаем

$$D = \frac{q}{4\pi r^2} \quad (\text{для } r > R), \tag{1.29}$$

т. е. за пределами объема шара индукция электрического поля убывает обратно пропорционально квадрату расстояний от центра шара. Таким образом, D имеет наибольшее значение на поверхности шара.

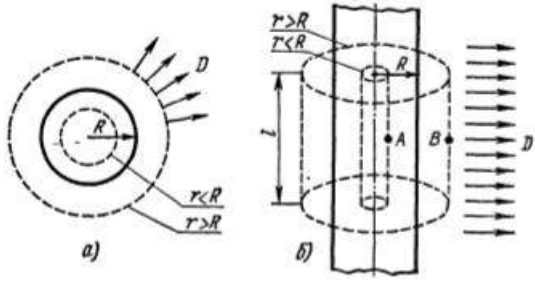


Рис. III.11

2. Поле шара, равномерно заряженного по поверхности (внутри шара избыточных зарядов нет). В этом случае для любой сферы, взятой внутри шара, $\Sigma q_i = 0$, следовательно, $\Psi = 0$ и $D = 0$, т. е. поле внутри шара отсутствует. Это означает, что в любой точке внутри шара сумма векторов индукции от всех точечных зарядов, лежащих на поверхности, равна нулю (если заряд распределен по поверхности шара неравномерно, то $\Psi = \Sigma D \Delta S \cos\alpha = 0$, но $D \neq 0$; выражение для Ψ содержит слагаемые противоположных знаков).

На самой поверхности равномерно заряженного шара, где $r = R$ (R — радиус шара), согласно формуле (1.29), имеем

$$D = \frac{q}{4\pi R^2} = \sigma,$$

где σ — отношение полного заряда шара к его поверхности — плотность зарядов на этой поверхности. Для точек, расположенных за пределами шара, применение теоремы Остроградского—Гаусса приводит к формуле (1.29).

3. Поле бесконечно длинного равномерно заряженного (по объему или по поверхности) цилиндра. В этом случае вектор D будет везде перпендикулярен оси цилиндра. Для расчета поля в какой-нибудь точке A внутри цилиндра или в B вне цилиндра проводим через эти точки цилиндрические поверхности радиуса r (на рис. III.11, *б* они показаны пунктиром). Потоки вектора D через основания этих цилиндров будут равны нулю ($\cos\alpha = 0$), а через боковые поверхности $\Psi = 2\pi r l D$. Если заряды расположены по поверхности, то для внутренних точек A будем иметь $\Psi = \Sigma q_i = 0$, поэтому $D = 0$. Для внешних точек B

$$\Psi = 2\pi r l D = \sum q_i = \tau l; \quad D = \frac{\tau}{2\pi r}, \tag{1.30}$$

где τ — заряд, приходящийся на единицу длины цилиндра. Если же заряды расположены по объему с плотностью ρ, то для точек B получится та же формула (1.30),

а для внутренних точек $\Psi = 2\pi rlD$; $\Sigma q_i = \pi r^2 l\rho$ и

$$D = \frac{1}{2}\rho r. \tag{1.31}$$

Заметим, что индукция D на поверхности равномерно заряженного цилиндра равна поверхностной плотности σ. Обозначив радиус цилиндра через R и рассмотрев участок длиной l, имеем

$$\sigma = \frac{\tau l}{2\pi Rl} = \frac{\tau}{2\pi R},$$

что совпадает с формулой (1.30) при $r = R$. Таким образом, на поверхности цилиндра

$$D = \sigma.$$

Итак, индукция электрического поля вдоль оси бесконечно длинного равномерно заряженного цилиндра равна нулю, а на его поверхности численно равна заряду, приходящемуся на единицу площади ($D = \sigma$). В том случае, если заряды расположены только на поверхности цилиндра, поле внутри цилиндра отсутствует; если же цилиндр равномерно заряжен по объему, то индукция поля, согласно формуле (1.31), возрастает от нуля до $D = \sigma$ пропорционально удалению от оси цилиндра. В обоих случаях поле за пределами цилиндра убывает (от значения $D = \sigma$ на поверхности) обратно пропорционально расстоянию от оси цилиндра согласно формуле (1.30).

4. Поле безграничной равномерно заряженной плоскости; теорема Остроградского—Гаусса приводит к результату (см. формулу (1.25))

$$D = \sigma/2.$$

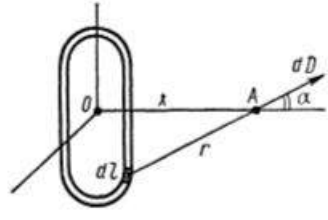


Рис. III.12

5. Поле равномерно заряженного тонкого кольца (вдоль его оси). Допустим, что τ — заряд на единице длины кольца; тогда индукция dD в точке A от заряда $dq = \tau dl$, находящегося на элементе кольца dl (рис. III.12), будет равна

$$dD = \frac{dq}{4\pi r^2} = \frac{\tau dl}{4\pi r^2}.$$

Ввиду равномерности распределения заряда по кольцу сумма проекций векторов $d\mathbf{D}$ на плоскость, перпендикулярную оси кольца, будет равна нулю. Найдем сумму проекций $d\mathbf{D}$ на ось кольца:

$$D_x = \int_L dD_x = \int_L \frac{\tau dl}{4\pi r^2}\cos\alpha = \frac{\tau}{4\pi r^2}2\pi R\cos\alpha = \frac{q}{4\pi r^2}\cos\alpha,$$

где $q = \tau \cdot 2\pi R$ — полный заряд кольца. Так как $r = x/\cos\alpha$, $\cos\alpha = x/\sqrt{R^2+x^2}$, то окончательно ($D_x = D$)

$$D = \frac{q}{4\pi x^2}\cos^3\alpha, \quad \text{или} \quad D = \frac{qx}{4\pi(R^2+x^2)^{3/2}},$$

т. е. индукция поля в центре кольца ($x = 0$) равна нулю, а вдоль оси увеличивается в сложной зависимости от расстояния до центра.

6. Поле равномерно заряженного диска (вдоль его оси). Задача может быть решена разделением диска на элементарные кольца толщиной dR, имеющие средний радиус R, площадь $2\pi R dR$ и заряд $dq = \sigma \cdot 2\pi R dR$, где σ — поверхностная плотность зарядов на диске. Очевидно, что в каждой точке A, взятой на оси диска, вследствие симметричного расположения зарядов индукция D будет направлена вдоль оси (для всех других точек вектор индукции будет лежать в одной

плоскости с осью диска). Индукция dD в точке A (x) от каждого элементарного кольца будет равна

$$dD = \frac{dq}{4\pi x^2}\cos^3\alpha = \frac{\sigma \cdot 2\pi R\,dR}{4\pi x^2}\cos^3\alpha.$$

Так как $R/x = \operatorname{tg}\alpha$, $d\alpha/\cos^2\alpha = dR/x$, то

$$dD = \frac{\sigma}{2}\sin\alpha\,d\alpha; \quad D = \int_0^{\alpha_0} dD = \frac{\sigma}{2}(1-\cos\alpha_0),$$

где $\cos\alpha_0 = \dfrac{x}{R_0} = \dfrac{x}{\sqrt{R_0^2+x^2}}$, R_0 — радиус диска. Окончательно

$$D = \frac{\sigma}{2}\left(1 - \frac{x}{\sqrt{R_0^2+x^2}}\right).$$

Если $x \ll R_0$, то $D \approx \sigma/2$, т. е. вблизи поверхности диска поле такое же, как у бесконечной плоскости. Если $x \gg R_0$, то

$$\sqrt{\left(\frac{R_0}{x}\right)^2+1} \approx 1 + \frac{1}{2}\left(\frac{R_0}{x}\right)^2; \quad D \approx \frac{\sigma}{2}\cdot\frac{R_0^2}{2x^2} = \frac{\sigma\pi R_0^2}{4\pi x^2} = \frac{q}{4\pi x^2},$$

где $q = \sigma\pi R_0^2$ — полный заряд диска; в этом случае поле диска будет мало отличаться от поля точечного заряда или шара.

7. Поле между обкладками шарового (сферического) конденсатора. Обозначим радиус наружного шара через r_1, внутреннего — через r_2. Заряды $q > 0$ и $q < 0$ будут распределены равномерно по поверхностям шаров с плотностями $\sigma_1 = q/4\pi r_1^2$ и $\sigma_2 = q/4\pi r_2^2$ (если плотности зарядов на обкладках равны, то количество зарядов на них q_1 и q_2 будет различным). Для применения теоремы Остроградского—Гаусса выберем сферическую поверхность радиуса r $(r_1 > r > r_2)$, концентричную с обкладками. Ввиду симметричного (относительно центра шаров) расположения зарядов на обкладках вектор индукции в зазоре будет везде направлен по радиусу, а по величине — одинаков вдоль выбранной поверхности, поэтому

$$\Psi = 4\pi r^2 D = q; \quad D = q/4\pi r^2,$$

т. е. индукция поля в зазоре между обкладками шарового конденсатора зависит от квадрата расстояния до центра. При этом различие между радиусами шаровых поверхностей r_1 и r_2 не имеет значения.

8. Поле между обкладками цилиндрического конденсатора. Обозначим через r_1 и r_2 радиусы наружного и внутреннего цилиндров, а через τ — заряд, находящийся на единице длины (внутреннего цилиндра). В зазоре между обкладками вектор индукции будет перпендикулярен оси конденсатора, а вдоль цилиндрической поверхности, соосной с обкладками, одинаков по величине. Для такой поверхности $(r_1 > r > r_2)$ имеем:

$$\Psi = 2\pi r l D = \tau l; \quad D = \frac{\tau}{2\pi r}. \tag{1.32}$$

Как и у сферического конденсатора, индукция электрического поля между обкладками цилиндрического конденсатора не зависит от радиусов внешнего и внутреннего цилиндров, а зависит только от расстояния r до оси конденсатора и от линейной плотности заряда на обкладках τ.

ФОРМУЛЫ ДЛЯ НАПРЯЖЕННОСТИ ПОЛЯ

Для расчета напряженности E электрического поля внутри однородного и изотропного диэлектрика необходимо полученные формулы для индукции D разделить на $\varepsilon_0\varepsilon$, например:

1) для точечного заряда или равномерно заряженного шара (за пределами объема шара)

$$E = \frac{q}{4\pi\varepsilon_0\varepsilon r^2} = k\frac{q}{\varepsilon r^2}; \tag{1.33}$$

2) для бесконечно длинной равномерно заряженной нити

$$E = \frac{\tau}{2\pi\varepsilon_0\varepsilon r} = k\frac{2\tau}{\varepsilon r}; \tag{1.34}$$

3) для бесконечной равномерно заряженной плоскости

$$E = \frac{\sigma}{2\varepsilon_0\varepsilon} = k\frac{2\pi\sigma}{\varepsilon}; \tag{1.35}$$

4) между обкладками плоского конденсатора

$$E = \frac{\sigma}{\varepsilon_0\varepsilon} = k\frac{4\pi\sigma}{\varepsilon}, \tag{1.36}$$

5) на оси равномерно заряженного тонкого кольца (R — радиус)

$$E = \frac{q}{4\pi\varepsilon_0\varepsilon}\frac{x}{(R^2+x^2)^{3/2}} = k\frac{q}{\varepsilon}\frac{x}{(R^2+x^2)^{3/2}};$$

6) на оси равномерно заряженного диска (R — радиус)

$$E = \frac{\sigma}{2\varepsilon_0\varepsilon}\left[1 - \frac{x}{\sqrt{R^2+x^2}}\right] = k\frac{2\pi\sigma}{\varepsilon}\left[1 - \frac{x}{\sqrt{R^2+x^2}}\right],$$

где $k = 1/4\pi\varepsilon_0 = 9\cdot10^9$ (Н·м²/Кл²).

В безграничном однородном, изотропном и линейном (см. § 2) диэлектрике можно сформулировать теорему Остроградского — Гаусса и для напряженности E:

$$\oint E\cos\alpha\,dS = k_0\frac{4\pi}{\varepsilon}\sum q_i = \frac{1}{\varepsilon_0\varepsilon}\sum q_i. \tag{1.37}$$

Электрическое поле графически изображается *силовыми линиями индукции* (или *напряженности*), вдоль которых вектор **D** (или **E**) направлен по касательной. У отдельного точечного заряда в вакууме (или в однородном и изотропном диэлектрике) силовые линии направлены по радиусам. Условились величину индукции D (или напряженности E) изображать числом линий, проводимых через единицу площади, перпендикулярной этим линиям. Например, через площадку ΔS (см. рис. III.9) проводят $D\Delta S$ силовых линий. Ввиду этого число силовых линий, проведенных через данную площадку, оказывается равным потоку электрической индукции через эту площадку; тогда, согласно теореме Остроградского — Гаусса, от каждого точечного заряда q следует провести q силовых линий. В изотропной среде силовые линии векторов D и E будут совпадать, отличаясь только числом линий, проводимых через единичные площадки; в анизотропной среде силовые линии этих векторов не будут совпадать.

Если заряженная частица, имеющая заряд e, движется в электрическом поле, то в каждой точке поля действующая на нее сила $F = eE$ будет направлена по касательной к силовой линии на-

пряженности E, проведенной через эту точку. По этой же касательной будет направлено и ускорение движения, однако скорость частицы может составлять любой угол с силовой линией. Очевидно, что направление скорости частицы будет совпадать с направлением силовой линии электрического поля только в частном случае, когда эти линии — прямые и, кроме того, начальная скорость частицы была ориентирована вдоль этих линий.

§ 4. РАБОТА ПЕРЕМЕЩЕНИЯ ЗАРЯДА В ЭЛЕКТРИЧЕСКОМ ПОЛЕ; РАЗНОСТЬ ПОТЕНЦИАЛОВ. ЭНЕРГИЯ ВЗАИМОДЕЙСТВИЯ ЗАРЯДОВ

Допустим, что в электрическом поле выделены две точки: *1* и *2*. Соединим их некоторой линией (рис. III.13) и разобьем эту линию на элементарные отрезки $\mathrm{d}l$, настолько малые, чтобы в пределах каждого из них напряженность поля **E** можно было бы принять постоянной по величине и направлению. Будем перемещать пробный заряд q_0 (точечный, малый, положительный) от первой точки ко второй и рассчитаем работу, совершаемую силой $q_0\mathbf{E}$, действующей на переносимый заряд; на участке $\mathrm{d}l$ работа перемещения равна

$$q_0 E\,\mathrm{d}l\cos\alpha.$$

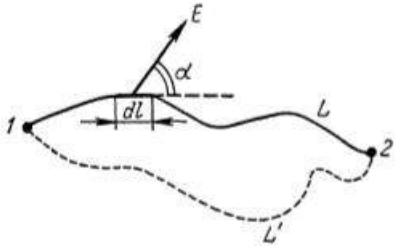


Рис. III.13

Полная работа переноса заряда q_0 из первой точки во вторую

$$A_{12} = q_0 \int_1^2 E\,\mathrm{d}l\cos\alpha. \quad (1.38)$$

Эта работа может быть положительной или отрицательной в зависимости от знаков $\cos\alpha$, т. е. величины углов между напряженностями поля и направлением переноса. (Элементы $\mathrm{d}l$ рассматриваются как векторы, ориентированные в направлении переноса.) Если работа перемещения положительная, то говорят, что она совершается силами поля; отрицательный знак этой работы означает, что для перемещения заряда вдоль выбранной линии к нему должны быть приложены внешние силы (не связанные с полем) и, следовательно, формула (1.38) определяет величину внешней работы, совершаемой при этом переносе.

Выясним, может ли эта работа зависеть от формы и длины траектории L (рис. III.13). Возьмем вторую траекторию L' и допустим, что работа перемещения вдоль этой траектории равна A'_{12}. Если перенести некоторый заряд от первой точки во вторую по траектории, вдоль которой совершается бóльшая работа, а обратный переход осуществить по траектории, вдоль которой совершается меньшая работа, то после этой операции будет получен выигрыш в работе, равный $A_{12}-A'_{12}$ (предполагается, что при переносе заряда напряженность поля со временем не изменяется, т. е. что поле является электростатическим). Так как при этом никаких изменений в электрическом поле, в расположении зарядов (создающих это поле) и т. п. не предполагается, то

указанную операцию можно многократно повторить и таким образом иметь неисчерпаемый источник работы. Такие источники в природе не обнаружены и теоретически не допускаются (согласно закону сохранения энергии). Ввиду этого следует полагать, что $A'_{12}=A_{12}$. Тогда *отношение* A_{12} *к переносимому заряду*, т. е. величина интеграла $\int_1^2 E\,dl\cos\alpha$, *не зависит ни от величины переносимого заряда, ни от формы и размеров пути переноса*, а определяется только расположением указанных точек в данном электрическом поле. Этот интеграл называется *разностью потенциалов точек 1 и 2*:

$$\varphi_1-\varphi_2=\frac{A_{12}}{q_0}=\int_1^2 E\cos\alpha\,dl=\int_1^2 \mathbf{E}\,d\mathbf{l}. \qquad (1.39)$$

Понятие разности потенциалов вводится для характеристики различных точек электрического поля, поэтому в точках, где поля нет, потенциал принимается равным нулю. Допустим, что вторая точка *2* находится в пространстве, где электрического поля нет, например на бесконечно большом расстоянии от электрических зарядов. Тогда $\varphi_2=0$ и

$$\varphi_1=\int_1^\infty E\cos\alpha\,dl. \qquad (1.40)$$

Следовательно, *потенциал данной точки электрического поля равен отношению работы переноса пробного заряда из данной точки поля в другую точку, где электрическое поле отсутствует* (например, в бесконечности), *к величине переносимого заряда*:

$$\varphi=A_0/q_0. \qquad (1.41)$$

Потенциал есть скалярная величина, положительная или отрицательная в зависимости от знака работы переноса A_0 (заряд q_0 условились всегда брать положительным); потенциалы всех точек поля вокруг положительного заряда положительные, а вокруг отрицательного заряда — отрицательные.

Потенциал выражается в *вольтах*; разность потенциалов между двумя точками поля равна одному вольту, если работа переноса одного кулона электричества из одной точки в другую равна одному джоулю: 1 В = 1 Дж/Кл.

В атомной физике работа перемещения элементарных зарядов в электрическом поле (а также энергия частиц) измеряется в *электронвольтах* (эВ); 1 эВ равен работе перемещения электрона, если разность потенциалов начальной и конечной точек перемещения равна 1 В:

$$1\text{эВ}=1{,}6\cdot10^{-19}\text{ Кл}\cdot1\text{В}=1{,}6\cdot10^{-19}\text{ Дж}=1{,}6\cdot10^{-12}\text{ эрг}.$$

Из формулы (1.39) следует, что работа перемещения некоторого заряда из одной точки поля в другую равна произведению этого заряда на разность потенциалов начальной и конечной точек перемещения:

$$A=q(\varphi_1-\varphi_2). \qquad (1.42)$$

Для двух очень близких точек поля, находящихся на расстоянии dl друг от друга, работа переноса $dA = q d\varphi$. Однако при применении аппарата высшей математики следует иметь в виду, что изменением (приращением) какой-нибудь величины x в математике называют не разность $x_1 - x_2$, а разность $x_2 - x_1$. Поэтому для двух бесконечно близких точек изменение (приращение) потенциала $d\varphi$ равно не $\varphi_1 - \varphi_2$, а $\varphi_2 - \varphi_1$, следовательно,

$$dA = -q\,d\varphi; \quad d\varphi = -E\,dl\cos\alpha. \tag{1.43}$$

В электрическом поле можно провести эквипотенциальную поверхность, все точки которой имеют одинаковый потенциал. Работа перемещения заряда вдоль такой поверхности, согласно формуле (1.42), равна нулю. Это возможно, если в каждой точке такой поверхности сила, действующая на заряд $F = qE$, перпендикулярна ей. Вместе с силой F вектор E также будет ориентирован по нормали к эквипотенциальной поверхности.

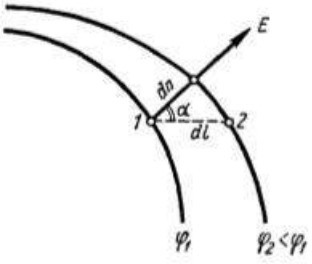


Рис. III.14

Допустим, что некоторый заряд q перемещается по направлению напряженности поля E на расстояние Δl. Тогда работа перемещения $A = qE\Delta l$ будет положительной. Согласно формуле (1.42), положительная работа перемещения означает, что $\varphi_1 > \varphi_2$, т. е. при таком перемещении конечная точка траектории имеет меньший потенциал, чем начальная. Таким образом, *по направлению напряженности поля потенциал убывает.*

На рис. III.14 показаны две эквипотенциальные поверхности, проходящие через точки *1* и *2*, имеющие потенциалы φ_1 и φ_2. Если точки *1* и *2* бесконечно близки друг к другу, то $dl\cos\alpha$ есть расстояние между эквипотенциальными поверхностями, измеренное по нормали к ним, т. е. по направлению вектора напряженности E. Тогда из формулы (1.43) следует

$$E = -\frac{d\varphi}{dn}, \tag{1.44}$$

т. е. *напряженность в данной точке поля численно равна изменению потенциала на единицу расстояния, взятого вдоль нормали к эквипотенциальной поверхности, проходящей через эту точку.* Знак минус показывает, что напряженность поля направлена в сторону убывания потенциала.

Работа перемещения заряда q между двумя точками электрического поля не зависит от пути перехода и поэтому может быть представлена, например, в виде суммы работ, совершаемых последовательно на перемещениях dx, dy и dz вдоль координатных осей. Представим эту работу через изменения потенциалов вдоль координатных осей ($d\varphi_x$, $d\varphi_y$ и $d\varphi_z$); тогда для элементарной работы dA можно написать:

$$-dA = q\,d\varphi = q\,d\varphi_x + q\,d\varphi_y + q\,d\varphi_z.$$

Разности потенциалов на элементарных участках пути dx, dy и dz можно представить в виде:

$$d\varphi_x = \frac{\partial \varphi}{\partial x} dx; \quad d\varphi_y = \frac{\partial \varphi}{\partial y} dy; \quad d\varphi_z = \frac{\partial \varphi}{\partial z} dz,$$

где $\frac{\partial \varphi}{\partial x}$, $\frac{\partial \varphi}{\partial y}$ и $\frac{\partial \varphi}{\partial z}$ показывают изменение потенциала на единицу длины вдоль соответствующих координатных осей. Таким образом,

$$-dA = q\frac{\partial \varphi}{\partial x} dx + q\frac{\partial \varphi}{\partial y} dy + q\frac{\partial \varphi}{\partial z} dz.$$

С другой стороны (см. ч. I, § 2), dA равна сумме работ, совершаемых компонентами силы $F = qE$, следовательно,

$$dA = qE_x dx + qE_y dy + qE_z dz.$$

Сравнивая эти выражения, получаем:

$$E_x = -\frac{\partial \varphi}{\partial x}; \quad E_y = -\frac{\partial \varphi}{\partial y}; \quad E_z = -\frac{\partial \varphi}{\partial z}.$$

Обозначая единичные («направляющие») вектора координатных осей через $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, можно представить вектор напряженности электрического поля в данной точке в зависимости от скоростей изменения потенциала вдоль координатных осей:

$$\mathbf{E} = -\left(\frac{\partial \varphi}{\partial x}\mathbf{i} + \frac{\partial \varphi}{\partial y}\mathbf{j} + \frac{\partial \varphi}{\partial z}\mathbf{k}\right).$$

Величина, записанная в скобках и равная сумме скоростей изменения какой-нибудь скалярной величины (характеризующей поле) вдоль координатных осей, называется *градиентом* этой величины и обозначается

$$\mathbf{grad}\,\varphi = \frac{\partial \varphi}{\partial x}\mathbf{i} + \frac{\partial \varphi}{\partial y}\mathbf{j} + \frac{\partial \varphi}{\partial z}\mathbf{k}.$$

Таким образом, *напряженность электрического поля в каждой точке равна градиенту потенциала поля в этой точке.*

Связь между напряженностью поля и потенциалом имеет наиболее простой вид для однородного поля, когда E постоянно (например, для поля между обкладками плоского конденсатора). Воспользуемся формулой (1.39) и будем интегрировать вдоль направления вектора E ($\cos\alpha = +1$):

$$\varphi_1 - \varphi_2 = \int_1^2 E\,dl\cos\alpha = E\int_1^2 dl = El;$$

$$E = \frac{\varphi_1 - \varphi_2}{l}. \qquad (1.45)$$

Рассчитаем потенциалы точек поля в вакууме ($\varepsilon = 1$) вокруг точечного заряда q, допустим положительного. Применим формулу (1.40), и так как выбор пути перемещения безразличен, то проведем интегрирование вдоль радиуса от точки A до бесконечности.

Тогда

$$E = kq/r^2; \quad \cos\alpha = +1; \quad dl = dr \quad \text{и}$$

$$\varphi = \int_r^\infty E\, dl \cos\alpha = k \int_r^\infty \frac{q}{r^2}\, dr = +k\frac{q}{r}, \tag{1.46}$$

т. е. *потенциал точечного заряда по абсолютной величине убывает обратно пропорционально расстоянию.* Если заряд q отрицательный, то $\cos\alpha = -1$ и $\varphi = -kq/r$. Любое заряженное тело, а также любую систему зарядов различных знаков можно представить как совокупность множества точечных зарядов, каждый из которых создает в данной точке поля напряженность E_i и потенциал φ_i. Тогда при соблюдении принципа суперпозиции полей $\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 + \ldots$ легко получить $\varphi = \varphi_1 + \varphi_2 + \ldots$, т. е. *потенциал нескольких точечных зарядов в каждой точке поля равен алгебраической сумме потенциалов, создаваемых в этой точке каждым зарядом в отдельности.*

Таким образом, потенциалы электрического поля вокруг заряженных тел различной формы и размеров можно вычислить либо суммируя потенциалы, создаваемые точечными зарядами, расположенными на этих телах, либо пользуясь связью между потенциалом и напряженностью поля (1.44). В частности, потенциал электрического поля вокруг равномерно заряженного шара, имеющего общий заряд q, выражается той же формулой (1.46), которая была получена для точечного заряда; в безграничном диэлектрике

$$\varphi = k\frac{q}{\varepsilon r} = \frac{q}{4\pi\varepsilon_0\varepsilon r}.$$

На самой поверхности шара (R — радиус шара) потенциал равен

$$\varphi = \frac{q}{4\pi\varepsilon_0\varepsilon R};$$

в том случае, когда заряд шара расположен только по его поверхности, эта формула может быть переписана в виде

$$\varphi = \frac{\sigma}{\varepsilon_0\varepsilon} R,$$

где $\sigma = q/4\pi R^2$ — плотность заряда на поверхности шара.

На рис. III.15 показаны графики зависимости потенциала электрического поля от расстояния до центра шара (*а*) и оси длинного цилиндра (*б*), равномерно заряженных по поверхности, а также изменение потенциала между обкладками плоского конденсатора (*в*); здесь обозначены: R — радиусы шара и цилиндра, τ — заряд на единице длины цилиндра, σ — поверхностная плотность зарядов на обкладках конденсатора.

Рассчитаем потенциалы в точках A и B электрического диполя (см. рис. III.5, *б*). В точке A сумма потенциалов от положительного φ_+ и отрицательного φ_- зарядов равна

$$\varphi_A = \varphi_+ + \varphi_- = k\frac{(+q)}{r+l/2} + k\frac{(-q)}{r-l/2} = -k\frac{ql}{r^2}\frac{1}{1-(l/r)^2} \approx -k\frac{ql}{r^2}.$$

Потенциал получился отрицательным, так как для точки A отрицательный заряд расположен ближе, чем положительный. В точке B потенциалы от обоих зарядов равны по величине и противоположны по знаку, поэтому $\varphi_B = 0$.

В разделе «Механика» (ч. I, § 12) было указано, что потенциальная энергия системы двух точечных взаимодействующих тел, в частности электрических зарядов q и q_0, равна

$$U = \pm k \frac{qq_0}{r}, \qquad (1.47)$$

где знак плюс берется для отталкивающихся тел, а знак минус — для притягивающихся тел. Сравнивая формулы (1.46) и (1.47), получаем, что

$$U = q_0\varphi; \quad \varphi = U/q_0,$$

т. е. *потенциал данной точки поля точечного заряда q численно равен потенциальной энергии системы, состоящей из заряда q и единичного заряда, помещенного в эту точку поля*. Однако следует подчеркнуть, что потенциал и потенциальная энергия есть различные физические величины и имеют различные размерности (вольт и джоуль).

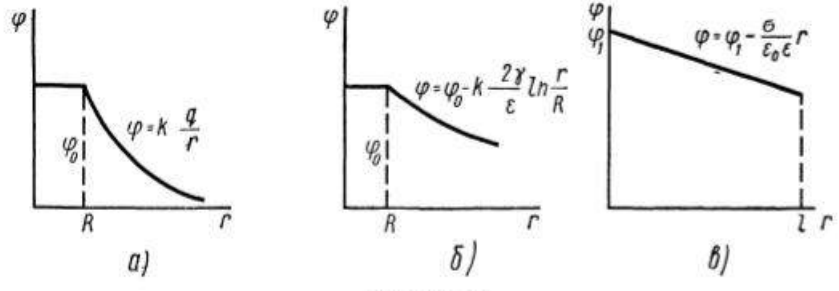


Рис. III.15

Воспользуемся связью между U и φ и рассчитаем потенциальную энергию диполя, помещенного в электрическое поле; если положительный заряд диполя $(+q)$ находится в точке с потенциалом φ_1, а отрицательный $(-q)$ — в точке с потенциалом φ_2, то

$$U = (+q)\varphi_1 + (-q)\varphi_2 = q(\varphi_1 - \varphi_2).$$

В поле с постоянным потенциалом ($\varphi = \text{const}$; $E = 0$) на диполь не будут действовать силы и поэтому $U = 0$. Если поле однородное ($E = \text{const}$), то, согласно (1.43),

$$\varphi_1 - \varphi_2 = -El\cos\alpha = -\mathbf{El},$$

где l — плечо диполя, и тогда

$$U = -qlE\cos\alpha = -q\mathbf{l}\mathbf{E} = -\mathbf{pE}.$$

Знак потенциальной энергии диполя зависит от того, какой из зарядов — положительный или отрицательный — находится в точке с бо́льшим потенциалом.

Если диполь находится в электрическом поле точечного (или сферически симметричного) заряда Q и если он ориентирован по радиусу, то

$$\varphi_1 = k\frac{Q}{r - l/2}; \quad \varphi_2 = k\frac{Q}{r + l/2}; \quad U = kqQ\frac{l}{r^2}\frac{1}{1-(l/2r)^2} \approx k\frac{qlQ}{r^2},$$

где r — расстояние от центра заряда Q до центра диполя.

Пользуясь этим результатом, можно легко рассчитать потенциальную энергию одного диполя в электрическом поле такого же диполя. Допустим, что диполи ориентированы одинаково и расположены на одной прямой. Обозначим через r_1 и r_2 расстояния между центром одного диполя и зарядами другого; тогда, суммируя потенциальные энергии положительных и отрицательных зарядов диполя, получим:

$$U = U_+ + U_- = k\frac{(+q)\,ql}{r_1^2} + k\frac{(-q)\,ql}{r_2^2} = kq^2l\,\frac{r_2^2 - r_1^2}{r_1^2 r_2^2}.$$

Если $r_2 - r_1 = l$ значительно меньше r_1 и r_2, то, полагая $(r_1 + r_2)/r_1^2 r_2^2 \approx 2/r^3$, получим

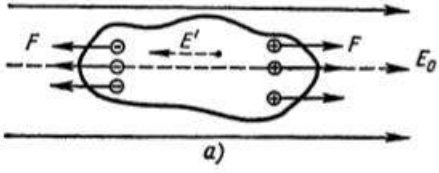
$$U \approx k \frac{2q^2 l^2}{r^3}$$

(r — расстояние между центрами диполей).

Это есть потенциальная энергия взаимодействия двух одинаковых диполей, ориентированных указанным выше образом.

§ 5. ПРОВОДНИКИ В ЭЛЕКТРИЧЕСКОМ ПОЛЕ

Если в электрическое поле внести проводник, то на находящиеся в нем свободные заряды будут действовать силы, которые для зарядов различных знаков направлены в противоположные стороны. Эти силы вызовут некоторое перераспределение свободных зарядов в пределах объема проводника (рис. III.16, *а*). По мере разделения положительных и отрицательных зарядов в объеме проводника образуется собственное электрическое поле с напряженностью $\mathbf{E}'$, противоположной напряженности внешнего поля $\mathbf{E}_0$. Смещение свободных зарядов будет происходить до тех пор, пока в объеме проводника суммарная напряженность



а)

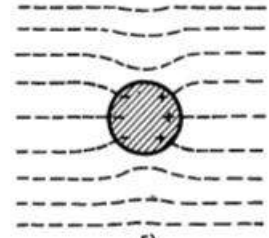

б)

Рис. III.16

поля $\mathbf{E}_0 + \mathbf{E}'$ отлична от нуля, т. е. пока на свободные заряды действуют силы. Как только $\mathbf{E}_0 + \mathbf{E}'$ сделается равной нулю в каждой точке объема проводника, перераспределение зарядов прекратится. Равенство нулю напряженности поля внутри проводника, согласно формуле (1.44), означает, что в пределах объема проводника потенциал либо везде равен нулю, либо везде одинаков ($d\varphi = 0$, если $\varphi = 0$ или $\varphi = \text{const}$).

Допустим, что какому-нибудь участку проводника сообщен некоторый электрический заряд. В течение короткого времени этот заряд независимо от того, существует ли внешнее электрическое поле или оно отсутствует, будет распределяться по объему проводника, пока в пределах этого объема напряженность поля не станет равной нулю, а потенциал — всюду одинаковым. Таким образом, *в равновесном состоянии проводник является эквипотенциальным телом* независимо от того, нейтрален ли он в целом или же в нем имеется избыточный заряд того или иного знака.

Перераспределение свободных зарядов в проводниках под действием внешнего электрического поля называется *электростатической индукцией*. Разделившиеся положительные и отрицательные заряды

в проводниках создают вне проводника поле, которое накладывается на внешнее поле и «искажает» его.

На рис. III.16, *б* показано изменение однородного поля при внесении незаряженного металлического шара. В равновесном состоянии напряженность суммарного поля вдоль поверхности проводника всюду перпендикулярна ей, так как только при этом условии может прекратиться перемещение свободных зарядов вдоль поверхности проводника. Таким образом, поверхность проводника, как и весь проводник, является эквипотенциальной.

Можно показать, что в объеме, ограниченном со всех сторон проводником, электрическое поле будет отсутствовать. Действительно, проводник есть эквипотенциальное тело, поэтому напряженность поля, если бы она существовала, должна быть повсюду перпендикулярной поверхности (рис. III.17). Проведем замкнутую поверхность, показанную пунктиром; внутри нее избыточных зарядов нет, поэтому поток электрической индукции через эту поверхность, по теореме Остроградского — Гаусса, должен равняться нулю. Так как индукция D направлена внутрь поверхности, то равенство нулю потока индукции возможно, если только $D = 0$, следовательно, $E = 0$. Отсутствием поля в пространстве, охваченном проводником, пользуются для *электростатической защиты*, т. е. предохранения тех или иных объектов от действия внешнего электрического поля.

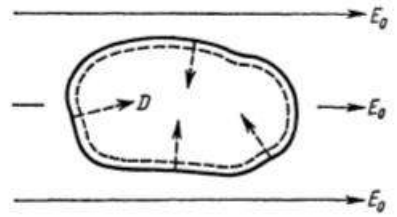


Рис. III.17

Индукция и напряженность поля вблизи поверхности заряженного проводника пропорциональны поверхностной плотности σ зарядов. Применим теорему Остроградского — Гаусса к элементарному цилиндру с основанием ΔS, ось которого ориентирована вдоль вектора D (рис. III.18, *а*). Так как поле внутри проводника отсутствует, то поток вектора D будет только на наружном основании цилиндра, следовательно, согласно формуле (1.28),

$$D\,\Delta S = \sum q_i = \sigma\,\Delta S;$$
$$D = \sigma; \quad E = \sigma/\varepsilon_0\varepsilon. \tag{1.48}$$

Следовательно, *индукция поля у поверхности заряженного проводника равна поверхностной плотности зарядов,* а напряженность поля прямо пропорциональна ей и, кроме того, обратно пропорциональна диэлектрической проницаемости среды.

Интересно обсудить вопрос о том, не противоречит ли соотношение (1.48) полученному ранее соотношению (1.32) в случае, если заряженный проводник имеет вид плоского тела больших размеров. Прежде всего заметим, что формула (1.48) ранее была получена также и для равномерного заряженного шара и цилиндра. Для того чтобы показать, что она применима и для равномерно заряженного плоского тела любой толщины, в частности и для заряженной плоскости (см. рис. III.7

и формулу (1.25)), применим теорему Остроградского — Гаусса к цилиндру с основанием ΔS, внутри которого находится заряд Δq (рис. III. 18, *б*). Полагая, что заряды расположены только по обеим поверхностям тела с одинаковой плотностью σ, получим

$$\Delta q = 2\sigma \Delta S.$$

По теореме Остроградского — Гаусса,

$$D \cdot 2\Delta S = \Delta q; \quad D = \frac{\Delta q}{2\Delta S} = \sigma.$$

Это соотношение, совпадающее с выражением (1.48), не зависит от толщины пластинки l. Если уменьшать эту толщину: $l \to 0$, то полученное соотношение должно сохраниться и в пределе. Однако в данном

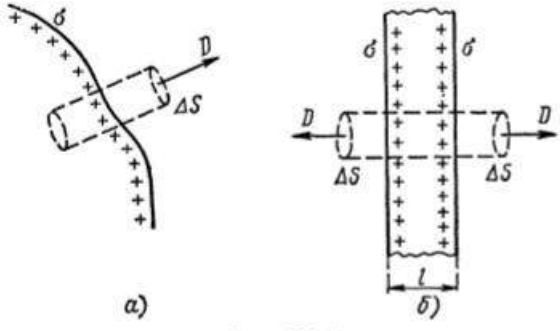


Рис. III.18

случае формула $D = \sigma$ содержит плотность зарядов $\sigma = \Delta q/2\Delta S$ только на одной стороне поверхности, тогда как при выводе формулы (1.25) для бесконечно тонкой пластинки под σ понимается отношение Δq не к $2\Delta S$, а только к ΔS. Ввиду этого поверхностная плотность зарядов в формуле (1.25) должна быть при одинаковом D вдвое больше той поверхностной плотности, которая написана выше для пластинки. Этим устраняется кажущееся противоречие между формулами (1.32) и (1.48). Таким образом, соотношение (1.48) дает индукцию электрического поля вблизи поверхности проводника л ю б о й ф о р м ы.

На поверхности проводника, имеющего форму шара, в равновесном состоянии поверхностная плотность избыточных зарядов σ везде одинакова. Однако у тел произвольной формы эта плотность различна в различных местах в зависимости от кривизны поверхности в данном месте. Выберем на поверхности тела некоторую малую площадку ΔS, которую с достаточным приближением можно было бы полагать частью сферы радиуса r. Применим к этой площадке соотношение между потенциалом и плотностью зарядов, которое было выведено выше для шара:

$$\varphi = \frac{\sigma}{\varepsilon_0 \varepsilon} r; \quad \sigma = \frac{\varphi}{r} \varepsilon_0 \varepsilon.$$

Так как потенциал φ одинаков во всех точках поверхности тела, то σ должна быть обратно пропорциональна радиусу r кривизны поверх-

ности в данном месте. Из этих рассуждений следует, что на остриях (где r очень мало) поверхностная плотность зарядов σ, а следовательно, и индукция $D = \sigma$ и напряженность электрического поля $E = D/\varepsilon_0\varepsilon$ должны быть очень велики, что подтверждается наблюдениями (так называемое «стекание» заряда с острия, сильная ионизация воздуха вблизи них и другие явления).

§ 6. ДИЭЛЕКТРИКИ В ЭЛЕКТРИЧЕСКОМ ПОЛЕ

В диэлектриках количество ионов и электронов (в единице объема), способных перемещаться в пределах объема тела, очень мало, поэтому при внесении диэлектрика в электрическое поле перераспределение этих зарядов не производит заметного эффекта. Существенное значение имеет воздействие электрического поля на с в я з а н н ы е з а р я д ы, имеющиеся в атомах и молекулах диэлектрика.

Рассмотрим сначала действие электрического поля на простейшую систему зарядов — *диполь* (рис. III.19).

Если поле однородное, то силы $F_{(+)} = +qE_0$ и $F_{(-)} = -qE_0$ равны и, следовательно, на диполь действует момент пары сил:

$$M = Fl \sin\alpha = qlE_0 \sin\alpha. \qquad (1.49)$$

Эта формула может быть записана в виде векторного произведения:

$$\mathbf{M} = [\mathbf{p}_э\mathbf{E}],$$

где $\mathbf{p}_э = q\mathbf{l}$ — вектор электрического момента диполя.

В однородном поле $F_{(+)}$ и $F_{(-)}$ составляют «пару сил».

Если же диполь находится в неоднородном поле, то силы $F_{(+)}$ и $F_{(-)}$ не будут равны и тогда кроме момента M на диполь будет действовать равнодействующая (F) сил $F_{(+)}$ и $F_{(-)}$, равная их разности. Очевидно, эта сила F будет сообщать диполю поступательное движение в направлении возрастания напряженности поля; величина F будет пропорциональна степени неоднородности поля в данном месте.

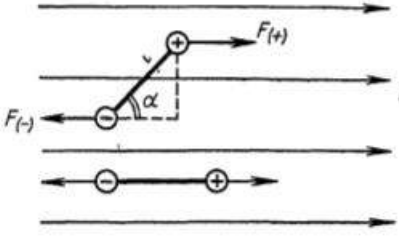


Рис. III.19

Момент M, действующий на диполи, стремится ориентировать их вдоль напряженности поля (рис. III.19). В этом состоянии силы «деформируют» диполь, т. е. увеличивают расстояние между зарядами. Диполи называются *жесткими,* если силы, связывающие заряды между собой, значительно больше сил, действующих со стороны внешнего электрического поля. Такие диполи в электрическом поле не испытывают заметной деформации. Диполи называются *упругими,* если они во внешнем электрическом поле заметно деформируются, т. е. изменяют l, а следовательно, и момент $p_э$. При удалении поля упругие диполи возвращаются в начальное состояние.

Допустим, что рассматриваемый диэлектрик состоит из молекул, которые во внешнем электрическом поле ведут себя как диполи, т. е. обладают некоторым электрическим моментом $p_э$. Выделим единичный объем и составим векторную сумму моментов всех молекул (диполей) этого объема:

$$\mathbf{P}_э = \sum \mathbf{p}_э. \qquad (1.50)$$

В отсутствие внешнего поля тепловое движение создает полный беспорядок в ориентировках диполей, поэтому $\mathbf{P}_э = 0$. Если же диэлектрик вводится в электрическое поле, то на его молекулы действует момент сил, ориентирующий их вдоль напряженности поля. Этой ориентировке препятствует беспорядочное тепловое движение, однако в среднем проекция векторов $\mathbf{p}_э$ на направление поля будет отлична от нуля, поэтому $\mathbf{P}_э \neq 0$. Очевидно, вектор $\mathbf{P}_э$ зависит от напряженности внешнего поля и температуры диэлектрика.

ПОЛЯРИЗАЦИЯ ДИЭЛЕКТРИКОВ

Появление электрического момента у элементарных объемов диэлектрика (содержащих, однако, настолько большое число молекул, чтобы можно было применять к этим объемам понятие температуры) называется поляризацией диэлектрика, а электрический момент единицы объема $\mathbf{P}_э$ — *поляризованностью* (*вектором* или *интенсивностью поляризации*). Укажем на некоторые виды поляризации:

1) упругая поляризация электронного и ионного смещения; внешнее поле смещает электронную оболочку атомов относительно положительно заряженных ядер; в кристаллической решетке, состоящей из ионов противоположных знаков (например, NaCl, KCl и др.), происходит смещение этих ионов в противоположных направлениях. Эти смещения делают вектор $\mathbf{P}_э$ отличным от нуля;

2) ориентационная, или дипольная, поляризация; если молекулы диэлектрика обладают собственным электрическим моментом (вода, нитробензол, аммиак и др.), то внешнее поле, как указывалось выше, ориентирует их. При этом возможна и деформация диполей, т. е. некоторое увеличение их собственного электрического момента;

3) спонтанная (самопроизвольная) поляризация; наблюдается у некоторых диэлектриков (сегнетова соль, титанат бария и др.), в объеме которых возникают сравнительно большие поляризованные области, называемые *доменами*. При отсутствии внешнего поля электрические моменты этих доменов ориентированы в объеме диэлектрика беспорядочно, поэтому весь диэлектрик в целом не поляризован. Внешнее поле создает некоторый порядок в ориентациях этих моментов.

Поляризация диэлектрика сопровождается появлением поверхностных зарядов на его границах, а также появлением внутреннего поля E', образованного этими зарядами.

На рис. III.20, *а* показан кубик из диэлектрика, на одной грани которого вследствие ориентирующего действия внешнего поля появился избыток положительных, а на другой грани — избыток отрицательных зарядов. Эти поверхностные заряды создают электрическое поле, напряженность которого E' внутри кубика направлена против напряженности внешнего поля E_0. На рис. III.20, *б* показана ориентация диполей диэлектрика вокруг положительно заряженного тела; отрицательные заряды диполей, обращенные к поверхности тела, будут заметно ослаблять поле этого тела (этим объясняется уменьшение силы взаимодействия заряженных тел в диэлектрических средах по сравнению с вакуумом). Если диэлектрик неоднородный, то упорядоченное расположение диполей приводит к появлению внутри его *объемных зарядов* (при этом, однако, весь диэлектрик остается электрически нейтральным). На рис. III.20, *в* показаны несколько участков

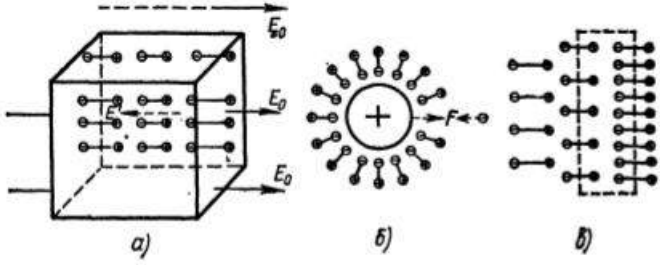


Рис. III.20

неоднородного диэлектрика с различным числом диполей в единице объема. В участке, очерченном пунктирной линией, число положительных и отрицательных зарядов различное; объемный заряд этого участка будет создавать дополнительное электрическое поле, пропорциональное степени неоднородности диэлектрика.

Поверхностные заряды и внутреннее поле E' появляются также у диэлектрика, помещенного между обкладками заряженного плоского конденсатора. На этом примере удобно показать, что введение диэлектрика в поле может изменить взаимное расположение зарядов, создающих это поле. При отсутствии диэлектрика (или когда диэлектрик занимает все поле; рис. III. 21, *а*) заряды на обкладках распределяются равномерно; вектор индукции поля **D** будет везде одинаков и численно равен $D = \sigma$. Если же диэлектрик занимает часть поля, то заряды на обкладках, вследствие взаимодействия с зарядами, появившимися на поверхности диэлектрика, будут перемещаться в сторону диэлектрика, и поэтому равномерное распределение зарядов вдоль обкладок конденсатора нарушится. Индукция поля (рис. III.21, *б*) в точке A ($D_A = \sigma_1$) будет отличаться от индукции в точке B ($D_B = \sigma_2$). Однако напряженности полей в этих точках будут равны, так как обкладки конденсатора (после перераспределения зарядов в них) являются эквипотенциальными поверхностями, и поэтому $E_A = E_B = (\varphi_1 - \varphi_2)/l$. Согласно формуле (1.36), $E_A = \sigma_1/\varepsilon_0$, $E_B = \sigma_2/\varepsilon_0\varepsilon$. Следовательно, $\sigma_2 = \varepsilon\sigma_1$. Так как заряды, появившиеся на поверхно-

сти диэлектрика, создают поле $E' = \sigma'/\varepsilon_0$, противоположное полю $E_0 = \sigma_2/\varepsilon_0$ зарядов на обкладках (которое является внешним полем), то

$$E_B = E_0 - E' = \frac{\sigma^2}{\varepsilon_0} - \frac{\sigma'}{\varepsilon_0}; \quad \sigma' = \sigma_2 - \varepsilon_0 E_B.$$

Следовательно, поверхностная плотность зарядов, появившихся на границах диэлектрика:

$$\sigma' = \sigma_2\left(1 - \frac{1}{\varepsilon}\right) = \sigma_1(\varepsilon - 1). \qquad (1.51)$$

Для расчета электрического момента единицы объема, взятого внутри диэлектрика $\mathbf{P}_э$, допустим, что в поляризованном диэлектрике молекулярные диполи составляют непрерывные цепочки, так что на участке l, взятом вдоль поля (рис. III. 22), находятся n_1 одинаковых диполей, каждый из которых имеет электрический момент $p_э = el_0$ (l_0 — расстояние между зарядами e диполя). Такой участок будет

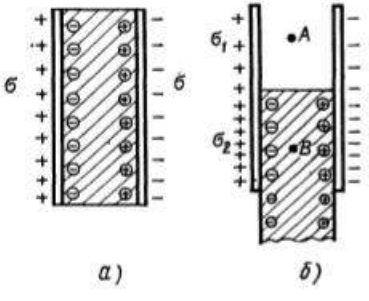


Рис. III.21

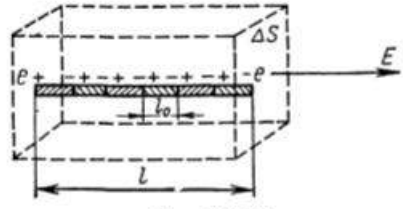


Рис. III.22

иметь общий электрический момент, равный $n_1 p_э = n_1 e l_0 = el$. Если площадка ΔS, перпендикулярная к полю, охватывает N таких цепочек, то электрический момент объема диэлектрика, выделенного на рис. III. 22 пунктиром, будет равен Nel. Заметим, что $Ne = q'$ (суммарный заряд на площадке ΔS) составлен только из зарядов, находящихся на концах мысленно отрезанных нами участков дипольных цепочек. Общий же заряд на площадке ΔS равен нулю, так как цепочки не оканчиваются на ней; на любой площадке ΔS будет находиться равное число положительных и отрицательных зарядов. Если бы площадка ΔS находилась на границе диэлектрика, то $Ne = q'$ были бы избыточными зарядами (одного знака), появившимися на этой площадке вследствие поляризации. Внутри же диэлектрика заряд q' имеет только вспомогательное (расчетное) значение (q' иногда называют «условным» или «фиктивным» зарядом, а $\sigma' = q'/\Delta S$ — плотностью этих зарядов).

Электрический момент единицы объема внутри поляризованного диэлектрика (т. е. значение вектора поляризации) в данном месте будет, по определению, равен:

$$P_э = Nel/l\Delta S = q'/\Delta S = \sigma'. \qquad (1.52)$$

Таким образом, для расчета $P_э$ в данном месте диэлектрика необходимо найти σ' — плотность условных зарядов на элементарной пло-

щадке ΔS, расположенной перпендикулярно к полю (т. е. лежащей на эквипотенциальной поверхности, проведенной в данном месте).

Для расчета напряженности внутреннего поля E' в зависимости от σ' и $P_э$, рассмотрим другой пример. Допустим, что равномерно заряженный шар с зарядом q находится внутри безграничного однородного и изотропного диэлектрика. Будем по-прежнему полагать, что при поляризации диэлектрика молекулярные диполи образуют цепочки, начинающиеся у поверхности шара и расположенные радиально. Так как внутри диэлектрика избыточных зарядов нет, а противоположные заряды диполей в объеме диэлектрика расположены весьма близко друг к другу, то можно полагать, что поле E' создается только зарядами, находящимися на концах цепочек, обращенных к поверхности шара (т. е. на ближайшей границе диэлектрика, так как другая граница находится в бесконечности).

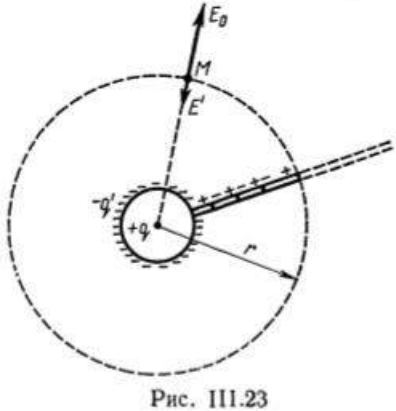


Рис. III.23

Проведем из центра рассматриваемого шара некоторую сферическую (эквипотенциальную) поверхность радиуса r (рис. III. 23). Обозначим суммарный заряд, появившийся на границе диэлектрика (вдоль поверхности шара) через q'. Такой же суммарный заряд (но уже «условный») будут иметь вторые концы цепочек на эквипотенциальной поверхности сферы. Следовательно, плотность условных зарядов на сферической эквипотенциальной поверхности (т. е. на расстоянии r от центра заряженного шара) будет

$$\sigma' = q'/4\pi r^2.$$

Напряженность электрического поля E', которая создается зарядами диэлектрика q', появившимися вдоль поверхности шара, на расстоянии r от центра шара равна

$$E' = q'/4\pi\varepsilon_0 r^2 = \sigma'/\varepsilon_0.$$

Следовательно, согласно формуле (1.52),

$$E' = P_э/\varepsilon_0; \quad P_э = \varepsilon_0 E'. \tag{1.53}$$

Таким образом, напряженность E' собственного электрического поля зарядов поляризованного диэлектрика в данном месте его объема пропорциональна электрическому моменту единицы объема диэлектрика в этом месте или поверхностной плотности условных зарядов на эквипотенциальной поверхности, проведенной через это место.

Если бы диэлектрик отсутствовал, т. е. электрическое поле было бы создано только «избыточными» зарядами q шара, то в точке M индукция D и напряженность E_0 были бы равны:

$$D = q/4\pi r^2; \quad E_0 = D/\varepsilon_0 = q/4\pi\varepsilon_0 r^2.$$

При наличии диэлектрика индукция D избыточных зарядов не изменяется, но напряженность поля уменьшается в ε раз: $E = E_0/\varepsilon$. Это уменьшение объясняется появлением собственного поля зарядов диэлектрика, всегда направленного против внешнего поля, поэтому E есть вектор напряженности суммарного поля:

$$E = E_0 - E' = E_0/\varepsilon.$$

Следовательно,

$$\begin{aligned} E' &= \left(1 - \frac{1}{\varepsilon}\right) E_0 = (\varepsilon - 1)\, E; \\ P_э &= \varepsilon_0 E' = \varepsilon_0 (\varepsilon - 1)\, E; \\ D &= \varepsilon_0 \varepsilon E = \varepsilon_0 E + P_э. \end{aligned} \tag{1.54}$$

Безразмерная величина $\varepsilon - 1 = k$ называется *относительной диэлектрической восприимчивостью*:

$$k = \varepsilon - 1; \quad \varepsilon = 1 + k. \tag{1.55}$$

Тогда соотношение

$$P_э = k\varepsilon_0 E \tag{1.56}$$

означает, что *вектор поляризации прямо пропорционален напряженности суммарного электрического поля в диэлектрике*. Произведение $k\varepsilon_0 = k_а$ называется *абсолютной диэлектрической восприимчивостью*.

ЭЛЕКТРОСТРИКЦИЯ

Силы, действующие со стороны внешнего электрического поля на заряды, содержащиеся в молекулах диэлектрика, вызывают деформацию (сжатие или растяжение) этого диэлектрика и появление вну-

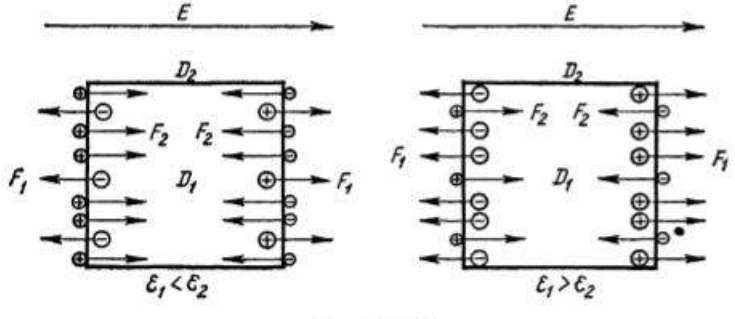


Рис. III.24

тренних механических напряжений. Это явление называется электростриксией. На рис. III.24 показаны силы, действующие на поверхностные заряды в случае, когда кусок диэлектрика D_1 с диэлектрической проницаемостью ε_1 находится в среде D_2 с диэлектри-

ческой проницаемостью ε_2. Кружками большого диаметра обозначены поверхностные заряды в D_1, малого диаметра — окружающей среды D_2. Число таких зарядов на единицу поверхности пропорционально величине диэлектрической проницаемости. Если $\varepsilon_1 < \varepsilon_2$, то напряжение сжатия от сил F_2 со стороны окружающей среды больше напряжения растяжения от сил F_1, приложенных к телу; тело D_2 сжимается. Если $\varepsilon_1 > \varepsilon_2$, то тело растягивается.

Деформация диэлектрика, помещенного в электрическое поле, может быть вызвана и другими причинами. Например, переориентация диполей и изменение электрического момента молекул вызовет некоторые изменения во взаимодействии молекул; изменение сил, связывающих молекулы диэлектрика, может вызвать изменение его размеров. Кроме того, если внешнее поле неоднородное, то силы, действующие на диполи, как было показано выше, имеют равнодействующую, направленную в сторону возрастания напряженности поля. Если тело не закреплено, то оно получит движение в сторону возрастания напряженности поля; этим объясняется притяжение кусков бумаги, ваты и т. д. наэлектризованными телами. Если же диэлектрик закреплен, то силы, действующие на диполи, вызывают дополнительную деформацию диэлектрика, зависящую от степени неоднородности поля.

ПЬЕЗОЭЛЕКТРИЧЕСКИЙ ЭФФЕКТ

У многих кристаллических тел (кварц, сегнетова соль, турмалин, сахар) наблюдается явление, называемое *пьезоэлектрическим эффектом*: если из кристалла вырезать пластинку, плоскости (грани) которой перпендикулярны так называемой *электрической оси кристалла,* то при сжатии этой пластинки на ее гранях появляются противоположные заряды; при растяжении знаки зарядов меняются. Наблюдается и *обратный пьезоэлектрический эффект*: если к плоскостям таких пластинок подать заряды противоположных знаков, то пластинка сжимается или растягивается. Поверхностная плотность зарядов σ на плоскостях и механическое давление p связаны соотношением

$$\sigma = \delta p,$$

где δ — пьезоэлектрическая постоянная; для кварца в системе СГС $\delta = 6{,}4 \cdot 10^{-8}$ ед. СГС$_\sigma \cdot$ см^2/дин.

Пьезоэлектрический эффект получил широкое применение в технике, в частности для возбуждения ультразвуковых колебаний. Для этого к граням кварцевой пластинки прилагается электрическое напряжение, меняющееся со временем с ультразвуковой частотой. Деформирующийся кристалл возбуждает в окружающей среде (в воздухе, жидкостях и твердых телах, с которыми он соприкасается) ультразвуковые волны. Обратный пьезоэлектрический эффект используется для измерения больших или быстроизменяющихся давлений, в частности для изучения ультразвуковых колебаний; электрическое напряжение, которое появляется на гранях деформируемого кристалла, усиливается и затем подается к измерительным приборам.

В газообразных, жидких и во многих твердых телах поляризация появляется и исчезает вместе с появлением и исчезновением внешнего электрического поля, причем вектор поляризации **P** прямо пропорционален напряженности поляризующего поля **E** (см. формулу (1.56)). Однако в некоторых твердых (кристаллических) телах диэлектрическая восприимчивость в формуле (1.56) не есть постоянная величина и заметно зависит от напряженности поля E; на характер этой зависимости сильно влияет температура тела. К таким телам относятся с е г н е т о э л е к т р и к и: сегнетова соль, титанат бария и др. Пока температура тела больше некоторой «критической» температуры θ, то диэлектрическая

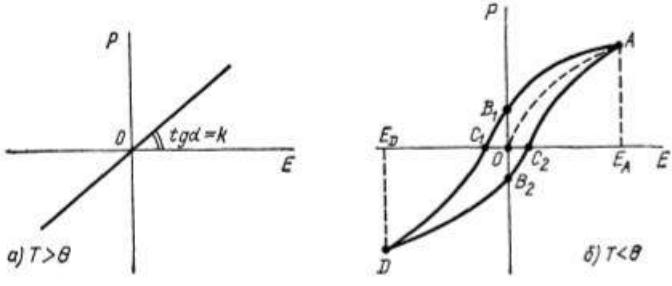


Рис. III.25

восприимчивость k не зависит от напряженности внешнего поля E, но зависит от температуры тела:

$$k\,(T-\theta)=\text{const};\quad k=\frac{\text{const}}{T-\theta}. \tag{1.57}$$

При $T<\theta$ наблюдается заметная зависимость k от E. На рис. III. 25, *а*, *б* показаны кривые зависимости P от E для температур, бо́льших и меньших θ.

Если при $T<\theta$ неполяризованный (точка O на рис. III.25, *б*) диэлектрик поместить в электрическое поле с медленно возрастающей напряженностью E, то вызванная этим полем поляризация возрастает по кривой OA. При ослаблении поля E поляризация убывает не по кривой AO, а по кривой AB_1, причем в точке B_1 внешнее поле отсутствует, но диэлектрик поляризован (эта остаточная поляризация сегнетоэлектриков аналогична остаточной намагниченности железа). При изменении направления внешнего поля вектор поляризации убывает до нуля в точке C_1 и затем, изменив свое направление, возрастает по величине вдоль кривой C_1D. Дальнейшее изменение внешнего поля от E_D до E_A сопровождается изменением вектора поляризации по кривой DB_2C_2A; в точке B_2 наблюдается остаточная поляризация. В точках C_1 и C_2 поляризация отсутствует, несмотря на наличие внешнего поля. Площадь $AB_1C_1DB_2C_2A$ в масштабе чертежа равна энергии, которая затрачивается внешним электрическим полем за один цикл переполяризации данного диэлектрика. Совершив такой цикл, диэлек-

трик возвращается в свое первоначальное электрическое состояние, а энергия переполяризации передается его атомам и молекулам в виде теплоты. Если диэлектрик помещен в переменное (например, синусоидальное) электрическое поле, то он будет нагреваться до тех пор, пока возрастающая теплоотдача в окружающую среду не приостановит дальнейшее повышение температуры. Таким образом, поляризованный диэлектрик содержит некоторое количество дополнительной энергии, которое зависит от его массы и интенсивности поляризации; эта энергия может быть выделена при деполяризации.

Сегнетоэлектрики имеют очень большие значения относительной диэлектрической проницаемости (доходящей до десятков тысяч), поэтому конденсаторы с такими диэлектриками имеют при очень малых размерах большие электроемкости. Такие конденсаторы применяются, например, в электронно-счетных машинах как «ячейки памяти». В некоторых приборах используется сильная зависимость диэлектрической проницаемости этих веществ от напряженности поля. Емкость конденсаторов с сегнетоэлектрическим диэлектриком зависит от приложенного к нему напряжения. Такие конденсаторы с нелинейной характеристикой, названные *варикондами*, применяются в радиотехнике в различных устройствах: в стабилизаторах напряжений, в диэлектрических усилителях и т. д.

ЭЛЕКТРЕТЫ

Существуют диэлектрики, которые, получив при изготовлении определенную поляризацию, в дальнейшем сохраняют ее так же, как постоянные магниты сохраняют свою намагниченность. Такие диэлектрики с постоянной поляризацией называются электретами. Пластинка из электрета имеет на одной стороне положительный, а на другой — равный отрицательный заряд и создает вокруг себя постоянное электрическое поле. Первоначальную поляризацию электретов получают различными способами; например, диэлектрик нагревают до определенной температуры (при которой облегчается та или иная форма поляризации — поворот диполей, смещение ионов и др.), помещают в сильное электрическое поле и в этом поле охлаждают. В твердом состоянии поляризация оказывается «замороженной», т. е. тепловое движение не в состоянии разрушить полученную веществом поляризацию. Другие диэлектрики (сера, сульфиды цинка и кадмия) помещают в сильное электрическое поле и подвергают действию света в течение некоторого времени, после чего убирается поле и устраняется освещение. Полученные таким образом «фотоэлектреты» сохраняют остаточную поляризацию только в темноте. Некоторые электреты получаются под действием излучения радиоактивных веществ; наконец, некоторым веществам можно сообщить остаточную поляризацию под действием одного только сильного электрического поля. Применение электретов основано главным образом на возбуждении переменного тока в проводниках, совершающих колебания в электрическом поле электрета (электретные микрофоны и телефоны, измерители вибраций и др.).

§ 7. ЭЛЕКТРОЕМКОСТЬ. ЭНЕРГИЯ ЗАРЯЖЕННОГО ПРОВОДНИКА. ПЛОТНОСТЬ ЭНЕРГИИ ПОЛЯ

Рассмотрим сначала уединенный проводник, находящийся достаточно далеко от других тел. Если этому проводнику сообщить заряды $q_1, q_2, \ldots$, то после их перераспределения по объему проводника он приобретает потенциалы $\varphi_1, \varphi_2, \ldots$ Отношение $q_1/\varphi_1, q_2/\varphi_2, \ldots$ для данного уединенного проводника оказывается постоянным, зависящим только от его формы и размеров, и называется его электроемкостью. Это отношение сохраняется и при бесконечно малых изменениях заряда и потенциала, так что

$$C = \frac{q}{\varphi} \quad \text{или} \quad C = \frac{dq}{d\varphi}. \tag{1.58}$$

Понятие электроемкости применимо только к проводникам, так как для них существует равновесное распределение зарядов по объему тела, при котором все точки проводника имеют один и тот же потенциал. Если же заряд сообщается изолятору, то он не растекается по нему и поэтому в различных местах изолятора потенциал может быть различен (в зависимости от расстояний до того места, где находится подведенный заряд).

Емкость уединенного шара радиуса r, находящегося в безграничном диэлектрике с проницаемостью ε, легко рассчитать, так как потенциал на его поверхности (а следовательно, и в любой точке его объема) $\varphi = kq/\varepsilon r$:

$$C = \frac{q}{\varphi} = \frac{\varepsilon}{k} r. \tag{1.59}$$

В системе СГС $k = 1$ и $C = \varepsilon r$; в СИ $k = 1/4\pi\varepsilon_0$ и $C = 4\pi\varepsilon_0\varepsilon r$.

При наличии вблизи данного проводника других тел — проводников или изоляторов — отношение (1.58) зависит также от формы, размеров и относительного расположения соседних тел. Если эти соседние тела — проводники, то в них происходит перераспределение свободных зарядов, электрическое поле которых накладывается на поле данного тела и изменяет его потенциал. Если же соседние тела — диэлектрики, то они поляризуются, вследствие чего на поле данного тела накладывается поле связанных зарядов диэлектрика; это опять-таки изменяет потенциал рассматриваемого проводника.

Таким образом, при наличии соседних тел данный проводник при сообщении ему заряда q приобретает иной потенциал, чем при их отсутствии.

Понятие электроемкости можно применять и к системе проводников; простейшей из них является система из двух одинаковых близко расположенных проводников, которым сообщаются равные и противоположные по знаку заряды. В частности, рассмотрим *плоский конденсатор,* состоящий из двух близко расположенных параллельных металлических пластинок (обкладок); при сообщении обкладкам конденсатора зарядов $+q$ и $-q$ они приобретают потенциалы φ_1 и φ_2. Электроемкостью конденсатора называется отношение заряда q на одной из его обкладок (по абсолютному значению, без учета знака) к раз-

ности потенциалов между обкладками:

$$C = q/(\varphi_1 - \varphi_2). \tag{1.60}$$

Допустим, что расстояние d между обкладками настолько мало, что электрическое поле между ними можно считать однородным; напряженность этого поля, согласно формуле (1.36),

$$E = k\frac{4\pi\sigma}{\varepsilon} = k\frac{4\pi q}{\varepsilon S},$$

где S — площадь обкладок; $\sigma = q/S$ — поверхностная плотность зарядов на обкладках. Для однородного поля выполняется соотношение (1.45), поэтому

$$\varphi_1 - \varphi_2 = Ed = k\frac{4\pi q}{\varepsilon S}d.$$

Подставив это выражение в формулу (1.60), получаем формулу для расчета емкости плоского (двухпластинчатого) *конденсатора*:

$$C = \frac{\varepsilon}{k}\frac{S}{4\pi d}. \tag{1.61}$$

У *шарового конденсатора* потенциалы φ_1 и φ_2 на обкладках определяются зарядами $\pm q$, которые имеются на этих обкладках, и их радиусами r_1 и r_2:

$$\varphi_1 = k\frac{q}{\varepsilon r_1};\quad \varphi_2 = k\frac{q}{\varepsilon r_2},$$

поэтому формула для расчета емкости такого конденсатора имеет вид

$$C = \frac{q}{\varphi_1 - \varphi_2} = \frac{\varepsilon}{k}\frac{1}{1/r_1 - 1/r_2} = \frac{\varepsilon}{k}\frac{r_1 r_2}{d}, \tag{1.62}$$

где $d = r_2 - r_1$ — величина зазора между обкладками. Если радиусы обкладок очень велики и d мало, то можно положить $r_1 r_2 = r^2$, $4\pi r^2 = S$ (площадь обкладок) и тогда полученная формула будет совпадать с (1.61).

У *цилиндрического конденсатора* определяется емкость, приходящаяся на единицу длины. Выведем сначала формулу для разности потенциалов между обкладками; согласно формулам (1.32), (1.13) и (1.39), имеем:

$$E = \frac{D}{\varepsilon_0\varepsilon} = \frac{\tau}{2\pi\varepsilon_0\varepsilon r};\quad \varphi_1 - \varphi_2 = \int_{r_1}^{r_2} E\,dr = \frac{\tau}{2\pi\varepsilon_0\varepsilon}\ln\frac{r_2}{r_1}.$$

(Интегрирование ведем вдоль перпендикуляра к оси конденсатора, т. е. вдоль направления силовой линии вектора $\mathbf{E}$. У очень длинного цилиндрического конденсатора вектор напряженности поля в зазоре перпендикулярен оси конденсатора: это условие не соблюдается на концах, но этим обстоятельством для достаточно длинных конденсаторов можно пренебречь.) Так как на единице длины каждой обкладки имеется заряд τ, то «погонная» емкость цилиндрического конденсатора будет равна

$$C = \frac{\tau}{\varphi_1 - \varphi_2} = \frac{2\pi\varepsilon_0\varepsilon}{\ln r_2/r_1} = \frac{\varepsilon}{2k\ln r_2/r_1}. \tag{1.63}$$

Если величина зазора $r_2 - r_1 = d$ очень мала, то $\ln(r_2/r_1) = \ln(1 + d/r_1) \approx d/r_1$. По этой формуле рассчитывается емкость электрического кабеля, состоящего из внутреннего провода и наружной металлической брони, между которыми находится слой диэлектрика.

В электротехнике приходится рассчитывать емкость *двухпроводной линии* — системы из двух параллельных проводов (обычно круглого сечения). Обозначим ра-

диусы сечений этих проводов через $r_1 = r_2 = r$, расстояние между осями проводов — через a и допустим, что $a \gg r$. В этом случае поле вокруг каждого провода можно с удовлетворительным приближением рассчитывать по формуле (1.34). Допустим, что на единице длины одного провода находится заряд $(+\tau)$, а другого — $(-\tau)$. В некоторой точке, расположенной на расстоянии x от оси первого провода, суммарная напряженность поля будет равна

$$E = E_+ + E_- = k\frac{2\tau}{\varepsilon x} + k\frac{2\tau}{\varepsilon(a-x)}.$$

Интегрируя Edx вдоль перпендикуляра, соединяющего оси проводников, получим разность потенциалов между проводами:

$$\varphi_1 - \varphi_2 = \int_r^{a-r} E\,dx = \frac{2k\tau}{\varepsilon}\Big|\ln x - \ln(a-x)\Big|_r^{a-r} = \frac{4k\tau}{\varepsilon}\ln\frac{a-r}{r}.$$

Следовательно, погонная емкость двухпроводной линии будет равна

$$C = \frac{\tau}{\varphi_1 - \varphi_2} = \frac{\varepsilon}{4k}\frac{1}{\ln[(a-r)/r]}.$$

Так как было предположено, что расстояние между проводами значительно больше радиуса их сечений, то

$$C \approx \frac{\varepsilon}{4k}\frac{1}{\ln(a/r)}. \qquad (1.64)$$

В приведенных выше расчетных формулах для электроемкости при использовании системы СГС следует положить $k = 1$, а в Международной системе (СИ) $k = 1/4\pi\varepsilon_0$. В частности, для плоского конденсатора:

1) в системе СГС $C = \frac{\varepsilon S}{4\pi d}$;

2) в СИ $C = \varepsilon_0\varepsilon\frac{S}{d}$.

Электроемкость выражается в *фарадах* (Ф) : 1 Ф = 1 Кл/В. В системе СГС единицей электроемкости является *сантиметр*:

$$[C] = \frac{[q]}{[\varphi]} = \frac{\text{г}^{1/2}\cdot\text{см}^{3/2}\cdot\text{с}^{-1}}{\text{г}^{1/2}\cdot\text{см}^{1/2}\cdot\text{с}^{-1}} = \text{см}.$$

Так как 1 Кл $= 3\cdot10^9$ ед. СГС заряда, а 1 В $= 1/300$ ед. СГС потенциала, то 1 Ф $= 9\cdot10^{11}$ см.

Рассмотрим параллельное (рис. III.26, *а*) и последовательное (рис. III.26, *б*) соединения конденсаторов. Если к точкам A и B *параллельно* соединенных конденсаторов подвести равные и противоположные заряды q, то они распределятся между обкладками конденсаторов так, что $q = q_1 + q_2 + q_3 + \ldots$ Разность же потенциалов между обкладками всех конденсаторов будет одна и та же (так как они соединены вместе проводниками); обозначим $\varphi_A - \varphi_B$ через φ. Емкостью такой системы конденсаторов называется отношение

$$C = \frac{q}{\varphi_A - \varphi_B} = \frac{q_1}{\varphi} + \frac{q_2}{\varphi} + \frac{q_3}{\varphi} + \ldots$$

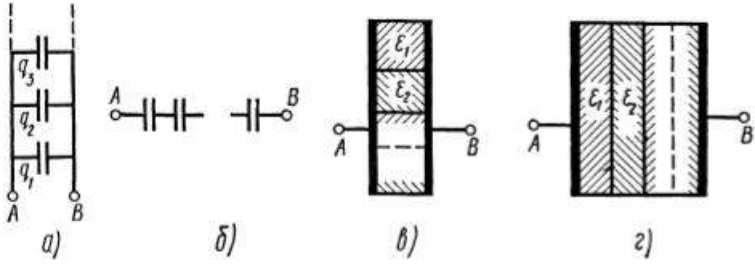
Однако отношение $q_1/\varphi = C_1$ — емкость первого конденсатора, $q_2/\varphi = C_2$ — емкость второго и т. д. Следовательно,

$$C = C_1 + C_2 + C_3 + \ldots \qquad (1.65)$$

Можно показать, что обычный многопластинчатый плоский конденсатор с числом пластин n представляет собой параллельное соединение $n-1$ плоских двухпластинчатых конденсаторов, поэтому

$$C = \frac{\varepsilon}{k}\frac{S}{4\pi d}(n-1). \qquad (1.66)$$

Если к точкам A и B последовательно соединенных конденсаторов подвести заряды $\pm q$, то вследствие электростатической индукции на обкладках конденсаторов появятся равные и противоположные по знаку заряды $q_1 = q_2 = q_3 = \ldots = q$. При этом пластинки соседних конденсаторов, соединенные между собой проводником, имеют одинаковый потенциал. Так как разность потенциалов на концах любой



а) б) в) г)

Рис. III.26

линии равна сумме разностей потенциалов на отдельных участках этой линии, то для линии AB, проходящей через электрические поля соединенных конденсаторов, можно написать:

$$\varphi_A - \varphi_B = (\varphi_1 - \varphi_2) + (\varphi_2 - \varphi_3) + (\varphi_3 - \varphi_4) + \ldots$$

Емкостью этой системы конденсаторов по-прежнему называется отношение

$$C = q/(\varphi_A - \varphi_B), \text{ откуда } \varphi_A - \varphi_B = q/C.$$

Так как для первого конденсатора $\varphi_1 - \varphi_2 = q/C_1$, для второго $\varphi_2 - \varphi_3 = q/C_2$ и т. д., то

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \ldots \qquad (1.67)$$

Заметим интересную деталь: если между обкладками плоского конденсатора поместить несколько металлических пластинок, расположенных параллельно обкладкам (т. е. вдоль эквипотенциальных поверхностей), и если суммарный зазор между ними $d_1 + d_2 + \ldots$ равен первоначальному зазору d, то емкость конденсатора не изменится. Действительно, такой конденсатор можно рассматривать как систему последовательно соединенных плоских конденсаторов, поэтому, применив формулу (1.64) и (1.67), получим

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} + \ldots = \frac{d_1}{\varepsilon_0 \varepsilon S} + \frac{d_2}{\varepsilon_0 \varepsilon S} + \ldots = \frac{d}{\varepsilon_0 \varepsilon S},$$

т. е. первоначальная емкость конденсатора $C = \varepsilon_0 \varepsilon S/d$ не изменилась. В частности, емкость конденсатора не изменится, если вдоль эквипотенциальных поверхностей поместить металлические пластинки бесконечно малой толщины.

Если между обкладками плоского конденсатора имеются различные диэлектрики, как это показано на рис. III.26, *в*, *г*, то для расчета емкости такого конденсатора можно воспользоваться формулами (1.65) и (1.67). Конденсатор (рис. III.26, *в*) можно представить как систему из параллельно соединенных конденсаторов, имеющих одинаковые расстояния d между пластинами, но различные S и ε, и тогда

$$C = \frac{\varepsilon_0}{d}(\varepsilon_1 S_1 + \varepsilon_2 S_2 + \ldots).$$

Конденсатор (рис. III.26, *г*) можно представить как систему последовательно соединенных плоских конденсаторов; так как введение или удаление бесконечно тонких металлических пластинок, параллельных обкладкам, не изменяет емкости конденсатора, то эти пластинки можно расположить вдоль границ между диэлектриками. Тогда, воспользовавшись формулами (1.61) и (1.67), получим ($k = 1/4\pi\varepsilon_0$)

$$C = \frac{\varepsilon_0 S}{d_1/\varepsilon_1 + d_2/\varepsilon_2 + \ldots}.$$

Если $\varepsilon_1 = \varepsilon_2 = \ldots = \varepsilon$, то эта формула перейдет в (1.61).

Для того чтобы сообщить проводнику некоторый заряд q, необходимо затратить определенную работу, так как каждая последующая порция подводимого заряда испытывает отталкивающее действие ранее поступивших на проводник одноименных зарядов. Допустим, что очередная порция dq заряда подводится из бесконечности, где потенциал $\varphi_1 = 0$, к проводнику, имеющему уже потенциал $\varphi_2 = \varphi$. Тогда элементарная работа, затрачиваемая на подведение заряда dq,

$$dA = dq(\varphi_1 - \varphi_2) = -\varphi\, dq.$$

Для определения всей работы, затрачиваемой на сообщение проводнику некоторого заряда q, надо просуммировать все эти элементарные работы:

$$A = -\int_0^q \varphi\, dq = -\int_0^q \frac{q}{C}\, dq = -\frac{q^2}{2C}.$$

Это выражение для работы можно видоизменить, если воспользоваться соотношением $q = C\varphi$.

Знак минус указывает, что для зарядки тела необходимо совершить некоторую внешнюю работу A. При разрядке электрическое поле заряженного тела само совершает работу, равную A. Таким образом, формула

$$W = q^2/2C = q\varphi/2 = C\varphi^2/2 \qquad (1.68)$$

выражает *энергию заряженного проводника*. Аналогичная формула получается и для заряженного конденсатора, причем φ означает разность потенциалов между клеммами конденсатора (или системы конденсаторов).

Энергию плоского конденсатора можно выразить через напряженность поля между обкладками. Подставим $\varphi = Ed$ и $C = \varepsilon S/k4\pi d$ в формулу (1.68), тогда

$$W = \frac{1}{2} C\varphi^2 = \frac{\varepsilon}{k} \frac{E^2}{8\pi} Sd. \tag{1.69}$$

Произведение Sd выражает объем электрического поля между обкладками конденсатора, а отношение W/Sd, которое обозначим через w, есть энергия конденсатора, приходящаяся на единицу объема его электрического поля:

$$w = \frac{1}{k} \frac{\varepsilon E^2}{8\pi}. \tag{1.70}$$

В системе СГС $w = \varepsilon E^2/8\pi$, в СИ $w = \varepsilon_0 \varepsilon E^2/2$.

Полагая, что энергия заряженных тел есть энергия их электрического поля, величину w можно назвать *объемной плотностью энергии электрического поля*.

Заметим, что постоянную ε_0 принято выражать не в Кл²/(Н·м²), а в Ф/м:

$$\text{Кл}^2/(\text{Н} \cdot \text{м}^2) = \text{Кл}^2/(\text{Дж} \cdot \text{м}) = \text{Кл}/(\text{В} \cdot \text{м}) = \text{Ф/м}.$$

Воспользуемся формулой для плотности энергии электрического поля и рассчитаем энергию, которая содержится в электрическом поле заряженного шара радиуса R. В тонком слое, ограниченном сферами радиусов r и $r + dr$, будет содержаться энергия $dW = {}^1/_2 \varepsilon_0 \varepsilon E^2 \cdot 4\pi r^2 dr$. Если заряд шара равен q (и так как напряженность поля на расстоянии r от центра шара равна $E = kq/\varepsilon r^2 = q/4\pi\varepsilon_0 \varepsilon r^2$), то

$$W = \int_R^\infty dW = \frac{q^2}{8\pi\varepsilon_0\varepsilon} \int_R^\infty \frac{dr}{r^2} = \frac{q^2}{8\pi\varepsilon_0\varepsilon} \frac{1}{R} = \frac{k}{\varepsilon} \frac{q^2}{2R}. \tag{1.71}$$

Эта формула показывает, что энергия электрического поля заряженного шара в основном сосредоточена вблизи его поверхности. Действительно, в шаровом слое, ограниченном поверхностью шара и сферой радиуса $r = 2R$, содержится столько же энергии, сколько имеется в остальном пространстве от $r = 2R$ до бесконечности:

$$W_1 = \int_R^{2R} dW = -\frac{q^2}{8\pi\varepsilon_0\varepsilon} \left(\frac{1}{2R} - \frac{1}{R} \right) = \frac{q^2}{8\pi\varepsilon_0\varepsilon} \frac{1}{2R} = \frac{k}{\varepsilon} \frac{q^2}{4R};$$

$$W_2 = \int_{2R}^\infty dW = -\frac{q^2}{8\pi\varepsilon_0\varepsilon} \left(-\frac{1}{2R} \right) = \frac{q^2}{8\pi\varepsilon_0\varepsilon} \frac{1}{2R} = \frac{k}{\varepsilon} \frac{q^2}{4R}.$$

Два заряженных шара при одинаковом заряде q имеют в окружающем их электрическом поле различную энергию; чем меньше радиус шара, тем больше эта энергия. Для точечного заряда энергия электрического поля была бы бесконечно большой.

Глава 2

ЭЛЕКТРИЧЕСКИЙ ТОК

§ 8. ЭЛЕКТРИЧЕСКИЙ ТОК ПРОВОДИМОСТИ; ПЛОТНОСТЬ ТОКА

Электрическим током называется упорядоченное перемещение электрических зарядов одного знака в каком-нибудь направлении. Движение положительных и отрицательных зарядов в противоположных направлениях есть электрический ток одного и того же направления. Условились, что направление тока совпадает с направлением движения положительных зарядов и противоположно направлению движения отрицательных зарядов. Для того чтобы получить односторонний перенос электрических зарядов, необходимо приложить к этим зарядам силы; в частности, если тела, содержащие свободные заряды, поместить в электрическое поле, то на положительные заряды будут действовать силы, направленные вдоль вектора E, на отрицательные — против E. Вследствие этого на тепловое движение зарядов накладывается некоторое направленное перемещение положительных зарядов в одном и отрицательных в обратном направлении. Однако при описании движения электрических зарядов обычно отвлекаются от их беспорядочного теплового движения и интересуются только результирующим переносом этих зарядов в том или ином направлении.

Основным признаком электрического тока является м а г н и т н о е п о л е, существующее вокруг движущихся зарядов. Кроме того, при прохождении электрического тока через вещество наблюдаются тепловые, оптические и химические явления, в которых имеет место превращение электрической энергии в другие виды энергии.

Электрические токи, вызванные в телах электрическими полями, называются *токами проводимости*. Например, в металлических проводниках электрический ток представляет собой упорядоченное перемещение свободных электронов (*электронная проводимость*), вызванное действующим внутри проводника электрическим полем (электроны, которые могут перемещаться в пределах объема проводника, в отличие от электронов, связанных внутри атомов вещества, называются *электронами проводимости*). Положительные ионы металла прочно связаны между собой в кристаллической решетке и в переносе зарядов не участвуют; вследствие этого электрический ток через металлы не сопровождается какими-нибудь изменениями их вещества. Упорядоченное движение электронов в металлических проводниках можно получить и без электрического поля, механическим путем: если быстро затормозить движущееся металлическое тело (например, вращающуюся катушку), свободные электроны некоторое время перемещаются по инерции по направлению движения и создают кратковременный электрический ток.

В некоторых твердых телах — ионных кристаллах типа NaCl, полупроводниках, в стекле и т. д. — электрический ток может быть образован также и упорядоченным движением ионов. С повышением температуры эта *ионная проводимость* твердых тел возрастает. Так,

например, стекло, нагретое до размягчения, обладает хорошей проводимостью. Наблюдается также и *смешанная проводимость*, создаваемая одновременным движением электронов и ионов.

В жидких проводниках (электролитах) электрический ток обусловлен движением положительных и отрицательных ионов в противоположных направлениях. При отсутствии электрического поля эти ионы совершают беспорядочное тепловое движение; при наличии поля электрические силы добавляют к тепловым скоростям ионов некоторые скорости вдоль поля, вследствие чего положительные ионы постепенно перемещаются к катоду, а отрицательные — к аноду.

Кратковременные электрические токи наблюдаются и в диэлектриках. При введении их в электрическое поле происходит поляризация (поворот элементарных электрических диполей); положительные заряды смещаются в одном, отрицательные — в противоположном направлении. Следовательно, пока происходит поляризация диэлектриков, в них имеется упорядоченное движение зарядов, соответствующее некоторому электрическому току; такие токи называются *токами поляризации*. Они прекращаются, когда заканчивается поляризация диэлектрика. Если удалить электрическое поле, вызвавшее поляризацию, то под действием хаотического теплового движения элементарные электрические диполи теряют свою преимущественную ориентировку; эта деполяризация диэлектрика сопровождается также упорядоченным «возвращением» положительных и отрицательных зарядов к исходным беспорядочным ориентировкам, что соответствует току поляризации противоположного направления.

Электрические токи могут быть получены при движении зарядов вместе с какими-нибудь телами, на которых они находятся. Например, движущиеся заряженные проводники или изоляторы создают вокруг себя магнитное поле и эквивалентны некоторому электрическому току; такие токи называются *конвекционными*.

Допустим, что через элементарную площадку S в некотором направлении за время Δt проходят q_1 положительных и q_2 отрицательных зарядов, а в обратном направлении — q_1' положительных и q_2' отрицательных зарядов. Тогда сила тока через площадку S равна

$$I = \frac{(q_1 - q_2) - (q_1' - q_2')}{\Delta t} = \frac{(q_1 - q_1') - (q_2 - q_2')}{\Delta t} = \frac{\Delta q_1}{\Delta t} - \frac{\Delta q_2}{\Delta t}.$$

Обычно интересуются избыточным переносом зарядов через площадку S, т. е. величиной $\Delta q = \Delta q_1 - \Delta q_2$. Тогда

$$I = \frac{\Delta q}{\Delta t} \quad \text{или} \quad I = \lim_{\Delta t \to 0} \frac{\Delta q}{\Delta t} = \frac{dq}{dt}. \tag{2.1}$$

Сила тока выражается в *амперах*: 1 А = 1 Кл/с.

Допустим, что через элементарную площадку S (например, сечение тонкого проводника) проходят только заряды одного знака. Обозначим через n число элементарных заряженных частиц в единице объема проводника, а через v — среднюю скорость их упорядоченного движения в направлении Ox. Тогда за время Δt через площадку S в этом направлении пройдут все частицы, находящиеся в объеме $Sv\Delta t$, т. е.

$nSv\Delta t$ частиц. Если заряд одной частицы равен e, то за время Δt через площадку S пройдет заряд $\Delta q = enSv\Delta t$. Тогда сила тока через площадку будет равна

$$I = \Delta q/\Delta t = nevS.$$

Если вектор площадки S (т. е. нормаль к ней) составляет с направлением тока угол α, то сила тока должна быть представлена в виде

$$I = nev\mathbf{S}. \qquad (2.2)$$

Заметим, что сила тока есть, по определению, скалярная величина; знак скалярного произведения $\mathbf{vS} = vS\cos\alpha$ зависит от угла α между направлением тока (т. е. направлением движения положительных зарядов) и направлением нормали к площадке.

Сила тока, приходящаяся на единицу площадки, ориентированной перпендикулярно направлению тока, называется *плотностью тока*:

$$j = \frac{I}{S} = nev.$$

Плотность тока есть *векторная* величина, ориентированная по направлению тока, т. е. по направлению вектора скорости упорядоченного движения положительных зарядов или против направления движения отрицательных зарядов:

$$\mathbf{j} = ne\mathbf{v} \quad \text{или} \quad \mathbf{j} = (\pm e)\, n\mathbf{v}. \qquad (2.3)$$

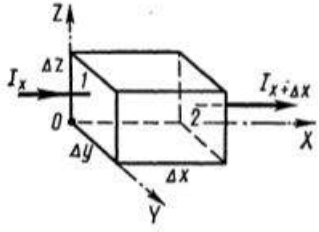


Рис. III.27

Для того чтобы сила тока через данную площадку оставалась постоянной во времени, необходимо, согласно формуле (2.2), сохранить постоянными величинами n и v.

Упомянутый в § 1 *закон сохранения заряда* математически записывается в виде некоторого соотношения между плотностью электрического тока в различных местах замкнутой поверхности, охватывающей рассматриваемую физическую систему, и изменением во времени объемной плотности зарядов ρ внутри этой поверхности. Для упрощения расчетов выберем элементарный объем ΔV в виде (рис. III.27) кубика с гранями Δx, Δy и Δz и допустим, что в направлении оси OX сила тока изменяется: на площадке *1* она равна I_x, а на площадке *2*

$$I_{x+\Delta x} = I_x + \Delta I_x = I_x + \frac{\partial I_x}{\partial x}\Delta x,$$

где $\frac{\partial I_x}{\partial x}$ есть изменение силы тока, приходящееся на единичную длину вдоль оси OX. Если $\frac{\partial I_x}{\partial x} > 0$, т. е. $I_{x+\Delta x} > I_x$, то, согласно закону сохранения электричества, внутри объема ΔV заряд будет уменьшаться (зарядов вносится меньше, чем выносится). За время dt это уменьшение будет равно

$$-\,dq = \Delta I_x\, dt = \frac{\partial I_x}{\partial x}\Delta x\, dt.$$

Разделив dq на объем ΔV, получим изменение плотности заряда ρ:

$$-\,d\rho = -\frac{dq}{\Delta V} = \frac{1}{\Delta x\,\Delta y\,\Delta z}\frac{\partial I_x}{\partial x}\Delta x\, dt.$$

Вместо силы тока введем в эту формулу плотность тока $j_x = I_x/\Delta S = I_x/\Delta y \Delta z$. Тогда получим простое соотношение

$$-\frac{d\rho}{dt} = \frac{\partial j_x}{\partial x}.$$

Если одновременно существуют токи и в направлениях OY и OZ, то, очевидно,

$$-\frac{d\rho}{dt} = \frac{\partial j_x}{\partial x} + \frac{\partial j_y}{\partial y} + \frac{\partial j_z}{\partial z}, \tag{2.4}$$

где j_x, j_y и j_z — проекции вектора плотности тока на координатные оси (начало этого вектора можно поместить либо в точке O, либо в центре рассматриваемого объема ΔV; результат расчета не изменится).

Величина

$$\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z},$$

где A — любой вектор, называется *дивергенцией* этого вектора в данной точке (составляющие этого вектора, как было выше принято для плотности тока, полагаются, вообще говоря, функциями координат). Дивергенция вектора показывает, как быстро изменяются компоненты этого вектора вдоль положительных направлений осей координат. Пользуясь этим понятием, можно сформулировать закон сохранения электрического заряда в виде соотношения

$$\operatorname{div} \mathbf{j} = -\frac{d\rho}{dt},$$

т. е. скорость изменения плотности заряда ρ в бесконечно малом объеме проводника зависит от того, как резко изменяется плотность тока j вдоль координатных осей в этом месте. В общем случае ρ может быть функцией не только времени, но и координат: $\rho = \rho(x, y, z, t)$; тогда в согласии с приведенными выше рассуждениями следует вместо полной производной $\frac{d\rho}{dt}$ написать частную производную $\frac{\partial \rho}{\partial t}$:

$$\operatorname{div} \mathbf{j} = -\frac{\partial \rho}{\partial t}. \tag{2.5}$$

§ 9. ЭЛЕКТРОННАЯ ТЕОРИЯ ПРОВОДИМОСТИ МЕТАЛЛОВ. ЗАКОНЫ ОМА И ДЖОУЛЯ—ЛЕНЦА В ДИФФЕРЕНЦИАЛЬНОЙ ФОРМЕ. СВЯЗЬ МЕЖДУ ЭЛЕКТРОПРОВОДНОСТЬЮ И ТЕПЛОПРОВОДНОСТЬЮ МЕТАЛЛОВ

Рассмотрим электрический ток в металлических проводниках, внутри которых существует электрическое поле с напряженностью E. Это поле действует на электроны проводимости с силой $F = eE$, сообщающей им ускорение $a = F/m = eE/m$, где m — масса электрона. Если бы движение электронов в металле происходило без потерь энергии, то их скорость, а следовательно, и сила тока в проводнике со временем увеличивались бы. Однако при столкновениях с атомами вещества, совершающими беспорядочное тепловое движение, электроны теряют часть своей кинетической энергии.

При постоянном токе, когда средняя скорость упорядоченного движения электронов остается со временем неизменной, вся энергия, получаемая электронами под действием электрического поля, должна быть передана атомам вещества, т. е. перейти в энергию их теплового движения.

Для простоты рассуждений предположим, что при каждом столкновении электрон полностью теряет ту энергию, которую он получил под действием силы $F = eE$ за время свободного пробега τ от одного столкновения до другого. Это означает, что в начале каждого свободного пробега электрон имеет только скорость своего теплового движения u, а в конце пробега, перед столкновением, его скорость под действием силы $F = eE$ увеличивается на некоторую величину v_1. Отвлекаясь от теплового движения, можно полагать, что перемещение электрона в направлении действия силы $F = eE$ является равноускоренным с начальной скоростью $v_0 = 0$. За время свободного пробега электрон приобретает скорость упорядоченного движения $v_1 = a\tau = (eE/m)\,\tau$, а средняя скорость этого движения

$$v = \frac{v_0 + v_1}{2} = \frac{1}{2}\frac{e}{m}E\tau.$$

Время свободного пробега определяется средней скоростью беспорядочного (теплового) движения электрона u и средней длиной свободного пробега λ: $\tau = \lambda/u$. Тогда плотность тока в металлическом проводнике

$$j = nev = \frac{1}{2}\frac{ne^2}{m}\frac{\lambda}{u}E.$$

Перепишем эту формулу в виде

$$j = \sigma E. \tag{2.6}$$

Величина

$$\sigma = \frac{1}{2}\frac{ne^2}{m}\frac{\lambda}{u} \tag{2.7}$$

характеризует свойства проводника и называется его *удельной электрической проводимостью* (электропроводностью).

Энергию W, приобретаемую одним электроном за время t под действием силы $F = eE$, можно рассчитать, умножив $mv_1^2/2$ на число свободных пробегов за время t, т. е. на t/τ:

$$W_1 = \frac{mv_1^2}{2}\frac{t}{\tau} = \frac{m}{2}\left(\frac{eE}{m}\tau\right)^2\frac{t}{\tau} = \frac{e^2}{2m}\tau E^2 t.$$

При постоянном токе вся эта энергия передается веществу проводника и переходит в тепловое движение его атомов. Рассчитаем энергию, которая в единицу времени передается единице объема проводника, содержащего n свободных электронов:

$$W = \frac{nW_1}{t} = \frac{1}{2}\frac{ne^2}{m}\tau E^2 = \sigma E^2, \tag{2.8}$$

или, так как σE равна плотности тока, то

$$W = \sigma E^2 = jE. \tag{2.9}$$

Формула (2.6) $j = \sigma E$ выражает закон Ома, а формула (2.9) — закон Джоуля — Ленца в дифференциальной форме. Приведенный выше вывод этих законов сделан на основе так называе-

мой *классической электронной теории проводимости металлов*. Эта теория утверждает, что в металлах имеется некоторое количество свободных электронов («электронный газ»), которые перемещаются в пределах объема проводника, но не могут покинуть этот объем. Для электронного газа поверхность металла представляет собой как бы «стенки сосуда». Физически это означает, что на поверхности металла для электронов существует особый *потенциальный барьер*, т. е. при переходе через поверхностный слой на электроны действуют значительные силы, направленные внутрь объема металла. Для того чтобы преодолеть этот «барьер», электроны должны обладать достаточной кинетической энергией.

Условия, при которых электроны могут преодолеть потенциальный барьер в поверхностном слое металла и покинуть его объем, рассматриваются в § 12 и 13.

Согласно выражению (2.7) для τ, электропроводность металлов прямо пропорциональна плотности электронного газа, т. е. числу n свободных электронов в единице объема металла. Однако от этого числа зависит и теплопроводность металла, так как, участвуя в тепловом движении, свободные электроны, весьма подвижные и обладающие большими скоростями, способствуют переносу тепла от одних (нагретых) мест металла к другим (холодным).

Измерения показали, что отношение коэффициента теплопроводности λ к электропроводности σ почти одинаково для всех металлов и пропорционально первой степени абсолютной температуры:

$$\lambda/\sigma = \text{const} \cdot T \qquad (2.10)$$

(закон Видемана — Франца).

Удельное сопротивление металлических проводников зависит от их внутренней структуры и от температуры. Чистый металл должен характеризоваться удельным сопротивлением, найденным для монокристалла этого вещества в различных направлениях (монокристалл является электрически анизотропным). Обычные технические металлы не являются химически чистыми и, кроме того, имеют поликристаллическую структуру, в которой монокристаллы весьма малых размеров ориентированы беспорядочно, и поэтому вещество оказывается изотропным. Удельное сопротивление сильно зависит от наличия неоднородностей в структуре металла. В монокристаллах химически чистых металлов неоднородность создается тепловым движением атомов в узлах кристаллической решетки; это движение совершенно беспорядочное, поэтому в объеме кристалла временно возникают и исчезают области с большей или меньшей концентрацией атомов (в единице объема). Наличие таких флуктуаций плотности увеличивает число столкновений электрона с атомами и между собой; это уменьшает длину свободного пробега электронов, вследствие чего, согласно формуле (2.7), уменьшается удельная проводимость. Кроме теплового движения неоднородности в кристаллической решетке данного металла возникают и при наличии примесей. Известно, что удельные проводимости некоторых чистых металлов (например, серебра) уменьшаются в несколько раз при добавлении небольшого количества других металлов.

Представление о том, что «электронный газ» в металлах хотя бы в первом приближении можно рассматривать как идеальный газ, заключенный в объем данного металлического тела, оказалось неприемлемым. Свойства электронного газа внутри металлов существенно отличаются от свойств обыкновенных газов и рассматриваются в квантовой теории металлов.

§ 10. РАБОТА ЭЛЕКТРИЧЕСКОГО ТОКА. СОПРОТИВЛЕНИЕ ПРОВОДНИКОВ. СВЕРХПРОВОДИМОСТЬ

Допустим, что в неоднородном проводнике с переменным сечением (рис. III.28) *1* и *2* изображают эквипотенциальные поверхности с потенциалами φ_1 и φ_2. При переносе заряда из первого сечения во второе электрические силы, действующие внутри проводника, совершают работу $A = q(\varphi_1 - \varphi_2)$. Эта работа при постоянном токе, как указывалось выше, идет не на увеличение кинетической энергии упорядоченного движения электронов, а выделяется в проводнике в виде тепла. Если сила тока с течением времени увеличивается, то часть работы электрических сил идет на увеличение скорости упорядоченного движения электронов, а остальная часть выделяется в виде тепла. При переменном токе работу электрических сил следует рассчитывать, разбивая время наблюдения t на элементарные отрезки dt, на протяжении которых силу тока и разность потенциалов можно полагать постоянными. Тогда за время dt через участок *1—2* пройдет электрический заряд $dq = I dt$, выделится энергия $dW = (\varphi_1 - \varphi_2)\, dq = I(\varphi_1 - \varphi_2)\, dt$. Обозначив $U = \varphi_1 - \varphi_2$, можем рассчитать энергию, выделяющуюся на участке *1—2* за время t, по формуле

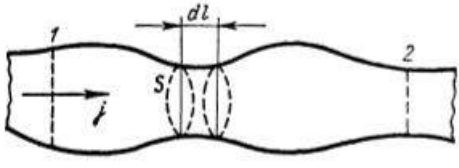


Рис. III.28

$$W = \int_0^t IU\, dt.$$

Сила тока выражается в амперах, разность потенциалов — в вольтах, время — в секундах, а энергия — в джоулях: 1 Дж = 1А·1В × × 1с = 1Кл·1В.

Энергию W можно выразить в зависимости от размеров и вещества проводника на рассматриваемом участке *1—2*. Допустим, на элементарном участке dl проводник однороден и имеет постоянное сечение S (рис. III.28). Кроме того, в пределах объема $S dl$ электрическое поле будем считать однородным, имеющим везде одинаковую напряженность E. Сила тока через сечение S, согласно формуле (2.6), равна $I = jS = \sigma ES$, откуда

$$E = \frac{I}{\sigma S}.$$

Умножим обе части этого равенства на dl и проинтегрируем для интересующего нас участка проводника между эквипотенциальными сечениями *1* и *2*:

$$\int_1^2 E\,dl = I\int_1^2 \frac{dl}{\sigma S}$$

(при постоянном токе сила тока одинакова для любого сечения проводника). Левый интеграл есть, по определению, разность потенциалов $\varphi_1 - \varphi_2$; правый интеграл зависит от свойств проводника (электропроводность σ) и его конфигурации. Обозначим этот интеграл через

$$R = \int_1^2 \frac{dl}{\sigma S}. \tag{2.11}$$

Это есть *электрическое сопротивление* проводника на участке *1—2*. Тогда предыдущее выражение перепишется в виде

$$\varphi_1 - \varphi_2 = IR; \quad I = (\varphi_1 - \varphi_2)/R = U/R. \tag{2.12}$$

Эта формула выражает *закон Ома для участка цепи*.

Пользуясь им, можно записать работу электрического тока в зависимости от сопротивления проводника:

$$W = qU = IUt = I^2Rt = \frac{U^2}{R}t, \tag{2.13}$$

или при переменном токе

$$W = \int_0^t I^2R\,dt = \int_0^t \frac{U^2}{R}\,dt. \tag{2.14}$$

Энергия W, выделяющаяся в проводнике в виде тепла,

$$Q = I^2Rt. \tag{2.15}$$

Эта формула выражает *закон Джоуля—Ленца* в обычной форме.

Электрическое сопротивление R однородного проводника с постоянным сечением зависит от его длины l и площади сечения S:

$$R = \int_1^2 \frac{dl}{\sigma S} = \frac{l}{\sigma S}.$$

Если длина и сечение проводника равны единице, то $R = 1/\sigma$. Величина

$$1/\sigma = \rho$$

есть *удельное электрическое сопротивление* вещества проводника. Для неоднородного проводника переменного сечения электрическое сопротивление необходимо рассчитывать по формуле (2.11) или же по приближенной формуле

$$R = \sum \rho_i \frac{\Delta l_i}{S_i}. \tag{2.16}$$

Сопротивление проводника, на концах которого при силе тока в один ампер существует разность потенциалов в один вольт, называется *омом*: 1 Ом = 1 В/А.

Удельное, а следовательно, и полное электрическое сопротивление проводников зависят от температуры. Эта зависимость имеет сложный вид. Для металлов можно пользоваться приближенными формулами:

$$\rho = \rho_0 (1 + \alpha t); \quad R = R_0 (1 + \alpha t), \tag{2.17}$$

где ρ_0 и R_0 относятся к нулевой температуре по шкале Цельсия, а α — *температурный коэффициент сопротивления*. Этот коэффициент можно считать постоянным только для небольших интервалов температур. При точных расчетах необходимо учитывать зависимость α от температуры.

Закон Ома, т. е. прямая пропорциональность между напряжением и силой тока (см. формулу (2.12)), имеет место для различных значений U и I только при условии $R = \text{const}$. Если же через проводник течет переменный ток и выделяющееся джоулево тепло не отводится так, чтобы обеспечить постоянство температуры проводника, то сопротивление проводника будет изменяться со временем в зависимости от того, как изменяется сила тока. Вследствие этого сопротивление проводника является функцией от силы тока: $R = R(I)$. Для каждого момента времени t можно рассчитать две величины:

$$R = \left(\frac{U}{I}\right)_t \quad \text{и} \quad R = \left(\frac{dU}{dI}\right)_t,$$

которые могут отличаться друг от друга в зависимости от вида функции $I = I(t)$ и от условий, в которых находится проводник. Если R есть сопротивление какого-нибудь сложного прибора, то функции $R(I)$ или $R(U)$ характеризуют электрические свойства этого прибора. Однако более удобными являются кривые $I = I(U)$, изображающие зависимость тока от приложенного напряжения; эти кривые называются «вольт-амперными характеристиками» прибора.

При очень низких температурах, близких к абсолютному нулю (около 0,5—8 К), сопротивление некоторых металлов скачком уменьшается практически до нуля. Например, алюминий при температуре 1,4 К теряет электрическое сопротивление. Состояние металла с нулевым электрическим сопротивлением называется *сверхпроводящим*, а само исчезновение сопротивления — *сверхпроводимостью*. Вследствие отсутствия сопротивления в сверхпроводниках можно вызвать очень большие токи (до 1200 А на 1 мм^2) без выделения теплоты. Если в замкнутой цепи из сверхпроводников вызвать электрический ток (например, при помощи электромагнитной индукции), то этот ток ввиду отсутствия потерь может существовать очень долго.

§ 11. ЭЛЕКТРОДВИЖУЩАЯ СИЛА; ЗАКОН ОМА И ПРАВИЛА КИРХГОФА

Допустим, что имеются два проводника A и B, заряженные разноименными зарядами до потенциалов φ_1 и φ_2. Если эти тела соединить проводником l, то в течение некоторого времени будет происходить перемещение электронов от A к B до тех пор, пока во всей

системе проводников не установится одинаковый потенциал; при этом напряженность электрического поля внутри проводников будет везде равна нулю. Сила тока в этом проводнике сначала возрастает от нуля (в момент соединения) до некоторого максимума, а затем постепенно убывает до нуля. Таким образом, система проводников, где действуют только электростатические силы, со временем переходит в равновесное состояние с одинаковым потенциалом, при котором напряженность поля внутри проводников всюду становится равной нулю, и упорядоченное движение зарядов прекращается. Для того чтобы поддержать в проводнике l направленное движение электронов в течение некоторого времени, необходимо иметь «устройство», которое «перекачивало» бы электроны обратно из тела B к телу A. Такой прибор, называемый *источником тока,* должен действовать на электроны (или на заряды вообще) особыми силами неэлектростатического происхождения. Силы, действующие на заряды со стороны источников тока и направленные против электростатических сил, называются *сторонними силами.* Сторонние силы возбуждаются в генераторах электрического тока при движении проводников в магнитном поле, в гальванических элементах и аккумуляторах (благодаря химическим реакциям), в термоэлементах и т. д.

ЗАКОН ОМА ДЛЯ ЗАМКНУТОГО КОНТУРА

Рассмотрим простейшую систему проводников, содержащую источник тока (рис. III.29). Допустим, что в приборе R, потребляющем электрическую энергию, необходимо поддержать определенную силу тока I, причем электроны должны двигаться в направлении, указанном стрелками. Очевидно, что при переносе через R электронов с общим зарядом, равным $-q$, электрические силы, действующие на электроны в направлении $2 \to R \to 1$, будут совершать положительную работу, которая, согласно формуле (1.42), зависит только от потенциалов начальной φ_2 и конечной φ_1 точек траектории переноса и равна

$$A_{21} = -q(\varphi_2 - \varphi_1) = q(\varphi_1 - \varphi_2). \quad (2.18)$$

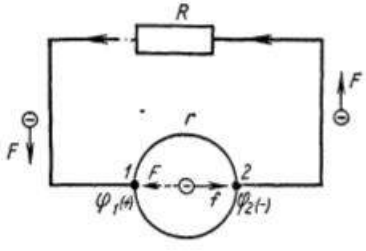


Рис. III.29

Для того чтобы поддержать потенциалы φ_1 и φ_2 постоянными, источник тока должен непрерывно перебрасывать электроны обратно от точки *1* к точке *2*. При этом необходимо преодолеть притяжение электронов к положительно заряженной точке *1* и отталкивание от отрицательно заряженной точки *2*, т. е. преодолевать электростатическую силу F, направленную внутри источника от точки *2* к точке *1*. Таким образом, источник тока должен приложить к электронам стороннюю силу f, направленную против электростатической силы F.

Далее следует иметь в виду, что движение электронов внутри источника тока происходит при наличии некоторого сопротивления, обу-

словленного столкновениями между электронами и атомами источника тока. При этих столкновениях теряется часть кинетической энергии упорядоченного движения электронов и поэтому, чтобы сохранить постоянной скорость этого движения, источник тока должен компенсировать указанную выше потерю энергии внутри самого источника.

Полная работа A, совершаемая сторонними силами внутри источника тока при переносе заряда q из точки *1* в точку *2*, равна сумме: 1) работы A_{12} против электростатических сил F, действующих внутри источника тока, и 2) потери энергии электронов W при их прохождении через источник тока:

$$A = A_{12} + W. \tag{2.19}$$

Это соотношение выражает *закон сохранения энергии*. Очевидно, что работа A_{12} сторонней силы равна работе A_{21}, совершаемой электростатическими силами вне источника тока. Это означает, что источник тока является также источником той энергии или работы, которая выделяется движущимися зарядами во внешнем участке цепи *1* — — R — *2*. Для того чтобы поддержать потенциалы φ_1 и φ_2 постоянными, источник тока должен непрерывно совершать работу A_{12}, компенсирующую потерю энергии во внешней цепи A_{21}.

Для оценки потери энергии электронов W при их перемещении внутри самого источника тока необходимо знать его электрическое сопротивление r; тогда, согласно формуле (2.13),

$$W = I^2rt = qIr. \tag{2.20}$$

Полная работа сторонних сил на основании закона сохранения энергии (см. формулу (2.19))

$$A = q(\varphi_1 - \varphi_2) + qIr. \tag{2.21}$$

Отношение работы, совершаемой сторонними силами внутри источника тока при перемещении через него заряда q к величине этого заряда, называется *электродвижущей силой* (э. д. с.) этого источника тока и обозначается $\mathscr{E}$:

$$\mathscr{E} = \frac{A}{q}; \quad \mathscr{E} = \varphi_1 - \varphi_2 + Ir. \tag{2.22}$$

На основании закона Ома для участка цепи *1* — R — *2*

$$\varphi_1 - \varphi_2 = IR. \tag{2.23}$$

Тогда

$$\mathscr{E} = IR + Ir; \quad I = \mathscr{E}/(R + r). \tag{2.24}$$

Эта формула выражает *закон Ома для замкнутого контура*, по которому течет постоянный ток. Называя IR падением напряжения во внешних участках цепи, а Ir — падением напряжения внутри источника тока, можно закон Ома выразить иначе:

электродвижущая сила, действующая в замкнутой цепи, равна сумме падений напряжения в этой цепи.

Ежесекундная работа, совершаемая источником тока, т. е. его мощность,

$$N = A/t = \mathscr{E}q/t = \mathscr{E}I; \quad \mathscr{E}I = I^2R + I^2r.$$

Эта работа равна той энергии, которая ежесекундно выделяется на всех сопротивлениях цепи.

Если источник тока не замкнут, то упорядоченное движение зарядов через него не происходит и потеря энергии внутри источника тока отсутствует. Сторонняя сила может только вызвать скопление зарядов на полюсах источника тока. Это скопление прекратится, когда внутри источника между его полюсами появится электрическое поле $\mathbf{E}$, в котором электростатическая сила $\mathbf{F} = q\mathbf{E}$ сделается равной сторонней силе, т. е. $\mathbf{F} = -\mathbf{f}$. Разность потенциалов между полюсами разомкнутого источника тока можно рассчитать по формуле (1.39):

$$\varphi_1^0 - \varphi_2^0 = \int_1^2 \mathbf{E}d\mathbf{l},$$

причем интегрирование можно произвести вдоль любой линии, соединяющей полюсы источника тока. Подставим $\mathbf{E} = \mathbf{F}/q_0$ (пробный заряд q_0, как обычно, положим положительным) и заменим $\mathbf{F}$ на $-\mathbf{f}$:

$$\varphi_1^0 - \varphi_2^0 = \frac{1}{q_0}\int_1^2 \mathbf{F}d\mathbf{l} = -\frac{1}{q_0}\int_1^2 \mathbf{f}d\mathbf{l} = \frac{1}{q_0}\int_2^1 \mathbf{f}d\mathbf{l}.$$

Однако $\int_2^1 \mathbf{f}d\mathbf{l}$ есть работа A, совершаемая сторонними силами против электростатических сил при переносе заряда q_0 из точки *2* в точку *1*; тогда, согласно указанному выше определению, э. д. с.

$$\mathscr{E} = A/q_0 = \varphi_1^0 - \varphi_2^0.$$

Таким образом, *электродвижущая сила источника тока равна разности потенциалов на его полюсах в разомкнутом состоянии*. Если же источник тока замкнуть на внешнюю цепь, то, согласно формуле (2.22), разность потенциалов между его полюсами будет меньше э. д. с. на величину падения напряжения Ir внутри самого источника:

$$\varphi_1 - \varphi_2 = \mathscr{E} - Ir.$$

Допустим, в электрическом контуре (рис. III.30) имеются два источника тока, которые могут быть включены так, что сторонние силы в них действуют либо в одном (*а*), либо в противоположных (*б*) направлениях. В первом случае (*а*) сторонние силы в обоих источниках действуют в направлении движении зарядов и совершают положительные работы A_1 и A_2. Общая работа этих сил $A = A_1 + A_2$, и тогда действующая в контуре э. д. с.

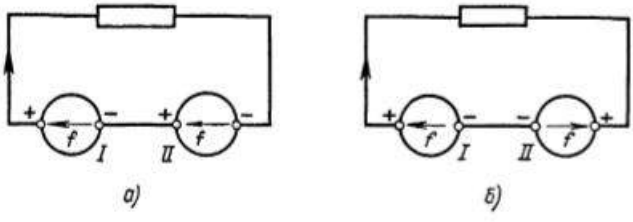
$$\mathscr{E} = \frac{A}{q} = \frac{A_1}{q} + \frac{A_2}{q} = \mathscr{E}_1 + \mathscr{E}_2. \qquad (2.25)$$

Энергия, выделяющаяся в контуре, равна сумме работ, совершаемых обоими источниками.

Во втором случае (*б*) у источника *I* сторонние силы действуют в направлении движения зарядов и совершают положительную работу; у источника *II* сторонние силы направлены против движения зарядов и совершают отрицательную работу. Суммарная работа сторонних сил в контуре (*б*) $A = A_1 - A_2$, и общая э. д. с. в контуре

$$\mathscr{E} = \frac{A}{q} = \frac{A_1}{q} - \frac{A_2}{q} = \mathscr{E}_1 - \mathscr{E}_2. \qquad (2.26)$$

Энергия, выделяющаяся в контуре, в данном случае равна разности работ, совершаемых источниками тока.



а) *б)*

Рис. III.30

Заметим, что электродвижущая сила, по определению, является *скалярной* величиной; ее берут с положительным или отрицательным знаком в зависимости от знака работы, совершаемой сторонними силами.

Выделим в электрической цепи участок, содержащий источник э. д. с. (рис. III.31). По определению, разность потенциалов между точками *1* и *2* равна отношению работы A_{12} электрических сил по переносу заряда q из точки *1* в точку *2* к величине переносимого заряда: $\varphi_1 - \varphi_2 = A_{12}/q$. В данном случае работа переноса слагается из работы $A' = I^2Rt = qIR$ на сопротивлении участка (в виде тепла) и работы A'', совершаемой электростатическими силами внутри источника э. д. с. против сторонних сил. Эта работа равна и противоположна по знаку работе сторонних сил A; последняя положительная, если сторонние силы действуют в направлении движения зарядов, и отрицательна, если заряды движутся в обратном направлении. Таким образом, $A_{12} = qIR \pm A'' = qIR \pm A$, и так как $A/q = \mathscr{E}$, то

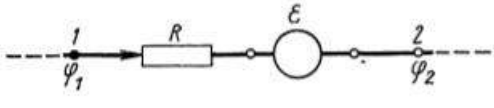


Рис. III.31

$$\varphi_1 - \varphi_2 = (qIR \pm A)/q = IR \pm \mathscr{E},$$

откуда получаем *закон Ома для участка цепи, содержащего источник э. д. с.*:

$$I = (\varphi_1 - \varphi_2 \pm \mathscr{E})/R, \qquad (2.27)$$

где R содержит в себе и сопротивление самого источника тока.

Знак плюс у работы сторонних сил, а следовательно, и у э. д. с. соответствует случаю, когда положительный полюс источника обращен к точке *2*, т. е. ток входит в отрицательный и выходит из положительного полюса источника. При обратном включении берется знак минус.

ПРАВИЛА КИРХГОФА

Для расчета сил токов в различных участках сложных разветвленных цепей (например, мостика Уитстона, рис. III.32) по заданным сопротивлениям этих участков и э. д. с. источников тока пользуются правилами Кирхгофа. Эти правила можно вывести на основе закона сохранения заряда и законов Ома (см. формулы (2.24) и (2.12)). Предполагается, что токи в цепи — установившиеся, т. е. силы токов, сопротивления и разности потенциалов в различных участках цепи с течением времени не изменяются. Окружим какую-нибудь точку разветвления, например A, замкнутой поверхностью (сферой); тогда для постоянства потенциала точки A необходимо, чтобы количество электричества, ежесекундно вносимое в эту сферу через одни проводники, было равно количеству электричества, выходящего из этой области по другим проводникам за то же время:

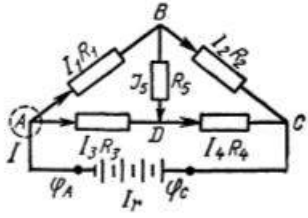


Рис. III.32

алгебраическая сумма сил токов должна равняться нулю (первое правило Кирхгофа):

$$\sum I_i = 0. \qquad (2.28)$$

Это утверждение, очевидно, применимо для любой точки цепи с постоянными токами; если в цепи имеется n точек разветвления, то, применяя условие (2.28), можно написать n уравнений. В частности, в мостике (см. рис. III.32) имеется четыре точки разветвления, однако число неизвестных величин — сил токов — шесть. Поэтому для нахождения этих величин необходимо составить еще два уравнения.

Для использования законов Ома выберем в цепи какую-нибудь замкнутую совокупность проводников, например $ABCDA$, в которой нет источника тока; напишем для нее очевидное тождество:

$$\varphi_A - \varphi_B + \varphi_B - \varphi_C + \varphi_C - \varphi_D + \varphi_D - \varphi_A = 0, \qquad (2.29)$$

где φ_A, φ_B, ... потенциалы начальных и конечных точек отдельных участков контура. На основании закона Ома, для каждого из участков цепи $\varphi_1 - \varphi_2 = IR$. Однако следует учесть, что в этой записи предполагается движение положительных зарядов от первой точки ко второй, т. е. от большего потенциала φ_1 к меньшему φ_2. Если же мы будем вычитать из меньшего потенциала больший, то сила тока получится отрицательной, т. е. $\varphi_2 - \varphi_1 = -IR$. Учитывая это обстоятельство,

заменим разности потенциалов в выражении (2.29) через произведения сил токов на соответствующие сопротивления. Так как $\varphi_A > \varphi_B$, $\varphi_B > \varphi_C$, $\varphi_A > \varphi_D$ и $\varphi_D > \varphi_C$, то

$$\varphi_A - \varphi_B = +I_1R_1; \quad \varphi_C - \varphi_D = -I_4R_4;$$
$$\varphi_B - \varphi_C = +I_2R_2; \quad \varphi_D - \varphi_A = -I_3R_3.$$

Тогда выражение (2.29) перепишется в виде

$$I_1R_1 + I_2R_2 - I_4R_4 - I_3R_3 = 0,$$

т. е. алгебраическая сумма падений напряжения в замкнутом контуре, не содержащем источника тока, равна нулю:

$$\sum I_iR_i = 0.$$

Рассмотрим теперь замкнутый контур, содержащий источник тока, например $ABDCA$. Напишем опять очевидное тождество:

$$\varphi_A - \varphi_B + \varphi_B - \varphi_D + \varphi_D - \varphi_C + \varphi_C - \varphi_A = 0. \qquad (2.30)$$

Согласно изложенному выше $\varphi_A - \varphi_B = +I_1R_1$; $\varphi_D - \varphi_C = +I_4R_4$. Допустим, $\varphi_B < \varphi_D$; тогда $\varphi_B - \varphi_D = -I_5R_5$. Для определения разности потенциалов между полюсами источника тока $\varphi_C - \varphi_A$ воспользуемся определением электродвижущей силы [см. формулу (2.22)], откуда

$$\varphi_1 - \varphi_2 = \mathscr{E} - Ir.$$

Здесь предполагается, что внутри источника положительные заряды движутся от точки с меньшим потенциалом (φ_2) к точке с большим потенциалом (φ_1). Так как в нашем случае $\varphi_A > \varphi_C$, то

$$\varphi_C - \varphi_A = -(\mathscr{E} - Ir) = -\mathscr{E} + Ir.$$

Подставим в выражение (2.30) найденные значения разностей потенциалов:

$$I_1R_1 - I_5R_5 + I_4R_4 - \mathscr{E} + Ir = 0; \quad \sum I_iR_i = \mathscr{E}.$$

Полученный результат означает, что в замкнутом контуре, содержащем источник тока, алгебраическая сумма падений напряжений равна э. д. с. источника тока.

Применяя те же рассуждения для более общего случая, когда в контуре имеется несколько источников тока и они по-разному включены, можно получить следующий результат:

алгебраическая сумма падений напряжения в замкнутом контуре равна алгебраической сумме э. д. с. источников тока, находящихся в этом контуре (второе правило Кирхгофа).

При применении этого правила обычно выбирается какое-нибудь направление обхода и соблюдается следующее условие знаков:

1) если токи I_i текут по направлению обхода, то соответствующие произведения I_iR_i берут со знаком плюс, в противном случае — со знаком минус;

2) если линия обхода направлена внутри источника тока от отрицательного полюса к положительному, то его э. д. с. берется со знаком плюс, в противном случае — со знаком минус.

При расчете сложных цепей чаще всего задаются только сопротивления отдельных участков и э. д. с. источников тока; направления и силы токов требуется определить. Для этого предварительно намечаются направления токов и решается система уравнений:

$$\sum I_i = 0; \quad \sum I_i R_i = \sum \mathscr{E}_i. \tag{2.31}$$

Если после расчета силы токов в каких-нибудь участках получаются отрицательными, то это означает, что направление токов в этих участках обратно намеченному до расчета.

ЭЛЕКТРОСТАТИЧЕСКИЙ ГЕНЕРАТОР ТОКА

Электрический ток можно получить также и в контуре, содержащем участок из диэлектрика. Рассмотрим цепь, состоящую из источника постоянного напряжения U и плоского конденсатора C, в котором диэлектрик можно вдвигать и выдвигать (рис. III.33, *а*). Изменение емкости конденсатора вызовет (при постоянном напряжении U) изменение зарядов на обкладках ($q = CU$), а следовательно, и ток в цепи. Сила этого тока будет равна

$$I = \frac{\mathrm{d}q}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}(CU) = U\frac{\mathrm{d}C}{\mathrm{d}t}.$$

При большой скорости изменения C ток I может достигнуть значительной величины. Направление тока зависит от того, увеличивается или уменьшается емкость конденсатора. Если этот процесс производить периодически, то в цепи будет возбужден переменный ток.

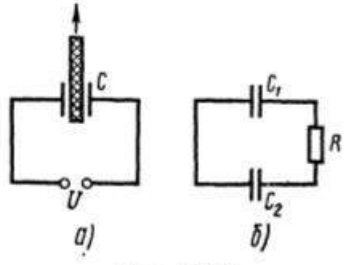


Рис. III.33

Однако такой «электростатический» способ возбуждения тока возможен также и в цепи, не содержащей источника напряжения (рис. III. 33, *б*). Допустим, что первоначально оба конденсатора C_1 и C_2 были заряжены; очевидно, что заряды на их обкладках $q_1 = C_1U_1$ и $q_2 = C_2U_2$ в равновесном состоянии (когда в цепи нет тока и $U_1 = U_2$) будут пропорциональны емкостям: $q_1/q_2 = C_1/C_2$.

Если увеличивать емкость одного конденсатора (можно при этом уменьшать емкость второго конденсатора), то, согласно этому соотношению, величины q_1 и q_2 будут изменяться и заряды будут перемещаться от одного конденсатора к другому. Возникающий в цепи ток будет пропорционален скорости изменения емкостей C_1 и C_2; если эти изменения производить периодически (то в одном, то в противоположном направлении), то в цепи будет протекать переменный ток. Источником энергии, которая будет выделяться на сопротивлении R (или на любом потребителе электрической энергии), является внешняя работа, совершаемая при увеличении емкости конденсатора (при уменьшении этой емкости источником энергии является электрическое поле между обкладками конденсатора). Таким образом, мы получаем «электростатический генератор электрического тока», в котором механическая работа может быть превращена в электрическую энергию.

§ 12. РАБОТА ВЫХОДА ЭЛЕКТРОНОВ ИЗ МЕТАЛЛА. КОНТАКТНАЯ РАЗНОСТЬ ПОТЕНЦИАЛОВ. ТЕРМОЭЛЕКТРИЧЕСКИЕ ЯВЛЕНИЯ

Выше упоминалось, что для «электронного газа» роль «стенок сосуда» играет поверхностный слой металла: силы, действующие на электроны в этом слое, направлены внутрь металла. На рис. III.34 схематически изображена ионная решетка металла, составленная

из прочно связанных между собой положительных ионов. Пока электрон находится внутри металла, силы, действующие на него со стороны положительных ионов, направлены во все стороны и в среднем взаимно уравновешены. Если же электрон оказался за пределами металла, то на него со стороны положительных ионов действуют силы, направленные только внутрь металла. Однако необходимо учесть также и силы отталкивания между данным электроном и остальными электронами, находящимися внутри металла и в пограничном слое.

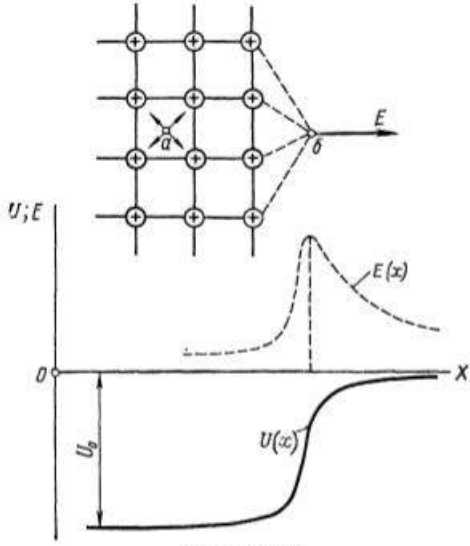


Рис. III.34

Во-первых, заметим, что на электрон *а*, оказавшийся вблизи поверхности (рис. III.34), равнодействующая сил притяжения к ионам не равна нулю и направлена внутрь металла, так как с той стороны ионов больше. Вследствие этого концентрация электронов в пограничном слое металла несколько меньше, чем в глубине. Во-вторых, электроны *б*, вылетевшие из металла, своим отталкивающим действием также будут уменьшать эту концентрацию. При таких условиях силы отталкивания, действующие на рассматриваемый электрон *б*, со стороны других электронов, находящихся в поверхностном слое металла, не будут уравновешиваться силами притяжения со стороны положительных ионов. Это означает, что вблизи поверхностного слоя металла существует электрическое поле, вектор E которого (равный нулю внутри металла) в этом слое возрастает до некоторого значения E_0 и затем убывает до нуля на больших расстояниях от тела (рис. III.34). Это также означает, что в поверхностном слое металла существует *потенциальный барьер*, график которого $U = U(x)$ в направлении, перпендикулярном к поверхности, изображен на рис. III.34; при этом потенциальная энергия электрона на больших расстояниях от металла равна нулю, а внутри металла — отрицательной величине, равной U_0.

ЭМИССИЯ ЭЛЕКТРОНОВ ИЗ МЕТАЛЛА

Для того чтобы электрон мог преодолеть силы, притягивающие его к ионной решетке металла, т. е. пройти через потенциальный барьер в поверхностном слое и удалиться из металла, необходимо затратить некоторую энергию. Максимальная кинетическая энергия, которую может иметь электрон внутри металла, недостаточна для этого. Поэтому для преодоления потенциального барьера U_0 к электрону необходимо приложить внешние силы или же каким-нибудь образом

сообщить ему дополнительное количество энергии. Работа, которая должна быть совершена на освобождение электрона из металла, называется работой выхода и является одной из важных характеристик металла; она связана с электрическим полем в поверхностном слое металла:

$$A = \int_0^\infty eE(x)\,dx; \quad \int_0^\infty E(x)\,dx = \frac{A}{e}, \tag{2.32}$$

где интегрирование (вдоль любой траектории) должно производиться от какой-нибудь точки, взятой внутри металла, до точек, расположенных достаточно далеко от его поверхности. Для чистой поверхности вольфрама эта работа равна 4,5 эВ. У других (чистых) металлов она колеблется в пределах 1,8—5,3 эВ. Если поверхностный слой металла содержит какие-нибудь примеси, то работа выхода уменьшается; например, покрытие поверхности вольфрама тонким слоем цезия уменьшает работу выхода до 1,36 эВ.

В равновесном состоянии металла некоторое количество электронов, участвующих в беспорядочном тепловом движении, ежесекундно выходит за пределы поверхности металла, но затем, под действием указанных выше сил, вновь втягиваются внутрь металла. Эти электроны образуют вблизи поверхности металла так называемое *электронное облако,* толщина и плотность которого (число электронов в единице объема) увеличиваются с повышением температуры.

Эмиссию («испарение») электронов с поверхности металла можно получить следующими способами:

1) нагреть металл до очень высокой температуры и тем увеличить число электронов, приобретающих при тепловом движении большие скорости. Такие электроны, обладая большой кинетической энергией, могут преодолеть силы, препятствующие их освобождению из металла (*термоэлектронная эмиссия*);

2) воспользоваться сильным электрическим полем, которое «подхватывало» бы электроны из поверхности металла. Такая эмиссия электронов называется *холодной*, или *автоэлектронной*; она может быть вызвана и при низких температурах;

3) произвести облучение поверхности металла световыми, ультрафиолетовыми, рентгеновскими и другими лучами, энергия которых поглощается электронами. Электроны, вылетающие из металла, затрачивают часть полученной энергии на работу выхода, а остальную часть сохраняют в виде кинетической энергии. Такой способ эмиссии электронов называется *фотоэлектрическим эффектом* (*фотоэлектронная эмиссия*);

4) произвести бомбардировку поверхности металла электронами, ионами или другими частицами. Если эта бомбардировка производится электронами, то эмиссия обусловлена тем, что число вылетающих электронов больше, чем число бомбардирующих электронов (для чистых поверхностей — в 1,2 — 1,8 раза, а для поверхностей, содержащих примеси и покрытых тонким слоем окислов, — в десятки раз); такой способ вырывания называется *вторичной электронной эмиссией.*

Если поверхность одного металла (*1*) привести в соприкосновение (контакт) с поверхностью другого металла (*2*), то происходит переход электронов из одного металла в другой, вследствие чего один из них заряжается положительно, другой — отрицательно. Возникающая при этом разность потенциалов между соприкасающимися телами называется контактной разностью потенциалов.

Появление контактной разности потенциалов обусловлено двумя причинами:

1) различием в работах выхода (A) электрона из металлов, приведенных в соприкосновение. В этом случае силы, действующие на электроны в пограничной области со стороны ионных каркасов обоих металлов (направленные в противоположные стороны), не уравновешены и поэтому вызывают переход электронов из одного металла в другой;

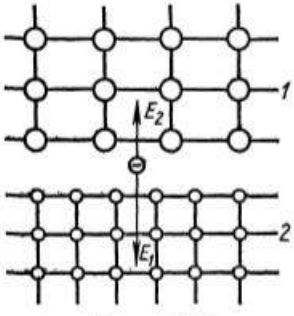


Рис. III.35

2) различием в плотностях (n) электронного газа в металлах, вследствие чего возникает диффузный переход электронов из металла, где плотность этого газа большая, в металл, где эта плотность меньше.

Рассмотрим действие каждой из этих причин в отдельности. На электрон, оказавшийся в пограничной области (между ионными решетками обоих металлов), действуют электрические поля $\mathbf{E}_1$ и $\mathbf{E}_2$, созданные поверхностными слоями каждого из металлов (рис. III.35) и направленные в противоположные стороны. Сумма векторов $\mathbf{E}_1$ и $\mathbf{E}_2$ лишь в частном случае может оказаться равной нулю в какой-нибудь точке пограничного слоя. При сложной зависимости этих векторов от расстояния (см. рис. III.34) их сумма на всем протяжении пограничного слоя будет отлична от нуля и поэтому на электрон, оказавшийся в этой области, действует некоторая результирующая сила, направленная в сторону одного из металлов.

На рис. III.35 показаны ионные решетки двух металлов *1* и *2* с различными объемными плотностями ионов. Очевидно, на электроны, оказавшиеся в пограничной области, со стороны металла *1* действуют меньшие силы, чем со стороны металла *2*, имеющего более плотную ионную решетку; работа выхода электрона из металла *1* (A_1) будет меньше, чем из металла *2* (A_2). Если $n_1 = n_2$, то вследствие теплового движения оба металла посылают в пограничный слой (в единицу времени с единицы поверхности) одинаковое число электронов N. Из этих $2N$ электронов большая часть втягивается в металл *2*. Таким образом, из пограничного слоя металл *1* получает меньше электронов, чем посылает сам, а металл *2* — больше; первый заряжается положительно, второй — отрицательно. Это вызовет появление в пограничной области внешнего электрического поля E, направленного от положительно заряженного металла к отрицательно заряженному; оно будет

препятствовать указанному выше переходу электронов от одного металла к другому. Следовательно, со временем должно наступить равновесное состояние, при котором числа электронов, проходящие через пограничную область в противоположных направлениях, сделаются равными. Очевидно, в этом состоянии на электроны в пограничной области не должны действовать односторонне направленные силы, т. е. сумма всех векторов $\mathbf{E}_1$, $\mathbf{E}_2$ и $\mathbf{E}$ должна быть равна нулю: $\mathbf{E}_1 + \mathbf{E}_2 + \mathbf{E} = 0$. Так как $\mathbf{E}_1$ и $\mathbf{E}_2$ имеют различные направления, то

$$E = E_1 - E_2.$$

Умножим это равенство на элементарное перемещение электрона dx и проинтегрируем вдоль какого-нибудь пути перехода от одного металла к другому. Интеграл $\int_1^2 E\,dx$ будет равен разности потенциалов обоих металлов; это есть, очевидно, контактная разность потенциалов; она обусловлена разностью работ выхода, поэтому ее обозначим через $(\varphi_1 - \varphi_2)_A$. Интегралы $\int_1^2 E_1\,dx$ и $\int_1^2 E_2\,dx$ будут, согласно формуле (2.32), равны отношениям соответствующих работ выхода A_1 и A_2 к заряду электрона. Следовательно,

$$(\varphi_1 - \varphi_2)_A = -(A_1 - A_2)/e,$$

знак «минус» учитывает направление векторов напряженности.

Рассмотрим теперь контактную разность потенциалов, которая вызывается различием плотностей электронного газа в металлах. Допустим, что у соприкасающихся металлов работы выхода равны $A_1 = A_2$, но числа свободных электронов в единице объема отличаются $n_1 \neq n_2$. Для упрощения рассуждений предположим, что в любой точке пограничной области электрические поля $\mathbf{E}_1$ и $\mathbf{E}_2$, созданные поверхностными слоями металлов, взаимно уравновешиваются, т. е. везде $\mathbf{E}_1 = -\mathbf{E}_2$ и поэтому электроны могут свободно переходить из одного металла в другой. При $n_1 \neq n_2$ каждый из металлов будет ежесекундно посылать в пограничную область различные количества электронов N_1 и N_2. Так как в этой области односторонне действующих сил пока нет, то из $N_1 + N_2$ электронов, ежесекундно поступающих в эту область, половина будет втянута в один металл, а другая половина — во второй металл. Таким образом, из первого металла ежесекундно выходят N_1 электронов, а поступают $1/2\,(N_1 + N_2)$ электронов; разница между этими числами вызовет недостаток или избыток электронов в данном металле, т. е. положительную или отрицательную зарядку его. Второй металл приобретает противоположный заряд. Вследствие этого появится электрическое поле $\Delta\mathbf{E}$, направленное от положительно заряженного металла к отрицательному, которое со временем приостановит диффузию электронов из одного металла в другой.

Таким образом, при $n_1 \neq n_2$ в пограничной области между металлами возникает контактное электрическое поле $\Delta\mathbf{E}$, которое препят-

ствует выравниванию плотностей электронного газа в металлах. Внутри этой области плотность электронного газа должна изменяться от n_1 в одном металле, до n_2 — в другом. По аналогии с изменением плотности газа в поле тяготения Земли (см. барометрическую формулу) можно написать

$$n_2 = n_1 e^{-\Delta A/kT},$$

где ΔA (которая в барометрической формуле равняется работе подъема одной молекулы газа на высоту H) в данном случае есть работа переноса электрона через пограничную область, т. е. равна $e(\varphi_1 - \varphi_2)$, где φ_1 и φ_2 — потенциалы металлов после наступления равновесного состояния. Тогда

$$\Delta A = e(\varphi_1 - \varphi_2)_n = kT \ln \frac{n_1}{n_2}; \quad (\varphi_1 - \varphi_2)_n = \frac{kT}{e} \ln \frac{n_1}{n_2}$$

(k — постоянная Больцмана, T — абсолютная температура металлов).

Если существуют обе причины, вызывающие контактную разность потенциалов, то

$$\varphi_1 - \varphi_2 = -\frac{A_1 - A_2}{e} + \frac{kT}{e} \ln \frac{n_1}{n_2}. \qquad (2.33)$$

Электрическое поле $\mathbf{E}$, которое создается между металлами вследствие скопления в них избыточных зарядов, будет складываться с внутренними электрическими полями $\mathbf{E}_1$ и $\mathbf{E}_2$, которые существуют в поверхностном слое каждого металла. Однако в равновесном состоянии сумма $\mathbf{E} + \mathbf{E}_1 + \mathbf{E}_2$ не должна равняться нулю, так как при соблюдении этого равенства возможен односторонний переход электронов из металла, где плотность электронного газа велика, в металл, где эта плотность мала. Поэтому равновесие может наступить тогда, когда сумма векторов $\mathbf{E} + \mathbf{E}_1 + \mathbf{E}_2$ дает некоторую результирующую $\Delta\mathbf{E}$, достаточную для приостановки диффузии электронов вследствие $n_1 \neq n_2$. Таким образом, в пограничной области между металлами при равновесии существует только контактное электрическое поле $\Delta\mathbf{E}$, обусловленное вторым членом формулы (2.33); разность потенциалов, соответствующая этому полю,

$$(\varphi_1 - \varphi_2)_n = \int_1^2 \Delta E \, dx = \frac{kT}{e} \ln \frac{n_1}{n_2}.$$

Измерения показывают, что для чистых металлов контактные разности потенциалов имеют значения от десятков долей до нескольких вольт, причем значительная часть этой разности потенциалов обусловлена разностью работ выхода.

Контактная разность потенциалов может быть получена при соприкосновении не только химически различных проводников, но и между одинаковыми проводниками; для этого необходимо только, чтобы они отличались по какому-нибудь признаку, от которого зависят работы выхода или плотности свободных электронов. Например, разность потенциалов наблюдается при соприкосновении одинаковых проводни-

ков, если они различным образом деформированы (сжаты или растянуты), имеют примеси в различных концентрациях и т. п., а также в том случае, если вдоль данного проводника имеется разность температур.

ЗАКОН ВОЛЬТА. ТЕРМОЭЛЕМЕНТЫ

Допустим, что имеется система из нескольких разнородных проводников, составляющих замкнутую цепь (рис. III.36, *а*). Если проводники — металлы (при их соприкосновении в системе не происходит химических изменений) и если температура в пределах системы везде одинаковая, то соблюдается з а к о н В о л ь т а: в замкнутой цепи из неизменяющихся проводников при одинаковой в цепи температуре контактные разности потенциалов не могут вызывать электрического тока; в разомкнутой цепи (рис. III.36, *б*) разность потенциалов между конечными проводниками определяется только веществом этих проводников и не зависит от вещества промежуточных тел (она будет точно такой же, как и при соприкосновении одних только крайних проводников).

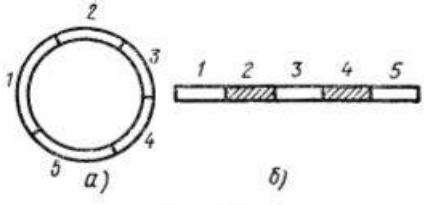


Рис. III.36

Закон Вольта является следствием (или подтверждением) закона сохранения энергии: если бы в замкнутой цепи абсолютно неизменяющихся проводников возникал электрический ток, то можно было бы от такой системы получить неограниченное количество электрической энергии без затраты других видов энергии. Закон Вольта можно получить и из формулы (2.33). Сложим контактные разности потенциалов, возникающие между каждой парой соприкасающихся металлов:

$$\varphi_1 - \varphi_2 = -\frac{A_1 - A_2}{e} + \frac{kT}{e} \ln \frac{n_1}{n_2};$$

$$\varphi_2 - \varphi_3 = -\frac{A_2 - A_3}{e} + \frac{kT}{e} \ln \frac{n_2}{n_3};$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\varphi_{n-1} - \varphi_n = -\frac{A_{n-1} - A_n}{e} + \frac{kT}{e} \ln \frac{n_{n-1}}{n_n}.$$

После сложения и сокращений получаем такой же результат, какой имелся бы при контакте только крайних проводников:

$$\varphi_1 - \varphi_n = -\frac{A_1 - A_n}{e} + \frac{kT}{e} \ln \frac{n_1}{n_n}.$$

В частности, если крайние проводники совершенно одинаковые, то $A_1 = A_n$; $n_1 = n_n$; $\varphi_1 = \varphi_n$. Следовательно, если такую цепь из разнородных металлов замкнуть, то сумма контактных разностей потенциалов будет равна нулю и в такой цепи электрического тока не будет.

В этих рассуждениях существенно *постоянство температуры* в системе. При наличии разностей температур в замкнутой цепи возникает электрический ток. Рассмотрим систему из двух различных металлов *1* и *2*, контакты (спаи) между которыми находятся при различных температурах (рис. III.37). Тогда разность потенциалов $\varphi_1 - \varphi_2$, возникающая в месте контакта с температурой T_1, будет отличаться от разности потенциалов $\varphi_2' - \varphi_1'$ в контакте с температурой T_2. Если для простоты предположить, что работы выхода A_1 и A_2 и плотности электронов n_1 и n_2 не изменяются с температурой, то суммарная разность потенциалов между точками B и C (при разомкнутой цепи) равна:

$$\varphi_B - \varphi_C = (\varphi_1 - \varphi_2) + (\varphi_2' - \varphi_1') = \left(-\frac{A_1 - A_2}{e} + \frac{kT_1}{e}\ln\frac{n_1}{n_2}\right) +$$
$$+ \left(-\frac{A_2 - A_1}{e} + \frac{kT_2}{e}\ln\frac{n_2}{n_1}\right) = (T_1 - T_2)\frac{k}{e}\ln\frac{n_1}{n_2} = \alpha(T_1 - T_2).$$

Эта разность потенциалов является электродвижущей силой; ее называют *термоэлектродвижущей силой*, а систему, изображенную на рис. III.37, — *термоэлементом*. Очевидно, если для данных двух металлов известна постоянная α термоэлектродвижущей силы

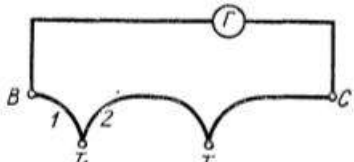


Рис. III.37

$$\mathcal{E} = \alpha(T_1 - T_2), \qquad (2.34)$$

то, измеряя $\mathcal{E}$ и зная температуру одного из спаев, например T_1, можно определить температуру T_2 среды, в которой находится другой спай. Этим пользуются для измерения очень высоких и очень низких температур, а также для измерения температур в тех местах, где нельзя использовать обычные ртутные термометры. Величина α измеряется стотысячными долями вольта на один градус разности температур. Например, для термоэлемента из железа и константана $\alpha = 5{,}2 \cdot 10^{-5}$ В/К; для различных интервалов температур α не остается постоянной. Заметим также, что измерение α позволяет найти отношение плотностей электронного газа в контактирующих металлах, так как

$$\alpha = \frac{k}{e}\ln\frac{n_1}{n_2}.$$

ЭФФЕКТ ПЕЛЬТЬЕ

Если через пограничную область между двумя соприкасающимися различными металлами пропустить электрический ток, то электроны, проходя через эту область, будут в зависимости от направления тока либо ускоряться указанным выше контактным полем $\Delta\mathbf{E}$, либо тормозиться им. В первом случае в пограничном слое наблюдается выделение тепла (электроны, получившие кинетическую энергию, будут при столкновениях передавать энергию атомам металла); во втором случае — поглощение тепла (электроны, потерявшие скорость, будут

при столкновениях с атомами получать от них энергию, т. е. охлаждать металл). Если, например, через термоэлемент (рис. III.37) при $T_1 = T_2$ пропустить электрический ток, то температуры контактов начнут изменяться, так как в одном из них поле $\Delta\mathbf{E}$ совершает положительную работу, а в другом — отрицательную. Выделение или поглощение тепла при прохождении тока через пограничную область, обусловленное работой контактного электрического поля $\Delta\mathbf{E}$, называется эффектом Пельтье; количество тепла за время t при постоянной силе тока I

$$\Delta Q = (\varphi_1 - \varphi_2)_n It = \left(\frac{kT}{e} \ln \frac{n_1}{n_2}\right) It = \alpha T q,$$

где $(\varphi_1 - \varphi_2)_n$ есть, как указывалось выше, разность потенциалов, обусловленная полем $\Delta\mathbf{E}$, $q = It$ — количество прошедшего электричества, а $\alpha = (k/e) \ln (n_1/n_2)$ — постоянная термоэлектродвижущей силы.

Итак, при прохождении электрического тока через пограничную область между двумя металлами происходят следующие явления:

1) выделение некоторого количества тепла по закону Ленца—Джоуля: $Q_1 = I^2Rt$, где R — сопротивление пограничной области. Это тепло не зависит от направления тока и пропорционально квадрату силы тока;

2) выделение или поглощение тепла, вызванное положительной или отрицательной работой контактного электрического поля $\Delta\mathbf{E}$: $Q_2 = \alpha TIt$ (эффект Пельтье). Это тепло пропорционально первой степени силы тока. При малых значениях тока теплота Q_2 может оказаться больше, чем Q_1;

3) перенос энергии из одного металла в другой вместе с перешедшими через пограничную область электронами. Средние кинетические энергии ε электронов в различных металлах (при одинаковой температуре) могут быть различными, поэтому $N = q/e = It/e$ электронов, перешедших из одного металла (ε_1) в другой (ε_2), перенесут с собой энергию

$$Q_3 = N(\varepsilon_1 - \varepsilon_2) = \frac{1}{e}(\varepsilon_1 - \varepsilon_2) It.$$

Заметим, что указанное выделение и поглощение тепла в контактах происходит и в том случае, когда в цепи термоэлемента течет электрический ток, вызванный не посторонним источником тока, а самой термоэлектродвижущей силой (при $T_1 \neq T_2$); при этом в контакте с высокой температурой происходит поглощение тепла, а в контакте с низкой температурой — выделение тепла. Таким образом, эффект Пельтье направлен к выравниванию разности температур в контактах. Если создание разности температур контактов рассматривать как внешнее воздействие на термоэлемент, то эффект Пельтье есть противодействующая ему реакция этого элемента.

Термоэлемент при $T_1 \neq T_2$ можно рассматривать как физическую систему, в которой происходит непосредственное превращение теплоты в электрическую энергию. Допустим, что в цепи термоэлемента сила тока равна I. Работа, совершаемая термоэлектродвижущей силой

(см. формулу (2.34)) за время t,

$$W = \mathscr{E} It = \alpha (T_1 - T_2) q.$$

Однако $\alpha T_1 q = Q_1$ — теплота, поглощенная (по эффекту Пельтье) в контакте с высокой температурой, а $\alpha T_2 q = Q_2$ — теплота, выделяющаяся в контакте с низкой температурой; разность между ними превращается в электрическую энергию. Поэтому коэффициент полезного действия термоэлемента

$$\eta = \frac{Q_1 - Q_2}{Q_1} = \frac{T_1 - T_2}{T_1}.$$

Таким образом, в полном согласии со вторым знаком термодинамики, в термоэлементе происходит получение тепла Q_1 от тела с высокой температурой, передача некоторого количества тепла Q_2 телу с низкой температурой и превращение разности $Q_1 - Q_2$ в другой вид энергии. Однако в случае термоэлемента из металлов значительная часть тепла переходит от горячего контакта к холодному путем теплопроводности, поэтому количество теплоты, превращенное в электрическую энергию, даже при большой разности температур контактов, составляет лишь весьма малую часть (доли процента) полного количества тепла, перешедшего от горячего контакта к холодному. При применении, по предложению А. Ф. Иоффе, полупроводниковых материалов с малой теплопроводностью можно приблизить коэффициент полезного действия термоэлектрической тепловой машины к ее идеальному значению.

Выше рассматривались только металлические проводники (проводники первого рода), у которых появление контактной разности потенциалов и прохождение электрического тока не сопровождается никакими химическими изменениями. Однако контактная разность потенциалов обнаруживается и в системе любых проводников, включая, например, электролиты (проводники второго рода), в которых возбуждение разности потенциалов и прохождение тока сопровождается химическими реакциями (гальванические элементы, аккумуляторы). В отличие от металлических проводников, в системе, содержащей электролиты, на заряды (электроны, ионы) действуют особые силы «химического» происхождения. Благодаря наличию этих сторонних сил в замкнутой системе проводников, содержащих электролиты, происходит непрерывный односторонний перенос зарядов, т. е. существует электрический ток.

§ 13. ТЕРМОЭЛЕКТРОННАЯ ЭМИССИЯ. ЭЛЕКТРИЧЕСКИЙ ТОК В ВАКУУМЕ. ЭЛЕКТРОННАЯ ЛАМПА

Интенсивность термоэлектронной эмиссии, т. е. число электронов, «испаряющихся» с поверхности нагретых металлов в единицу времени, зависит главным образом от температуры металла и работы выхода электронов из него. С повышением температуры увеличивается число электронов, у которых кинетическая энергия теплового движения делается больше работы выхода. Можно увеличить число «испаряющихся» электронов, если уменьшить работу выхода электронов из ме-

талла; это достигается путем специальной обработки поверхности металла, добавлением примесей, изменяющих структуру ионной решетки, или нанесением на поверхность металлов тонких слоев другого вещества, способствующих уменьшению работы выхода.

Термоэлектронная эмиссия используется для получения потоков электронов, направляемых затем на те или иные объекты, подвергаемые бомбардировке электронами. Например, в рентгеновских трубках потоки электронов, которым сообщаются большие скорости, бомбардируют металлы и возбуждают рентгеновское излучение. В электронных микроскопах потоки электронов служат для «освещения» изучаемых объектов; они дают на экране микроскопа «электронное изображение» этого объекта. *Поток электронов в вакууме является разновидностью электрического тока.* Такой электрический ток в вакууме

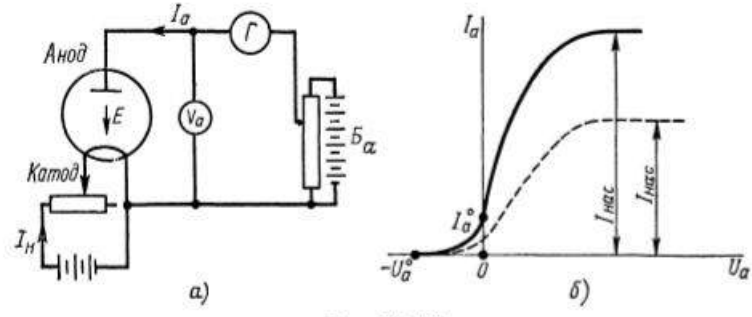


Рис. III.38

можно получить, если в сосуд, откуда тщательно откачивается воздух (до давлений порядка 10^{-6} мм рт. ст.), поместить нагреваемый катод, являющийся источником «испаряющихся» электронов, и анод. Между катодом и анодом создается электрическое поле, сообщающее электронам скорости в определенном направлении.

Рассмотрим схему (рис. III.38, *а*), в которой можно изменять температуру катода (путем подачи различного тока накала $I_н$) и изменять напряженность электрического поля E у поверхности катода (путем изменения разности потенциалов U_a между анодом и катодом). Пока температура катода мала, значительная часть вылетающих из его поверхности электронов возвращается обратно к катоду. К аноду идут лишь те электроны, которые, обладая большой скоростью, удаляются от поверхности катода настолько далеко, что электрическая сила $F = eE$ оказывается больше силы, втягивающей эти электроны обратно в катод. С увеличением тока накала $I_н$, т. е. с повышением температуры катода, число таких электронов увеличивается. Поток электронов от катода к аноду непрерывно поддерживается внешним источником тока, создающим внутри лампы разность потенциалов U_a; измеряя анодный ток I_a, можно обнаружить зависимость тока эмиссии в лампе от температуры катода.

Допустим теперь, что температура катода сохраняется постоянной; изменяется только анодное напряжение U_a и, следовательно,

напряженность поля E у поверхности катода. При малых значениях U_a и E, когда действующая на электроны сила eE мала, лишь незначительная часть электронов может быть оторвана от электронного облака, окружающего катод (электроны, наиболее удалившиеся от катода). По мере увеличения E число электронов, отрываемых от этого электронного облака, увеличивается. Наконец, при очень больших значениях U_a и E почти все электроны, покидающие катод, подтягиваются к аноду. Наибольшая величина термоэлектронного тока, которая может быть получена из катода при данной его температуре, называется *током насыщения*.

На рис. III.38, *б* показана зависимость анодного тока I_a от анодного напряжения U_a при двух температурах катода: высокой (сплошная линия) и низкой (пунктирная линия). При высоких температурах и близком расположении анода и катода возможно существование анодного тока I_a и при $U_a = 0$; это означает, что имеется некоторое количество электронов, энергия теплового движения которых оказывается достаточной для преодоления работы выхода и достижения анода без помощи электрического поля. Для того чтобы приостановить движение этих электронов, необходимо к лампе приложить обратное поле, т. е. подать к аноду отрицательный, а к катоду — положительный потенциал (напряжение U_a^0).

До наступления насыщения зависимость силы тока через лампу от приложенной разности потенциалов между анодом и катодом выражается формулой Богуславского — Лэнгмюра (закон «трех вторых»):

$$I_a = BU_a^{3/2}, \tag{2.35}$$

где B зависит от формы, размеров и относительного расположения катода и анода. Для токов насыщения интересуются не силой тока $I_{нас}$, а плотностью тока насыщения, т. е. $i_{нас} = I_{нас}/S$, где S — поверхность катода, испускающая электроны. Плотность тока насыщения уже не зависит от U_a и определяется по формуле Ричардсона:

$$j_{нас} = BT^2 e^{-A/kT},$$

где B — постоянная величина (теоретически — одинаковая для всех чистых металлов); T — абсолютная температура катода; A — работа выхода электронов из катода; k — постоянная Больцмана.

Для чистого вольфрама $B = 60$ А/(см$^2 \cdot$К^2); работа выхода $A = 4{,}5$ эВ, поэтому при $T = 1000$ К плотность тока насыщения составляет $1{,}3 \cdot 10^{-15}$ А/см^2, а при $T = 3000$ К — около 11 А/см^2. Этот пример иллюстрирует, как резко возрастает термоэмиссионный ток с повышением температуры катода.

В радиотехнике и в устройствах по автоматическому управлению различными процессами широкое применение получили *электронные* или *катодные лампы*, в которых используется электронный поток от нагреваемого катода к аноду. Эти лампы применяются для:

1) выпрямления переменных токов;
2) усиления слабых колебаний токов или потенциалов,
3) генерирования электромагнитных колебаний.

Выпрямление переменных токов осуществляется двухэлектродными лампами (*диоды, кенотроны*). При включении кенотрона в цепь синусоидального тока по схеме, изображенной на рис. III.39, *а*, через прибор R, потребляющий электрическую энергию, проходит пульсирующий ток, график которого показан рядом; пунктиром изображен ток обратного направления, не пропускаемый («отрезанный») кенотроном. Таким образом, через R ток проходит только в одном направлении, но с перерывами через каждые полпериода.

Более удобна схема, изображенная на рис. III.39, *б*. В течение одной половины периода ток течет в направлении AK_1DB, в течение следующей половины периода — в направлении CK_2DB. В течение обоих полупериодов через прибор R, потребляющий электрическую

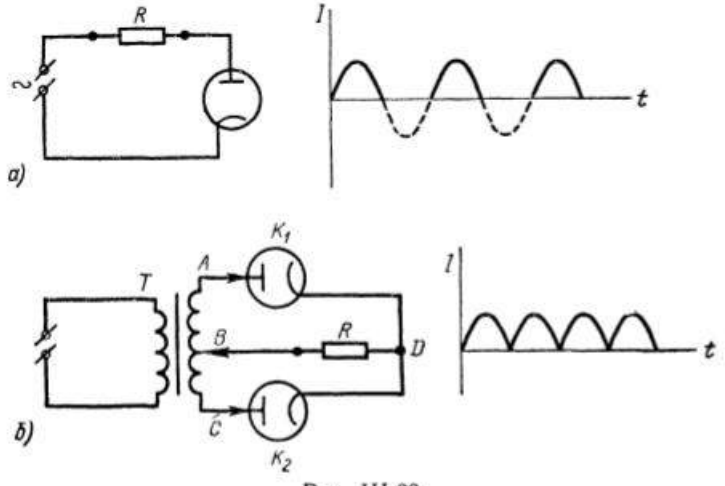


Рис. III.39

энергию, ток проходит в одном и том же направлении. Здесь B есть средняя точка вторичной обмотки трансформатора T; подбирая числа витков этого трансформатора, можно одновременно с выпрямлением производить повышение или понижение напряжения тока.

Усиление слабых колебаний токов и потенциалов производится при помощи электронных ламп, у которых между катодом и анодом помещен дополнительный электрод — *сетка*. Принципиальная схема включения такой лампы — *триода* — в качестве усилителя показана на рис. III.40. Переменное напряжение U_c создает между сеткой и катодом электрическое поле, напряженность которого E_c может быть направлена либо от сетки к катоду, либо же от катода к сетке. В первом случае поле E_c подхватывает электроны, вылетевшие из катода, и сообщает им дополнительные скорости по направлению к аноду. Во втором случае поле E_c отбрасывает электроны обратно к катоду; при больших E_c (или U_c) поток электронов от катода к аноду прекращается совсем: лампа «запирается» сеткой.

Таким образом, в зависимости от величины и знака U_c (т. е. от величины и направления $\mathbf{E}_c$) будет изменяться сила «анодного» тока I_a,

протекающего через лампу. Следует заметить, что при переменном U_c анодный ток I_a является пульсирующим, т. е. течет только в одном направлении и изменяется по величине. Усиление заключается в том, что разность потенциалов на сопротивлении R, равная $I_a R = U_a$, значительно больше разности потенциалов U_c, приложенной между катодом и сеткой.

На рис. III.41 показана зависимость между анодным током I_a и разностью потенциалов между катодом и сеткой U_c, называемая *сеточной характеристикой электронной лампы*; пунктиром показана

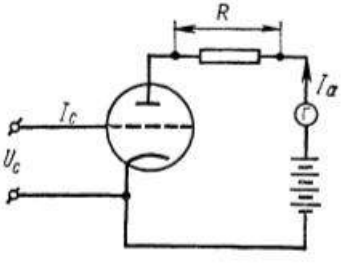


Рис. III.40

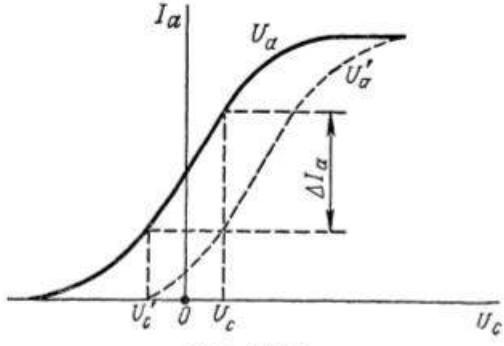


Рис. III.41

та же зависимость для меньшего значения анодного напряжения $U'_a < U_a$. Таким образом, если при неизменном U_c увеличивать анодное напряжение на $\Delta U_a = U_a - U'_a$, то анодный ток изменится на ΔI_a; такое же увеличение анодного тока можно получить, если изменить сеточное напряжение от U'_c до U_c, т. е. на $\Delta U_c = U_c - U'_c$. Отношение

$$\mu_c = \Delta U_a / \Delta U_c$$

называется *статическим коэффициентом усиления лампы*. При этом разность потенциалов на сопротивлении R изменится на величину $\Delta U = R\Delta I_a$. Отношение

$$\mu_д = \Delta U / \Delta U_c$$

называется *динамическим коэффициентом усиления лампы*.

§ 14. ЭЛЕКТРИЧЕСКИЙ ТОК В ГАЗАХ. ПОНЯТИЕ О ПЛАЗМЕ

Допустим, что изучаемый газ заключен в сосуд C с двумя электродами, к которым прикладывается разность потенциалов U. Электрическое поле E между электродами можно изменять, перемещая движок $Д$ потенциометра $П$, замыкающего батарею $Б$ (рис. III.42). Если в газе нет свободных зарядов (положительных или отрицательных ионов или электронов), то в цепи гальванометра тока не будет. Заметим, что в газах всегда будет находиться некоторое количество зарядов, так как газ ионизируется как при неизбежных тепловых столкновениях молекул, так и под действием различных излучений, в частности,

от радиоактивных веществ. Однако одновременно с процессом *ионизации*, т. е. разделения нейтральных молекул на заряженные ионы, в газе происходит и обратный процесс молизации или *рекомбинации*, т. е. соединения ионов в нейтральные молекулы. В равновесном состоянии газов оба эти процесса уравновешиваются: число ежесекундно ионизирующихся молекул равно числу нейтральных молекул, вновь образовавшихся из ионов за то же время.

Если на газ внешнего ионизирующего воздействия не оказывается, то естественная концентрация ионов в нем будет очень малой, и ток через газ практически не обнаруживается. Вызвать заметный электрический ток в газе (так называемый *газовый разряд*) можно, если: 1) при помощи постороннего воздействия (*ионизатора*) непрерывно разбивать нейтральные молекулы газа на ионы и тем самым увеличивать концентрацию свободных зарядов в газе. Это можно сделать, подвергая газ интенсивному облучению потоком быстрых частиц (электронов и др.), ультрафиолетовыми, рентгеновскими лучами, лучами радиоактивных веществ, а также повышая температуру газа, чтобы увеличить интенсивность ионизации при тепловых столкновениях. В этом случае вместе с прекращением действия внешнего ионизатора прекращается и ток через газы; такая проводимость газа называется *несамостоятельной*; 2) приложить настолько большую разность потенциалов, чтобы имеющиеся в газе ионы, разгоняясь в электрическом поле, приобретали энергии, достаточные для ионизации нейтральных молекул при столкновениях с ними. В этом случае каждый ион при одном столкновении вызывает появление двух или нескольких ионов; эти ионы в свою очередь разгоняются в поле и разбивают нейтральные молекулы на ионы. Таким образом, число ионов в газе быстро растет, и газ приобретает заметную проводимость; такая проводимость называется *самостоятельной*.

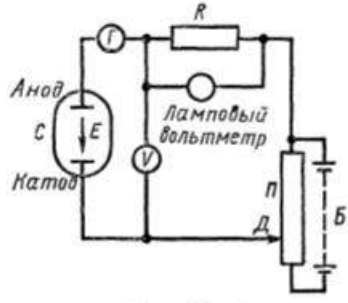


Рис. III.42

Следует различать два вида столкновений между частицами, в частности между ионами, электронами и нейтральными молекулами. При одних столкновениях частицы не испытывают никаких внутренних изменений, а только обмениваются кинетическими энергиями движения. Такие столкновения называются *упругими*; сумма кинетических энергий частиц до и после удара остается постоянной.

При других — *неупругих* — столкновениях атомы и молекулы испытывают изменения в своем строении; происходит переход кинетической энергии соударяющихся частиц в потенциальную энергию взаимодействия составных частей этих атомов и молекул — ядер и вращающихся вокруг них электронов. Такой процесс называется *возбуждением* атомов или молекул; при обратном переходе в нормальное состояние поглощенная энергия возвращается в виде энергии излучения. Наконец, при неупругих столкновениях возможно также изме-

нение состава атомов и молекул; в частности нейтральная молекула может быть разбита на два иона или от атома может быть оторван электрон и т. д. Ионизация газов при соударениях является результатом неупругих столкновений.

Для проводимости газов при некоторых условиях (в частности при малых давлениях газа в сосуде) заметное значение имеет выбивание электронов с поверхности катода при падении на него положительных ионов. Каждый такой ион может освободить из катода несколько электронов в зависимости от энергии, приобретенной им в электрическом поле, а также — от работы выхода электрона из вещества катода. Освобожденные из катода электроны, подхваченные электрическим полем, могут на пути к аноду вызвать ионизацию газа; кроме того, этот упорядоченный поток электронов составляет некоторую (иногда значительную) долю всего тока, протекающего через газ.

Если сила тока, проходящая через газы, мала и не может быть непосредственно обнаружена гальванометром $Г$, то прибегают к косвенным методам. В частности, как это показано на рис. III.42, в цепь последовательно с газовым промежутком включается резистор с сопротивлением R порядка десятков и сотен миллионов ом. На концах этого резистора образуется разность потенциалов $U = IR$, которую измеряют, например, ламповым вольтметром, не замыкающим концы этого резистора. Тогда, зная R и измерив U, можно рассчитать силу тока через газ I. Например, если $U = 0{,}01\ В$, $R = 10^7$ Ом, то $I = 0{,}01/10^7 = 10^{-9}$ А.

ДВИЖЕНИЕ ИОНОВ В ГАЗЕ

Для расчета силы тока через газы допустим, что в единице объема газа содержится n_+ положительных ионов, несущих заряды q_+ и n_- отрицательных ионов с зарядами q_-; числа n_+ и n_- могут быть не равны друг другу, но суммы зарядов $n_+q_+ = n_-q_-$. Так как движение положительных зарядов в одном, а отрицательных зарядов в противоположном направлении означает электрический ток одного и того же направления, то, согласно формуле (2.2), сила тока через газ равна:

$$I = q_+n_+v_+S + q_-n_-v_-S, \tag{2.36}$$

где v_+ и v_- — скорости упорядоченного движения соответствующих ионов в газе; S — площадь, через которую происходит движение ионов. Эта площадь может быть различной в различных сечениях сосуда с газом; рассмотрим плотность тока $j = I/S$:

$$j = q_+n_+v_+ + q_-n_-v_-. \tag{2.37}$$

Движение ионов через газ сопровождается потерей их энергии при столкновениях, которая при установившемся движении ионов непрерывно компенсируется за счет работы электрического поля. Можно оценить сопротивление, испытываемое ионами при их упорядоченном движении в газе, если ввести «силы трения» F_+ и F_-, пропорциональные скоростям движения соответствующих ионов

$$F_+ = k_+v_+; \quad F_- = k_-v_-,$$

где k_+ и k_- — коэффициенты трения для положительных и отрицательных ионов (размеры и конфигурация разноименных ионов могут сильно отличаться). При установившемся движении ионов эти силы трения должны быть уравновешены электростатическими силами q_+E и q_-E; следовательно,

$$k_+v_+ = q_+E; \quad v_+ = \frac{q_+}{k_+}E = b_+E;$$
$$k_-v_- = q_-E; \quad v_- = \frac{q_-}{k_-}E = b_-E. \tag{2.38}$$

Величины

$$b_+ = q_+/k_+ = v_+/E; \quad b_- = q_-/k_- = v_-/E \tag{2.39}$$

называются *подвижностями ионов*; они численно равны скоростям установившегося движения ионов при напряженности поля E, равной единице. Подставив выражение (2.38) в формулу (2.37), получаем для плотности тока:

$$j = (q_+n_+b_+ + q_-n_-b_-)E = \sigma E.$$

Таким образом, коэффициент электропроводности газов

$$\sigma = q_+n_+b_+ + q_-n_-b_- \tag{2.40}$$

зависит от концентрации ионов и их подвижностей (или от коэффициентов трения).

При прохождении электрического тока через газы исчезновение ионов происходит, во-первых, вследствие молизации, т. е. случайных встреч противоположных ионов и соединения их в нейтральные молекулы, во-вторых вследствие нейтрализации ионов при их соприкосновении с противоположно заряженными электродами. Первый процесс-молизацию — можно считать пропорциональным концентрации как положительных, так и отрицательных ионов, т. е. пропорциональным произведению n_+n_-; тогда ежесекундное число ионов одного знака, исчезающих вследствие молизации в единице объема газа, может быть представлено в виде

$$\Delta n'_+ = \beta_+n_+n_-; \quad \Delta n'_- = \beta_-n_+n_-,$$

где $\beta_\pm$ — некоторые коэффициенты, зависящие от вещества газа, давления, температуры и т. д., а также от структуры иона.

Второй процесс — нейтрализация ионов у электродов — пропорционален силе тока I, проходящего через газ. Допустим, что в единицу времени нейтрализуется $\Delta N''$ положительных и отрицательных ионов, причем $\Delta N'' \sim I$. В единице объема газа

$$\Delta n'' = \frac{\Delta N''}{V} \sim \frac{I}{V},$$

где V — объем газа. Если полагать $V = Sl$, где S — сечение; l — длина сосуда с газом, то $I/Sl = j/l$ (j — плотность тока). В равновесном состоянии число ионов $\Delta n_\pm$, образующихся в единице объема газа в единицу времени, должно компенсироваться исчезновением

ионов соответствующих знаков от молизации и нейтрализации у электродов:

$$\Delta n_{\pm} = \Delta n'_{\pm} + \Delta n''_{\pm}.$$

При постоянной интенсивности внешнего ионизатора эти числа сохраняются неизменными. Если дополнительно к этому сила или плотность тока через газ очень мала, то величиной $\Delta n''_{\pm}$ можно пренебречь по сравнению с величиной $\Delta n'_{\pm}$. Тогда из постоянства $\Delta n_{\pm} \approx \Delta n'_{\pm} = \beta_{\pm} n_+ n_-$ следует постоянство n_+ и n_-. При этих условиях электропроводность σ в формуле (2.40) — постоянная величина (не изменяющаяся при увеличении или уменьшении тока через газ), и тогда формула $j = \sigma E$ совпадает с законом Ома (2.6). Таким образом, при слабых токах в газах соблюдается прямая пропорциональность между силой тока и приложенным напряжением.

При очень сильных токах, наоборот, можно пренебрегать молизацией ($\Delta n'$) по сравнению с нейтрализацией ионов у электродов ($\Delta n''$). Тогда $\Delta n_{\pm} \approx \Delta n''_{\pm}$ при различных значениях приложенного к газу напряжения. Так как сила тока пропорциональна $\Delta n''$, то, следовательно, она пропорциональна Δn. Таким образом, при указанном выше условии сила (а следовательно, и плотность) тока не зависит от приложенной к газу разности потенциалов, а определяется только числом ионов $N_{\pm} = \Delta n_{\pm} V$, образующихся в единицу времени во всем объеме газа, т. е. определяется интенсивностью внешнего ионизатора. Этот максимально возможный в газе ток (который нельзя увеличить путем увеличения приложенной разности потенциалов) называется *током насыщения*.

В некоторых случаях проводимость газов обусловлена главным образом движением электронов. При большой разности потенциалов положительные ионы приобретают в электрическом поле большие скорости и, ударяясь о катод, выбивают из него электроны. Существенно, что один тяжелый ион может выбить из катода несколько, иногда значительное число электронов. При очень больших плотностях тока в газе удары положительных ионов нагревают катод до очень высокой температуры и тем вызывают дополнительно интенсивную термоэлектронную эмиссию. Кроме того, электрическое поле у поверхности катода может быть настолько большим, что почти все электроны, вылетающие за пределы поверхности катода, будут подхвачены полем и направлены к аноду. Все эти электроны, двигаясь ускоренно к аноду, могут вызывать как возбуждение атомов газа, так и их ионизацию. Вновь образовавшиеся при этом положительные ионы, достигнув катода, снова выбивают из него электроны и т. д. При таких условиях та часть тока, которая обусловлена упорядоченным движением электронов, может составлять весьма значительную часть общей силы тока, т. е. ток через газ в основном обусловливается движением электронов от катода к аноду. Движение же положительных ионов будет составлять малую часть силы тока, но их удары о катод играют существенную роль.

Возбужденные (например, при столкновениях с электронами) атомы газа, возвращаясь в нормальное состояние, излучают свет;

этим объясняется свечение газа при газовых разрядах. Характер этого свечения (искровой, дуговой) зависит в основном от плотности (давления) газа и от величины прилагаемой разности потенциалов; спектральный состав свечения обусловлен химической природой газа, через который проходит ток.

Прохождение электрического тока через разреженные газы имеет некоторые особенности. При очень больших разрежениях («высокий вакуум»), т. е. при весьма малом числе частиц n_0 в единице объема газа, длина свободного пробега частицы (молекулы или иона)

$$\lambda = \frac{1}{\sqrt{2\pi d^2 n_0}}$$

(d — эффективный диаметр частицы) делается соизмеримой с размерами сосуда. Тогда число столкновений частиц между собой становится значительно меньше, чем число столкновений этих частиц со стенками сосуда. При таких условиях процессы ионизации и молизации в объеме сосуда не играют заметной роли. Чтобы газ мог проводить ток, необходимо ввести в пространство между электродами свободные заряды каким-нибудь другим путем. В частности, это можно сделать при помощи термоэлектронной эмиссии, если нагреть электрическим током электрод, присоединенный к отрицательному полюсу батареи (т. е. катод). В этом случае ток через газ будет почти целиком электронным. По мере увеличения давления (плотности) газа постепенно возрастает роль ионизации, происходящей внутри газа вследствие столкновений между нейтральными атомами, ионами и электронами.

ВИДЫ РАЗРЯДОВ В ГАЗАХ

Характер газовых разрядов определяется составом газа, его давлением и температурой, приложенным напряжением, формой, размерами, относительным расположением и веществом электродов. Различают следующие виды газовых разрядов:

1) т и х и й р а з р я д; весьма слабые плотности тока в газе; не сопровождается испусканием света или звука;

2) т л е ю щ и й р а з р я д; представляет собой ток через разреженный газ при высоких напряжениях (давление газа — порядка 1 мм. рт. ст., напряжение — несколько сотен вольт). Проводимость обусловлена главным образом движением электронов и положительных ионов, выбивающих электроны из катода. Сопровождается характерным (холодным) свечением, в котором можно выделить области, где электроны разгоняются полем, и области, где они, неупруго сталкиваясь с атомами газа, возбуждают и ионизируют их. Газосветные трубки, наполненные гелием, неоном, криптоном и другими газами, испускают излучение различного цвета;

3) д у г о в о й р а з р я д; характеризуется большими плотностями тока (тысячи ампер на 1 мм^2), сравнительно небольшими напряжениями (десятки вольт), сильным свечением и высокой температурой (3000—6000° С). Электрический ток представляет собой движение электронов, вылетающих из раскаленного катода, и ионов, образую-

щихся при тепловых столкновениях. Применяется для целей сварки, для освещения (дуговые фонари, прожекторы), для получения высоких температур и т. д.;

4) искровой разряд, представляющий собой электрический пробой газов при кратковременном лавинообразном возрастании числа ионов в нем. Для этого необходимо приложить к газу достаточно высокое напряжение, необходимое для обеспечения ионизации при столкновениях;

5) коронный разряд; весьма слабые электрические токи в газах при атмосферном давлении, вызванные сильно неоднородным полем при очень больших значениях напряженности поля; это имеет место, в частности, возле остриев, возле тонких проводов высокого напряжения и т. д. Наблюдается слабое свечение около электродов.

ПЛАЗМА

Газ, в котором имеется большое количество положительных и отрицательных ионов, а также свободных электронов, называется *плазмой*. Плазма может быть получена, во-первых, при прохождении электрического тока через газы. Вследствие столкновений ионов и электронов (разгоняемых электрическим полем) с нейтральными атомами газа непрерывно происходит ионизация газа; концентрация ионов и электронов при большой силе тока через газ, может быть значительной. Такая плазма состоит из смеси нейтральных атомов, ионов и электронов; ее существование поддерживается током, проходящим через газ. При удалении приложенного электрического поля противоположно заряженные частицы газа рекомбинируют и плазменное состояние газа исчезает.

Плазма может быть получена при нагревании любого вещества до очень высоких температур (свыше 10 000° С). При этих условиях вещество находится в газообразном состоянии, причем вследствие тепловых столкновений почти все атомы вещества превращаются в ионы; концентрация нейтральных атомов оказывается очень малой. По мере повышения температуры плазмы сначала происходит постепенное удаление электронов с внешних оболочек атомов, а затем при температурах порядка миллионов градусов все атомы оказываются лишенными своих электронных оболочек. В этом состоянии плазма содержит смесь положительно заряженных атомных ядер и отрицательно заряженных электронов. Обратное образование нейтральных атомов при столкновении ядер и электронов при этих температурах имеет весьма малую вероятность.

Средние кинетические энергии теплового движения различных типов частиц плазмы в отличие от идеального газа оказываются различными. Наибольшей энергией обладают электроны; энергия ионов и нейтральных атомов меньше. Если воспользоваться связью между средней кинетической энергией частицы и температурой газа

$$\left(\frac{mv^2}{2}\right)_{\text{сред}} = \frac{3}{2}kT,$$

то можно утверждать, что различные компоненты плазмы имеют различные температуры. Обозначим температуру электронного газа в плазме через $T_э$, ионного газа — $T_и$ и газа, составленного из нейтральных атомов, — $T_а$. Ввиду большого различия в энергиях, «электронная температура» $T_э$ оказывается значительно больше «ионной» ($T_и$) и «атомной» ($T_а$) температур; таким образом

$$T_э \gg T_и > T_а.$$

В газоразрядных трубках и в электрической дуге электронная температура (достигающая десятков тысяч температур) в несколько раз (2—10) превышает ионную и атомную температуры. Однако в некоторых специальных способах получения сильно ионизированной плазмы можно получить $T_и > T_э$. Существенное значение имеет утечка электронов через стенки сосуда, в котором находится плазма. Эта утечка может привести к тому, что суммарный положительный заряд ионов плазмы может оказаться больше суммарного отрицательного заряда электронов. Первоначально нейтральная (в целом) плазма может приобрести положительный заряд.

Плазма отличается рядом важных свойств; она является хорошим проводником электрического тока, взаимодействует с электрическим и магнитным полем, обладает магнитными свойствами и т. д. В последнее время изучению плазмы уделяется очень большое внимание (см. ч. IV, § 22).

§ 15. ЭЛЕКТРИЧЕСКИЙ ТОК В ЖИДКОСТЯХ. ЗАКОНЫ ФАРАДЕЯ

Электропроводность жидких веществ (электролитов) обусловлена наличием в них положительных и отрицательных ионов, образовавшихся при *электролитической диссоциации* как молекул самой жидкости, так и молекул растворенного вещества, если они имеются. Степенью, или *коэффициентом диссоциации*, называется отношение числа молекул n, распавшихся на ионы, к полному числу нейтральных молекул в данном объеме жидкости (в дальнейшем будем иметь в виду только единичный объем жидкости):

$$\alpha = n/N.$$

Коэффициент α зависит от вещества жидкости и ее температуры, а для растворенных веществ еще и от концентрации раствора (для растворов следует различать коэффициенты диссоциации самой жидкости и растворенного вещества; обычно первый коэффициент значительно меньше второго и диссоциацией молекул растворителя пренебрегают). Степень диссоциации возрастает с увеличением температуры и уменьшением концентрации растворенного вещества. В сильно разбавленных растворах $\alpha \approx 1$.

Число молекул Δn, распадающихся на ионы в единицу времени, пропорционально оставшемуся числу нейтральных молекул, которое равно $N - n = N - \alpha N = N(1 - \alpha)$; тогда, обозначив коэффициент этой пропорциональности через β, можно написать:

$$\Delta n = \beta N (1 - \alpha).$$

Величину β называют *коэффициентом ионизации*; он также зависит от вещества и температуры.

Допустим, что в единице объема жидкости указанные выше n молекул разделились на n_+ положительных и n_- отрицательных ионов. Эти числа могут быть и не равны; например, одна молекула H_2SO_4 распадается на два положительных иона H^+ и один отрицательный ион SO_4^{--}. Обозначим заряды ионов через q_+, их валентность — через z_+, а элементарный электрический заряд — по-прежнему через e; тогда $q_+ = z_+e$; $q_- = z_-e$. Так как суммарные заряды ионов должны быть равны, то $n_+q_+ = n_-q_-$, следовательно, $n_+z_+e = n_-z_-e$. Числа ионов n_+ и n_- пропорциональны числу диссоциированных молекул n, т. е. пропорциональны αN.

Кроме ионизации, в жидкости происходит еще и обратный процесс молизации или рекомбинации. Интенсивность этого обратного процесса, т. е. числа $\Delta n'$ нейтральных молекул, образовавшихся из ионов в единицу времени, пропорциональна как n_+, так и n_-, т. е. пропорциональна их произведению:

$$\Delta n' = \gamma_0 n_+ n_-,$$

где γ_0 — *коэффициент молизации*. Обычно для простоты рассуждений имеют в виду простейший случай, когда $n_+ = n_- = n$, и тогда $\Delta n' = \gamma_0 (\alpha N)^2$. При образовании $\Delta n'$ нейтральных молекул исчезает Δn_+ положительных и $\Delta n'_-$ отрицательных ионов. Если через жидкость электрический ток не проходит, то молизация является единственным процессом, при котором исчезают ионы; в равновесном состоянии жидкости должно соблюдаться условие $\Delta n' = \Delta n$, и тогда

$$\gamma_0 (\alpha N)^2 = \beta N (1 - \alpha); \quad \frac{1-\alpha}{\alpha^2} = \frac{\gamma_0}{\beta} N.$$

Однако, если через жидкость проходит электрический ток, то исчезновение ионов происходит не только внутри объема жидкости вследствие молизации, но и на электродах, при соприкосновении ионов с противоположно заряженными электродами. При силе тока через жидкость, равной I, в единицу времени на электродах нейтрализуется $\Delta n''_+ = I/q_+ = I/z_+e$ положительных и $\Delta n''_- = I/q_- = I/z_-e$ отрицательных ионов. Очевидно, при установившемся протекании тока должно соблюдаться условие равенства числа появляющихся и исчезающих ионов в единицу времени:

$$\Delta n_\pm = \Delta n'_\pm + \Delta n''_\pm.$$

Подставив в формулу (2.37) значения зарядов $q_\pm = z_\pm e$ и заменив $n_\pm z_\pm = n = \alpha N$, получим для плотности тока следующее выражение:

$$j = \alpha N e (v_+ + v_-).$$

Введем подвижности ионов $b_\pm = v_\pm/E$ при помощи соотношений (2.38) и обозначим $\eta = N/N_A$, где N_A — постоянная Авогадро (число молекул в одном моле вещества); η — показывает число молей вещества в единице объема, где содержится N молекул. Кроме того, $N_A e = F$

есть *число Фарадея* ($F \approx 9{,}65 \cdot 10^4$ Кл/моль); тогда

$$j = \alpha \frac{N}{N_A} N_A e (b_+ + b_-) E = \alpha \eta F (b_+ + b_-) E = \sigma E. \qquad (2.41)$$

Отношение

$$\frac{\sigma}{\eta} = \Lambda = \alpha F (b_+ + b_-) \qquad (2.42)$$

называется *эквивалентной электрической проводимостью*, т. е. проводимостью, приходящейся на единицу концентрации η. Повышение электропроводности жидкостей при нагревании объясняется увеличением коэффициента диссоциации α и, кроме того, увеличением подвижности ионов. Подвижности ионов зависят от состава и структуры этих ионов, от вещества растворителя и от температуры; для ионов серной кислоты при 18° С сумма подвижностей составляет около 0,004 см²/(с·В).

Допустим, что ионы какого-нибудь знака, нейтрализуясь у соответствующего электрода, осаждаются на нем. Определим количество вещества, выделяющегося на этом электроде в единицу времени. Число нейтрализующихся ионов в единицу времени $\Delta n_{\pm}^{''} = I/z_{\pm}e$; обозначим массу одного иона через m_0, тогда ежесекундное количество осажденного вещества

$$\frac{\Delta M}{\Delta t} = m_0 \Delta n_{\pm}^{''} = \frac{m_0 I}{z_{\pm} e}.$$

Умножим и разделим это выражение на постоянную Авогадро N_A и учтем, что $m_0 N_0 = A$ есть масса одного моля вещества, а $eN_A = F$ — число Фарадея. Тогда выражение

$$\Delta M = \frac{1}{F} \frac{A}{z} I \Delta t \qquad (2.43)$$

представляет собой объединенный (первый и второй) з а к о н Ф а р а д е я для электролиза.

Электролиз применяется для получения ряда важных металлов из расплавленных солей (цинк, медь, алюминий, магний, калий, натрий и др.), для покрытия металлических изделий тонким слоем коррозионноустойчивых веществ — хромом, никелем, алюминием, серебром, покрытия железных листов слоем цинка, для получения очень чистых металлов, для изготовления типографских клише, электролитической полировки поверхностей различных деталей (например, лопастей паровых турбин), для изготовления труб (медных) без швов, для резки металлов (так называемая анодно-механическая резка очень твердых сплавов при помощи вращающихся медных или железных дисков, смачиваемых специальным электролитом) и др.

§ 16. ПОЛУПРОВОДНИКИ И ИХ ПРИМЕНЕНИЯ

Полупроводниками называются вещества, обладающие особым характером электрической проводимости, у которых удельная электропроводность σ сильно зависит от температуры, от действующего

внутри них электрического поля, от наличия примесей, от интенсивности облучения, внешнего давления и т. д. Особенно характерна для них зависимость электропроводности от температуры

$$\sigma = \sigma_0 e^{-b/T}, \qquad (2.44)$$

где b — величина, зависящая от вещества. У металлов электропроводность сравнительно слабо зависит от температуры и с повышением ее уменьшается (на 1° С — 0,3% и менее); у полупроводников с повышением температуры σ резко возрастает по экспоненциальному закону (2.44) — до 5—6% на 1°С. При высоких температурах полупроводники по электропроводности приближаются к металлам, а при низких температурах являются хорошими изоляторами.

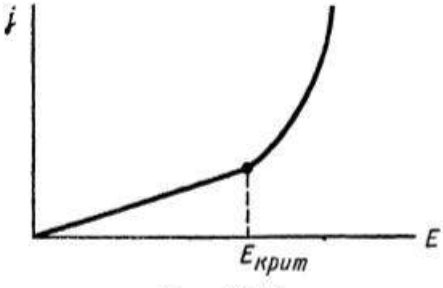


Рис. III.43

Важной отличительной чертой полупроводников является также заметная зависимость электропроводности от напряжённости E электрического поля, вызывающего в них ток. У металлов σ не зависит от E, поэтому, согласно дифференциальному закону Ома ($j = \sigma E$), плотность тока j прямо пропорциональна напряженности электрического поля. У полупроводников такая пропорциональность существует только до некоторого критического значения ($E_{крит}$) этой напряженности; при $E > E_{крит}$ закон Ома (прямая пропорциональность между j и E) не соблюдается. Зависимость j от E у многих чистых полупроводников имеет вид, показанный на рис. III. 43, причем $E_{крит}$ имеет значение около 10^4 В/см. У других полупроводников отступления от закона Ома наблюдаются при меньших значениях $E_{крит}$ (10^2—10^3 В/см) и имеют более сложный характер.

Свойства полупроводников сильно зависят от их химической чистоты. У металлов примеси уменьшают электропроводность; у полупроводников незначительное количество примесей резко повышает электропроводность. Например, у германия заметное увеличение электропроводности обнаруживается при изменении концентрации индия в нем на миллионную долю процента; таким образом, добавление 1% примеси может увеличить σ в миллионы раз.

ЭЛЕКТРОПРОВОДНОСТЬ ПОЛУПРОВОДНИКОВ

Электропроводность многих полупроводников резко увеличивается под действием рентгеновского, ультрафиолетсвого, светового и др. излучений и зависит от интенсивности поглощенного ими излучения; у металлов такой зависимости не наблюдается.

Для объяснения этих, а также и других свойств полупроводников необходимы подробные сведения о структуре твердых тел, характере процессов, которые происходят в них при прохождении электриче-

ского тока, при различных воздействиях на эти тела, влиянии примесей и т. д. Ниже будет дано только схематическое изложение основных представлений о «механизме» электропроводности в полупроводниках.

В твердом веществе связь между соседними атомами осуществляется благодаря валентным электронам, которые цементируют атомы в прочное тело. Установлено, что связь между двумя атомами наиболее прочная, если она осуществляется двумя электронами, одновременно взаимодействующими с обоими атомами (двухэлектронная или валентная связь). Например, молекула двухатомного водорода обнаруживает большую прочность благодаря двухэлектронной связи между ее атомами. В элементарной ячейке (тетраэдре) кристаллической решетки германия каждый атом связан с четырьмя соседними атомами при помощи восьми электронов: четыре электрона взяты из своей валентной оболочки, а остальные четыре — из оболочек соседей. Принято говорить, что электроны «заполняют связи» между атомами. Таким образом, для прочности твердого тела необходимо определенное количество электронов, осуществляющих связи между атомами. Это число зависит от кристаллической структуры вещества. Если данное тело располагает бо́льшим количеством валентных электронов, чем это необходимо для «заполнения связей» между атомами, то избыточные электроны оказываются слабо связанными с атомами, т. е. являются почти свободными. Наоборот, при нехватке электронов, связи между атомами окажутся «незаполненными» («вакантными»). Иллюстрируем это на примере кристаллической решетки германия; допустим, что один из атомов германия заменен атомом другого вещества, имеющим в валентной оболочке не четыре, а пять электронов (мышьяк). Тогда из этих пяти электронов четыре используются для прочной связи с четырьмя соседними атомами германия, а пятый электрон атома мышьяка останется лишним, слабо связанным с атомами. Если же в кристаллической решетке германия один из его атомов заменить атомом другого вещества, имеющего три валентных электрона (индий), то для осуществления двухэлектронных связей этого атома с четырьмя соседними атомами германия недостает одного электрона: в этом случае одна из связей оказывается незаполненной (вакантной, пустой).

До сих пор мы считали, что вещество обладает электропроводностью, если в нем имеются свободные электроны (или почти свободные), весьма слабо связанные с атомами решетки. Однако можно показать, что электропроводность может быть осуществлена и при отсутствии свободных электронов (при помощи связанных электронов), если только:

1) в веществе имеются незаполненные связи,

2) температура вещества достаточно высокая.

На рис. III. 44, *а* большими кружками обозначены пять четырехвалентных атомов германия и один трехвалентный (примесный) атом индия, связанные между собой электронами (боковые связи между атомами не показаны). При этом одна связь, показанная квадратиком, остается незаполненной. При наличии теплового движения возможно, что один из соседних электронов перейдет на это вакантное место.

Тогда связи примесного атома индия окажутся заполненными, но появится «пустое место» в связях соседнего атома германия (рис. III.44,*б*). Если внешнего электрического поля нет и существует только тепловое движение, то переход электронов от заполненных связей к пустым происходит совершенно беспорядочно. При наличии же внешнего электрического поля переход от заполненных связей к пустым в направлении действующей на них внешней электрической силы будет происходить чаще, чем переход в противоположном направлении (в котором электрическое поле тормозит движение электрона). Таким образом, благодаря наличию вакансий в связях между атомами, внешнее электрическое поле при содействии теплового движения вызывает упорядоченное перемещение связанных электронов в направлении, противоположном напряженности поля. При этом вакантная связь, получившая название «дырки», перемещается в направлении поля, как перемещались бы положительно заряженные частицы.

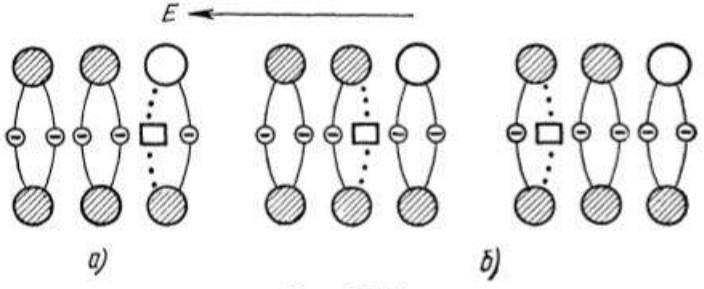


Рис. III.44

Электропроводность, обусловленная движением свободных электронов, называется *электронной*, а движением связанных электронов (т. е. перемещением вакантных связей или «дырок») — *дырочной* электропроводностью.

Незаполненные связи («дырки») могут быть получены не только при помощи примесей, т. е. замены атомов данного вещества атомами другого вещества с меньшим числом валентных электронов, но и при помощи облучения или под воздействием интенсивного теплового движения, при котором могут быть «разорваны» связи между соседними атомами. В этом случае одновременно образуются и свободный электрон и «дырка». По мере увеличения их концентрации в объеме вещества происходит рекомбинация, т. е. заполнение вакантных связей свободными электронами. При каждой температуре существует определенное равновесное состояние между противоположными процессами освобождения и заполнения связей; число свободных электронов и число вакантных связей в единице объема германия при нормальной температуре имеют порядок 10^{13} см$^{-3}$.

Таким образом, свободные электроны и вакантные связи существуют в чистых полупроводниках при любой температуре; с повышением температуры концентрации их растут. Электропроводность вещества, обусловленная свободными электронами и «дырками», образовавши-

мися в равных количествах при тепловых движениях атомов, называется *собственной*.

Примеси, добавление которых к полупроводнику приводит к увеличению концентрации свободных электронов, называются *донорами*; примеси, увеличивающие концентрацию «дырок», называются *акцепторами*. Если проводимость полупроводников обусловлена преимущественно свободными электронами, то они называются полупроводниками *n*-типа; для них «дырочная» проводимость является неосновной, побочной. Полупроводники, у которых основными «носителями тока» являются «дырки», а побочными — свободные электроны, называются полупроводниками *p*-типа.

В любом полупроводнике имеются и свободные электроны и вакантные связи — «дырки». Обозначим через $n_э$ и $n_д$ числа электронов и дырок в единице объема вещества, а $b_э$ и $b_д$ — их подвижности; удельная электропроводность полупроводника

$$\sigma = n_э e b_э + n_д e b_д^0.$$

При собственной проводимости (в химически чистом веществе) $n_э = n_д$; при наличии примесей эти числа будут отличаться. Например, в германии с примесью мышьяка $n_э > n_д$, а с примесью индия $n_э < n_д$. Подвижности электронов и дырок зависят от структуры кристаллической решетки (у ионных решеток они меньше, чем у атомных), а также от температуры; в данном веществе $b_э \neq b_д$ (у большинства полупроводников $b_э > b_д$: например, у германия при $t = 20°$ С $b_э = 0{,}26$, $b_д = 0{,}19$ м²/(В · с). Изменение σ под действием различных физических факторов (температуры, электрического поля, излучения, давления и т. д.) происходит преимущественно вследствие того, что эти факторы изменяют $n_э$.

ЯВЛЕНИЯ НА ГРАНИЦЕ ДВУХ ПОЛУПРОВОДНИКОВ

Рассмотрим явления, происходящие в месте контакта двух полупроводников *n*- и *p*-типа (рис. III. 45, *а*). Очевидно, вследствие теплового движения через место контакта происходит диффузия свободных

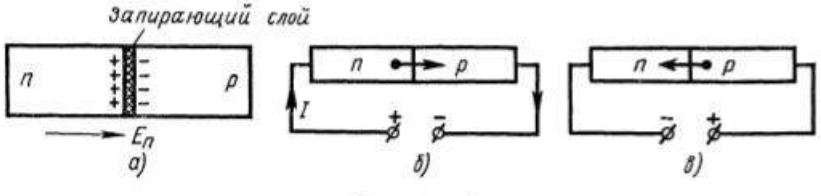


Рис. III.45

электронов и дырок в обоих направлениях $n \to p$ и $p \to n$. Так как концентрация свободных электронов в полупроводнике *n*-типа больше, чем в полупроводнике *p*-типа, то при этой диффузии число электронов, переходящих в единицу времени из *n* в *p*, будет больше, чем переходящих в обратном направлении. Однако электроны, перешедшие

в p-полупроводник, не увеличивают концентрацию свободных электронов в этом теле (как это происходило при контакте двух различных металлов); так как в полупроводнике p имеется большая концентрация вакантных связей, то электроны, перешедшие из n-полупроводника, занимают часть «дырок» и оказываются связанными.

Точно так же переход «дырок» из p в n (т. е. переход связанных электронов из n в p) не увеличивает концентрацию «дырок» в n-полупроводнике; вследствие рекомбинации со свободными электронами, в избытке имеющимися в n-полупроводнике, эти «дырки» немедленно заполняются. В результате этих процессов n-полупроводник теряет электроны и поэтому заряжается положительно; p-полупроводник заряжается отрицательно.

На границе между этими полупроводниками, в слое небольшой толщины образуется электрическое поле, препятствующее дальнейшей односторонней диффузии электронов в направлении $n \rightarrow p$; слой, в котором образовался этот «потенциальный барьер» для электронов, называется *запирающим*. После установления равновесия переход электронов в направлении $n \rightarrow p$ компенсируется обратным переходом $p \rightarrow n$.

Если к этим полупроводникам приложить внешнее электрическое поле, то равновесие нарушится. Положительный полюс источника тока можно присоединить либо к n-, либо к p-полупроводнику. В первом случае (рис. III. 45, *б*) внешнее поле препятствует движению основных «носителей тока» (электронов — в n и дырок — в p) и способствует движению побочных носителей тока, концентрация которых в полупроводниках невелика. Можно сказать, что источник тока «отсасывает» электроны из n и «посылает» их в p, вследствие чего уменьшается концентрация основных носителей тока в обоих полупроводниках и их сопротивление сильно возрастает. Это направление тока называется *обратным* или запирающим.

Во втором случае (рис. III. 45, *в*) внешнее электрическое поле способствует движению основных «носителей тока» и препятствует движению побочных. Можно сказать, что внешнее поле ослабляет (понижает) потенциальный барьер между полупроводниками и тем самым увеличивает диффузию электронов из n в p и «дырок» — в обратном направлении. Сопротивление полупроводников уменьшается, и сила тока получается большой. Это направление тока называется *прямым*.

Таким образом, в одном («прямом») направлении тока полупроводниковый диод имеет малое сопротивление, а в обратном («запирающем») направлении — очень большое сопротивление. Это свойство диода используется для выпрямления переменного тока. Зависимость силы тока I через диод от приложенного внешнего напряжения U(вольтамперная характеристика) показана на рис. III. 46.

Рассмотрим свойства полупроводникового триода, состоящего из комбинации переходов n-p-n или p-n-p. На рис. III. 47 показана схема включения триода, предназначенная для получения на выходном сопротивлении R усиленных колебаний переменного напряжения $\mathscr{E}$. В переходе n_1-p электрический ток, вызываемый батареей $Б_1$, течет в прямом направлении (напряжение этой батареи должно быть

больше $\mathscr{E}$, чтобы при перемене знака $\mathscr{E}$ переход n_1-p не запирался). В переходе $p \rightarrow n_2$ действует обратное (запирающее) напряжение батареи $Б_2$. Допустим, что напряжение, подаваемое на переход n_1-p, изменилось на величину $\Delta U_э = \Delta \mathscr{E}$. Тогда сила тока через этот переход $I_э$ изменится на $\Delta I_э$. Изменение силы тока между полу-

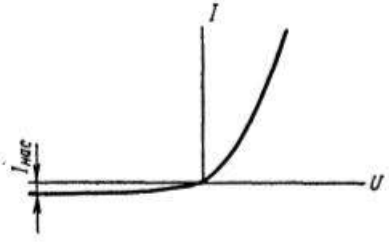


Рис. 46

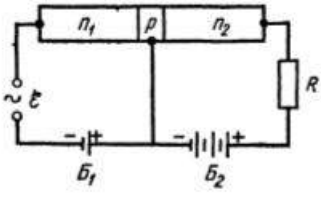


Рис. III.47

проводниками n_1 и p будет сопровождаться изменением концентрации носителей тока в полупроводнике p, вследствие чего изменится сопротивление в n_2-p переходе. Это вызовет изменение силы тока $I_к$ в цепи n_2-p на величину $\Delta I_к$, а падение напряжения на выходном сопротивлении R изменится на $\Delta U_к = R\Delta I_к$. Отношение $\Delta I_к/\Delta I_э$ зависит от конструкции триода (например, для триодов с плоскими поверхностями контактов это отношение ≈ 1, при точечных контактах $\approx 2 + 2{,}5$). Усиление напряжения, которое можно получить в триоде, равно:

$$\mu = \frac{\Delta U_к}{\Delta U_э} = \frac{\Delta I_к}{\Delta I_э}\frac{R}{r},$$

где r — сопротивление в цепи переменного напряжения $\mathscr{E}$. Величины R и r можно подобрать такими, чтобы получить необходимое усиление. Полупроводник p называется *базой* или *основанием триода*; полупроводник n_1, имеющий с базой открытый переход, называется *эмиттером* и является поставщиком носителей тока в базу: полупроводник n_2, у которого переход с базой заперт, называется *коллектором*. Весь триод, работающий как усилитель, называется т р а н з и с т о р о м. Потери энергии в транзисторе очень малы по сравнению с электронной лампой (в которой необходима затрата энергии на подогрев катода).

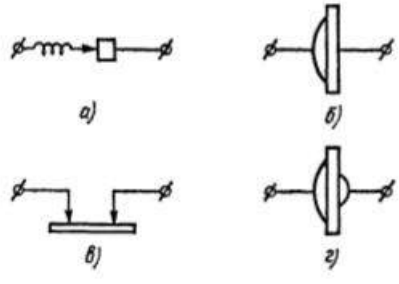


Рис. III.48

Схематическое устройство полупроводниковых диодов и триодов показано на рис. III. 48. В маломощном точечном диоде (*а*) металлическая (вольфрамовая или др.) пружина заостренным концом прижимается к специально обработанной поверхности маленького германиевого кристаллика с электронной проводимостью; в небольшом объеме кристалла вблизи контакта образуется дырочная проводимость.

В плоскостном диоде (*б*) на пластинку германия наплавляется индий, атомы которого проникают на небольшую глубину в германий и создают тонкий слой с дырочной проводимостью; остальная часть германиевой пластинки имеет электронную проводимость. Полупроводниковые триоды *p-n-p* имеют с одной (*в*)или с двух (*г*) сторон германиевого кристалла (основания) пружинные контакты или индиевые наплавки.

ПРИМЕНЕНИЯ ПОЛУПРОВОДНИКОВ

Перечислим важнейшие применения полупроводников:

1) *полупроводниковые диоды и триоды* с большим успехом заменяют электронные лампы, так как они более экономичны, компактны, отличаются простотой устройства, надежностью, механической прочностью и большим сроком работы. Применяемые в электро- и радиотехнике селеновые выпрямители имеют к. п. д. до 70%, германиевые — до 98%. Имеются выпрямители, работающие при высоких температурах. Полупроводниковые триоды имеют к. п. д. до 50% (тогда как у вакуумных электронных ламп — около 1%). Полупроводниковые приборы употребляют мало энергии и требуют для питания низкое (по сравнению с электронными лампами) напряжение, поэтому необходимые для них источники питания могут иметь очень малые габариты. Это позволило решить ряд важных задач радиотехники (создание миниатюрных радиоприемников и передатчиков и др.);

2) *фотосопротивления* — полупроводники (селен, сернистые кадмий и свинец и др.), у которых электрическое сопротивление резко уменьшается при облучении их светом, ультрафиолетовыми, рентгеновскими и другими лучами; они используются для измерения световых потоков, освещенности, воспроизводства звука, записанного на кинопленку в различных устройствах контроля, сигнализации, автоматического регулирования и т. д. Имеются фотосопротивления, чувствительные к инфракрасному излучению;

3) *термисторы* — полупроводники (смеси окислов различных металлов: магния, никеля, титана и др.), у которых электрическое сопротивление сильно зависит от температуры; они применяются для измерения температур (в таких условиях, при которых другие способы не применимы: химически активная среда, наличие вибраций, необходимость очень малых размеров датчика и др.), для автоматического регулирования температуры, в качестве ограничителей начального значения силы тока в пусковых устройствах и т. д.;

4) *варисторы* — полупроводники (карбид кремния и др.), у которых электрическое сопротивление сильно зависит от напряженности приложенного электрического поля; применяются для защиты электрических цепей от нерегулярных высоких перенапряжений, например от грозовых разрядов.

В электрических печах вместо дорогих и недолговечных металлических спиралей используются стержни из тугоплавких полупроводников, допускающих нагрев до 1300° С. Термоэлементы, составленные из двух полупроводников с *n* и *p*-проводимостями, имеют

более высокое значение коэффициента термоэлектродвижущей силы. Они могут быть использованы как преобразователи теплоты непосредственно в электрическую энергию (с к. п. д., доходящим до 10%), а также при использовании эффекта Пельтье, для целей охлаждения (полупроводниковые термоэлектрогенераторы и холодильники).

Глава 3

ЭЛЕКТРОМАГНЕТИЗМ

§ 17. МАГНИТНОЕ ПОЛЕ

Электрические заряды, покоящиеся относительно выбранной системы отсчета, имеют вокруг себя только электрическое поле. Движущиеся же заряды (например, поток ионов в вакуумных трубках, электронов в металлах и полупроводниках и т. п.) имеют вокруг себя еще и магнитное поле.

Магнитное поле обнаруживается и изучается по его воздействию на тела и измерительные приборы. Если выбрать миниатюрные «пробные тела», то можно каждую точку магнитного поля характеризовать некоторой векторной величиной (обозначаемой **B**); например:

1) на магнитную стрелку действует механический момент, стремящийся ориентировать ее таким образом, чтобы этот момент был равен нулю. В этом положении направление вдоль оси стрелки от южного полюса к северному принимается за направление магнитного поля (т. е. вектора **B**) в данном месте. Величина же вектора **B** полагается пропорциональной максимальному механическому моменту, действующему на эту стрелку в перпендикулярном полю положении;

2) на плоский виток проводника с током также действует механический момент, зависящий от размеров витка и силы тока в нем. При одном положении витка этот момент равен нулю (в этом положении за направление поля выбирается перпендикуляр к плоскости витка), а в другом (перпендикулярном) — имеет максимальное значение (им определяется величина **B** в данном месте).

Кроме того, для экспериментального изучения магнитного поля можно воспользоваться также и другими воздействиями этого поля на тела, например: на прямолинейные участки проводников с токами, на электрическое сопротивление некоторых металлов (особенно заметное у висмута), на размеры тел (например, из никеля), на пучки электронов в вакуумных трубках и т. п. Во всех способах измерения вектора **B** необходимы «пробные тела», на которые действует магнитное поле, и измерительные приборы, отмечающие результаты этих воздействий.

При экспериментальном изучении полей важно установить, не изменяет ли внесенный прибор или пробное тело ту характеристику поля, которую мы желаем измерить. Например, при использовании магнитной стрелки или витка с током их собственное магнитное поле может накладываться на изучаемое поле и искажать его (магнитное поле пробного тела может вызвать изменение намагниченности тех

тел, магнитное поле которых мы изучаем). Следовательно, необходимо позаботиться, чтобы влияние пробных тел на изучаемое поле было бы достаточно малым. Далее необходимо быть уверенным, что свойства пробных тел не изменяются при их переносе с одного места изучаемого поля в другое. Например, намагниченность магнитной стрелки (от которой зависит действующий на нее механический момент) может измениться при переходе от слабого поля к сильному; магнитные свойства витка с током зависят от площади витка, которая, однако, может измениться в зависимости от интенсивности изучаемого поля (см. § 20), и т. п. Важно отметить, что в процессе измерения имеются две неизвестные величины: искомые характеристики изучаемого поля и поведение пробных тел в этом поле; одно из них можно найти только в том случае, если с удовлетворительной точностью известно другое.

Ввиду этого имеет смысл сначала найти такие характеристики магнитного поля, которые могут быть выбраны и рассчитаны *до производства измерений* (по значениям зарядов и их скоростей, по форме и размерам проводников и силе тока в них) независимо от выбора «пробных тел» и измерительной аппаратуры. Использование таких характеристик может оказаться весьма полезным при последующем экспериментальном изучении магнитных полей, их воздействия на тела и магнитных свойств различных веществ.

§ 18. НАПРЯЖЕННОСТЬ МАГНИТНОГО ПОЛЯ ВОКРУГ ДВИЖУЩЕГОСЯ ЗАРЯДА И ПРОВОДНИКОВ С ТОКАМИ

Допустим, что по прямолинейному бесконечно длинному проводнику очень малого сечения течет ток I. Магнитное поле вокруг проводника можно полагать прямо пропорциональным току I. Проведем через проводник две бесконечные плоскости I и II, составляющие между собой угол $d\alpha$, а в плоскости, перпендикулярной проводнику, проведем окружность некоторого радиуса r (рис. III. 49), так что $d\alpha = dl/r$. Так как магнитное поле вокруг проводника имеет цилиндрическую симметрию, то величина $(d\alpha/2\pi)\,I = \frac{I}{2\pi r}\,dl$ будет показывать, какая часть всего магнитного поля заключена между плоскостями I и II. Обратим внимание на величину

$$H = \frac{I}{2\pi r}. \tag{3.1}$$

Очевидно, что ее можно рассматривать как некоторую характеристику магнитного поля на расстоянии r от бесконечно длинного прямолинейного проводника с током I. Будем полагать ее вектором, направленным по касательной к окружности (в соответствии с правилом правого винта, ввинчиваемого по направлению тока в проводнике; рис. III. 49). Однако в общем случае проводник может иметь любую форму; поэтому для расчета вектора **H** в каждой точке магнитного поля вокруг проводника произвольной формы необходимо найти некоторый общий метод. Таким методом может быть следующий. Будем полагать,

что вектор **H** в данной точке магнитного поля вокруг проводника с током I слагается из элементарных векторов $d\mathbf{H}$, соответствующих бесконечно малым участкам проводника dl (рис. III. 50, *а*), причем величина dH должна определяться по формуле

$$dH = \frac{I\,dl\,\sin\alpha}{4\pi r^2}, \qquad (3.2)$$

а направление — по правилу правого винта в зависимости от направления тока. Формула (3.2) подобрана таким образом, чтобы

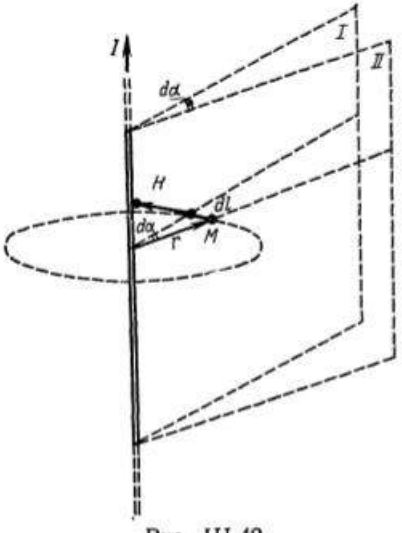


Рис. III.49

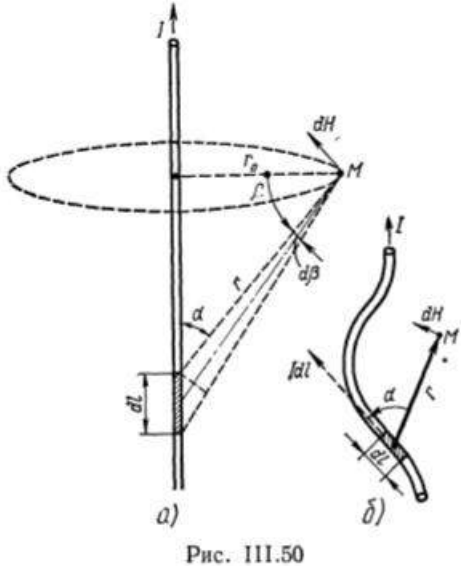


Рис. III.50

она для бесконечно длинного прямолинейного проводника приводила к формуле (3.1). Действительно (рис. III. 50, *а*),

$$dl\,\sin\alpha/r = d\beta; \quad r = r_0/\cos\beta;$$

$$H = \int_L \frac{I\,dl\,\sin\alpha}{4\pi r^2} = \int_{+\pi/2}^{-\pi/2} \frac{I}{4\pi r_0}\cos\beta\,d\beta = \frac{I}{4\pi r_0}.$$

Произведение $I d\mathbf{l}$ рассматривается как вектор, ориентированный по направлению тока, и называется *элементом тока*. Вектор **H** называется *напряженностью магнитного поля* в данной точке.

Утверждение, что расчет напряженности магнитного поля dH в любой точке вокруг элемента тока должен производиться по формуле (3.2), называется законом Био—Савара—Лапласа. Дополнительное указание, что напряженность магнитного поля в любой точке вокруг проводника произвольной формы должна определяться путем суммирования $\Sigma d\mathbf{H}_i = \mathbf{H}$, называется принципом суперпозиции. Если имеется несколько проводников различной формы и с различными токами и если $\mathbf{H}_1$, $\mathbf{H}_2$, ...

есть напряженности поля в данной точке, создаваемые каждым из этих проводников в отдельности (*при отсутствии других*), то, согласно принципу суперпозиции, суммарная напряженность поля в этой точке равна

$$\mathbf{H} = \mathbf{H}_1 + \mathbf{H}_2 + \ldots \tag{3.3}$$

Укажем на одно важное следствие из закона Био—Савара—Лапласа, которое облегчает расчеты магнитных полей. Допустим, что по проводнику произвольной формы (рис. III. 51) течет ток I. По закону Био—Савара—Лапласа, напряженность магнитного поля H, создаваемого этим током в точке M, можно рассчитать, вычислив интеграл по замкнутому контуру тока:

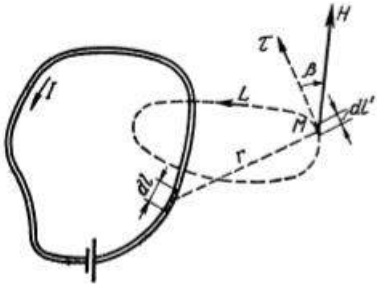


Рис. III.51

$$H = \oint \frac{I\,\mathrm{d}l \sin\alpha}{4\pi r^2}. \tag{3.4}$$

Проведем теперь в магнитном поле некоторую воображаемую замкнутую линию L (любой формы и размеров) и разделим ее на элементарные участки $\mathrm{d}l'$. Для каждого участка составим произведение $H\mathrm{d}l' \cos\beta$ (β — угол между направлением H и касательной τ к линии) и найдем их сумму вдоль всей линии L, т. е. вычислим интеграл $\oint H\mathrm{d}l' \cos\beta$. Расчет, который мы не приводим ввиду его сложности, дает весьма простой результат — для линии обхода любой формы и размеров:

$$\oint H\,\mathrm{d}l' \cos\beta = I. \tag{3.5}$$

Если изменить направление тока в проводнике, то в каждой точке поля вектор H изменит свое направление на обратное, косинусы углов β будут иметь противоположный знак и интеграл сделается отрицательным. Знак этого интеграла изменится также и при перемене направления обхода по линии L, вследствие чего изменятся направления касательных τ. Ввиду этого направление обхода и направление тока должны быть связаны между собой правилом знаков, а именно, если вращать правый винт по выбранному нами направлению обхода линии L, то его перемещение должно соответствовать положительному направлению тока I.

Полученный результат (3.5) не зависит ни от формы контура с током, ни от формы замкнутой линии L. Если линия L охватывает несколько проводников с токами I_1, I_2, ..., то, согласно принципу суперпозиции, интеграл (3.5) будет равен алгебраической сумме этих токов. Если линия L охватывает один и тот же ток n раз, то интеграл (3.5) равен nI; наконец, если L не охватывает тока, то интеграл (3.5) оказывается равным нулю, т. е. токи, протекающие за пределами контура L, в правую часть формулы (3.5) не входят.

Интеграл $\oint H\mathrm{d}l \cos\beta = \oint \mathbf{H}\mathrm{d}\mathbf{l}$ называется *циркуляцией вектора напряженности вдоль данной замкнутой линии обхода*. Утверждение,

что эта циркуляция равна алгебраической сумме сил токов, охватываемых линией обхода,

$$\oint \mathbf{H}\, d\mathbf{l} = \sum I_i, \tag{3.6}$$

называется *теоремой о циркуляции напряженности магнитного поля* или законом полного тока. Этим законом можно воспользоваться для расчета напряженности магнитного поля, причем выбирается удобный контур обхода; в частности, удобно производить обход вдоль такой линии, в каждой точке которой $\cos\beta = 1$. Такие линии называются *силовыми линиями* магнитного поля.

Магнитное поле называется *однородным*, если во всех точках поля напряженность одинакова и по величине и по направлению; поэтому поле вокруг прямолинейного проводника с током является неоднородным. При графическом изображении геометрических характеристик магнитного поля пользуются силовыми линиями, вдоль которых напряженность поля ориентирована по касательным. Силовые линии вокруг бесконечно длинного прямолинейного проводника с током имеют вид концентрических окружностей (рис. III. 52). Силовым линиям придается направление, соответствующее направлению напряженности поля; очевидно, что в однородном магнитном поле силовые линии будут параллельными прямыми. Условились, что число силовых линий, проводимых через единицу площади (перпендикулярной этим линиям), должно быть численно равно величине напряженности поля H в этом месте. Если в магнитном поле выбрана площадка ΔS, а α есть угол между нормалью к этой площадке и вектором H, то произведение

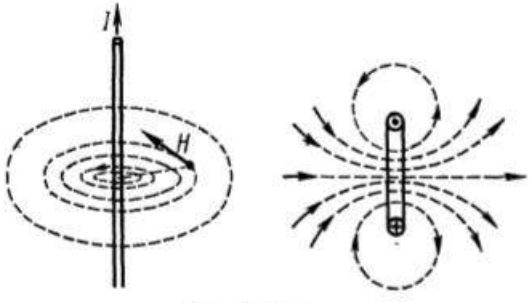


Рис. III.52

$$\Delta N = H \Delta S \cos\alpha \tag{3.7}$$

будет «числом» силовых линий, проведенных через эту площадку (это число называют также *потоком вектора напряженности через данную площадку*).

Приведем некоторые расчеты магнитных полей:

1) магнитное поле движущегося заряда. Полагая, что в элементе тока $Id\mathbf{l}$ содержится Δn электронов, имеющих скорости упорядоченного движения v, найдем напряженность поля $H = \Delta H/\Delta n$, создаваемую в данной точке одним движущимся электроном. Так как сила тока $I = nevS$ (e — заряд электрона), $n = \Delta n/(S\,dl)$, то

$$H = \frac{ev \sin\alpha}{4\pi r^2}; \tag{3.8}$$

2) напряженность магнитного поля в центре плоского кругового тока. В формуле (3.4) r будет одинаково для всех элементов

тока, расположенных по окружности (рис. III.53), $\alpha = \pi/2$, $\sin\alpha = 1$, а все векторы $d\mathbf{H}$ направлены в одну сторону, поэтому вместо их векторного сложения можно произвести интегрирование:

$$H = \int \frac{I\, dl \sin\alpha}{4\pi r^2} = \frac{I}{2r}. \tag{3.9}$$

Для «плоской катушки», имеющей n витков одинакового радиуса,

$$H = n\frac{I}{2r}.$$

При расчете напряженности поля в точке M, лежащей на оси кругового тока на расстоянии b от центра, необходимо учесть направления $d\mathbf{H}$. Расчет, который мы не приводим, показывает, что суммарная напряженность направлена вдоль оси и равна (рис. III.53)

$$H' = \frac{Ir^2}{2(r^2+b^2)^{3/2}}; \tag{3.10}$$

3) напряженность магнитного поля внутри длинного соленоида с током. Воспользуемся законом полного тока; выберем линию обхода так, чтобы участок *1—2* (рис. III.54) проходил вдоль силовой линии внутри

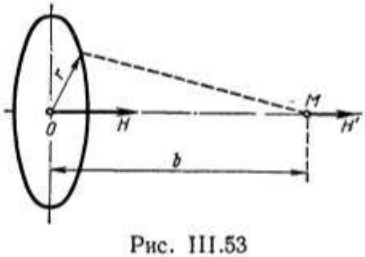


Рис. III.53

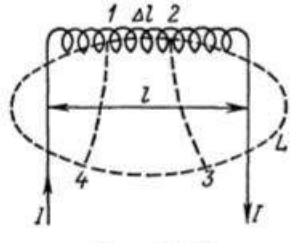


Рис. III.54

соленоида, участки *2—3* и *4—1* были бы на всем своем протяжении перпендикулярны силовым линиям, а участок *3—4* совпадал бы с силовой линией, проходящей достаточно далеко от соленоида, где напряженность поля очень мала (по сравнению с напряженностью поля внутри соленоида). Длину Δl участка *1—2*, охватывающего Δn витков, выберем такой, чтобы на протяжении этого участка величину напряженности можно было считать одинаковой; для этого плотность обмотки, т. е. число витков на единицу длины $n_1 = \Delta n/\Delta l$, должна быть достаточно большой.

Циркуляция вектора H по контуру *1—2—3—4* равна $\oint \mathbf{H}\, d\mathbf{l} = \int_1^2 \mathbf{H}\, d\mathbf{l} + \int_2^3 \mathbf{H}\, d\mathbf{l} + \int_3^4 \mathbf{H}\, d\mathbf{l} + \int_4^1 \mathbf{H}\, d\mathbf{l}$. Второй и четвертый интегралы равны нулю ввиду перпендикулярности $\mathbf{H}$ и $d\mathbf{l}$, а третьим интегралом пренебрегаем ввиду малости H вне соленоида. Тогда $\oint \mathbf{H}\, d\mathbf{l} \approx \int_1^2 H\, dl = H\,\Delta l$. На основании формулы (3.6) получаем:

$$H\,\Delta l = \sum I_i = \Delta n \cdot I; \quad H = I\,\Delta n/\Delta l = n_1 I. \tag{3.11}$$

Результат расчета не зависит от того, в каком месте сечения соленоида проходит участок *1—2*, следовательно, в любой точке этого сечения напряженность поля будет одинаковой. Таким образом, внутри очень длинного соленоида с плотной обмоткой магнитное поле можно полагать однородным (вблизи концов соленоида поле не являет-

ся однородным). Иногда в качестве линии обхода берут всю силовую линию соленоида и тогда формула

$$H = In/l = In_1 \qquad (3.12)$$

(где l — длина соленоида; n — полное число витков) даст среднюю напряженность магнитного поля внутри соленоида;

4) напряженность магнитного поля внутри толстых проводников с током. Для расчета напряженности на расстоянии r от оси проводника выберем линию обхода в виде окружности радиуса r (рис. III.55) с центром на оси проводника. Если проводник прямолинейный и бесконечно длинный, то вдоль этой линии обхода напряженность магнитного поля будет везде одинакова и в каждой точке направлена по касательной ($\cos\beta = 1$). Тогда $\oint \mathbf{H}d\mathbf{l} = H \cdot 2\pi r$. Эта линия охватывает площадь $S = \pi r^2$. Если плотность тока в различных местах сечения проводника одинакова и равна j, то суммарный ток, проходящий через S и, следовательно, охватываемый линией обхода, равен jS. Тогда на основании формулы (3.6)

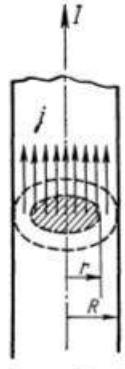


Рис. III.55

$$H \cdot 2\pi r = j\pi r^2; \quad H = \frac{1}{2} jr. \qquad (3.13)$$

Так как $j = \frac{I}{\pi R^2}$, то

$$H = \frac{I}{2\pi R^2} r.$$

Таким образом, на оси проводника ($r = 0$) напряженность поля $H = 0$, а по мере удаления от оси растет прямо пропорционально расстоянию вплоть до боковой поверхности проводника. В точках, лежащих за пределами объема проводника, применение соотношения (3.6) приводит к формуле (3,1), согласно которой напряженность магнитного поля обратно пропорциональна расстоянию от оси проводника.

Выше рассматривались магнитные поля вокруг потоков электрических зарядов — внутри проводников или в вакууме; пользуясь законом Био—Савара—Лапласа, можно рассчитать напряженность в любой точке их полей, т. е. найти функции $H = H(x, y, z)$. Однако магнитные поля создаются также и связанными зарядами (электронами и ионами), движущимися в атомах и молекулах вещества. Если электрон описывает замкнутую траекторию, то он создает вокруг себя магнитное поле, которое аналогично магнитному полю проводника такой же формы и размеров, как и траектория, с силой тока в нем $I = -ne$, где n — число оборотов в секунду по этой орбите (n показывает, сколько раз в секунду электрон проходит через одну и ту же точку орбиты, поэтому ne есть количество электричества, проходящее через эту точку в секунду, т. е. сила тока).

Токи, вызванные движением зарядов в атомах и молекулах вещества, называются *молекулярными*. Так как напряженности полей складываются векторно, то при беспорядочной ориентации молекулярных токов какого-нибудь тела общее поле вокруг этого тела будет практически отсутствовать. В телах, имеющих вокруг себя магнитное поле (постоянные магниты, электромагниты), молекулярные токи имеют некоторую упорядоченность; однако расчет общего магнитного поля этих токов затруднен, так как неизвестны ни силы этих токов, ни их расположение и ориентации в объеме тел; кроме

того, форма и размеры намагниченных тел могут быть различными. Поэтому при изучении магнитного поля вокруг намагниченных тел прибегают к некоторым упрощенным приемам (см. § 25, 26).

§ 19. ДЕЙСТВИЕ МАГНИТНОГО ПОЛЯ НА ЗАРЯДЫ И ПРОВОДНИКИ С ТОКОМ; СИЛА ЛОРЕНЦА И ЗАКОН АМПЕРА. ИНДУКЦИЯ МАГНИТНОГО ПОЛЯ

Наблюдения показывают, что постоянное магнитное поле действует только на движущиеся заряды; переменное же магнитное поле действует также и на покоящиеся заряды. Однако необходимо прежде всего уточнить, что понимается под «постоянным» и «переменным» магнитным полем и как определить движение заряда относительно магнитного поля.

Магнитное поле всегда связано с определенными телами, например с системой проводников, по которым текут электрические токи или с намагниченными телами из железа или магнитных сплавов. Каждая интересующая нас точка магнитного поля имеет определенное расположение относительно этих тел. Если напряженности во всех точках магнитного поля с течением времени не изменяются, то поле называют постоянным; если же напряженности магнитного поля со временем меняют свои значения, то поле называют переменным. В частности, вокруг проводника, по которому течет переменный ток, существует переменное магнитное поле; напряженность поля в различных местах увеличивается и уменьшается вместе с увеличением и уменьшением силы тока.

Связывая магнитные поля с определенными телами, можно утверждать, что электрические заряды покоятся или движутся относительно магнитных полей, если они покоятся или движутся относительно тел, которые образуют эти поля. Выбрав какую-нибудь систему отсчета, можно магнитное поле назвать «движущимся» или «покоящимся», если движется или покоится тело, с которым связано это поле. При этом движение электрического заряда относительно неподвижного магнитного поля неотличимо от движения магнитного поля относительно неподвижного заряда. Заметим также, что *относительное движение заряда и неоднородного магнитного поля эквивалентно нахождению неподвижного заряда в переменном магнитном поле.*

Ниже рассматривается только действие постоянного магнитного поля на движущиеся заряды и проводники с током. Это действие определяется на основании следующих утверждений:

1) на заряд e, движущийся со скоростью v в постоянном магнитном поле напряженностью H, действует сила (называемая *силой Лоренца*), равная (в вакууме)

$$F = \mu_0 evH \sin\alpha, \tag{3.14}$$

где α — угол между направлениями скорости и напряженности, а μ_0 — коэффициент, зависящий от выбора единиц измерения силы, заряда, скорости и напряженности. Направление силы $\mathbf{F}$ зависит от

знака заряда и перпендикулярно плоскости, в которой лежат векторы **v** и **H**. Сила Лоренца записывается в виде векторного произведения (рис. III. 56, *а*):

$$\mathbf{F} = \mu_0 e [\mathbf{vH}];$$

2) на прямолинейный участок длиной dl проводника с током I в постоянном магнитном поле в вакууме действует сила (закон Ампера)

$$dF = \mu_0 HI\, dl \sin\alpha, \qquad (3.15)$$

где α — угол между направлениями тока и напряженности поля. Сила $d\mathbf{F}$ перпендикулярна плоскости, в которой лежат векторы элемента тока $I d\mathbf{l}$ и напряженности поля; в векторной записи (рис. III.56, *б*)

$$d\mathbf{F} = \mu_0 I [d\mathbf{l}\mathbf{H}].$$

Формула (3.15) может быть получена из выражения (3.14), и, наоборот, зная закон Ампера, можно получить силу Лоренца. Обозначим

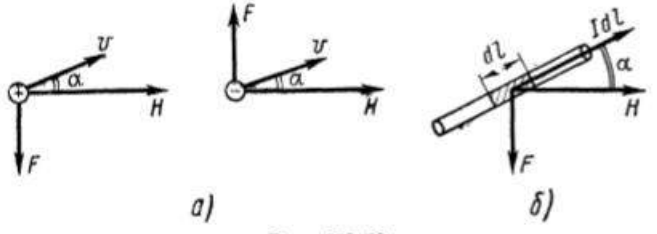


Рис. III.56

через n число упорядоченно движущихся электронов в единице объема проводника, v — скорость этого движения, S — площадь сечения проводника. Тогда сила тока $I = nevS$, а произведение $Id\mathbf{l}$ («элемент тока») равно $nevSd\mathbf{l} = Nev$, где N — число упорядоченно движущихся электронов в объеме участка тока. Подставив $Id\mathbf{l} = Nev$ в формулу (3.15) и разделив на N, получим силу, действующую на один электрон (см. выражение (3.14)). При этом предполагается, что скорости упорядоченного движения электронов параллельны между собой, а поле в пределах объема проводника является однородным; эти условия соблюдаются, если проводник достаточно тонкий. Значение μ_0 содержится в определении ампера:

ампер — сила неизменяющегося тока, который, проходя по двум параллельным прямолинейным проводникам бесконечной длины и ничтожно малого кругового сечения, расположенным на расстоянии 1 м *один от другого в вакууме, вызывает между этими проводниками силу, равную* $2 \cdot 10^{-7}$ Н *на каждый метр длины.*

Так как у бесконечно длинного прямолинейного проводника с током напряженность магнитного поля на расстоянии r (рис. III. 57) всюду одинакова и равна $H_1 = I_1/2\pi r$, то на участок l второго (парал-

лельного) проводника, по закону Ампера, будет действовать сила F:

$$F = \mu_0 H_1 I_2 l = \mu_0 \frac{I_1 I_2}{2\pi r} l. \qquad (3.16)$$

Такая же сила действует на равный участок первого проводника со стороны магнитного поля второго проводника. Следует подчеркнуть, что формула (3.16) показывает не силу взаимодействия между рассматриваемыми участками проводников, а силу, с которой прямолинейный *бесконечно длинный* проводник с током I_1 действует на прямолинейный участок конечной длины l другого проводника с током I_2.

Воспользовавшись определением ампера, получим ($F = 2 \cdot 10^{-7}$ Н, $I_1 = I_2 = 1$А, $r = 1$ м, $l = 1$ м)

$$\mu_0 = 4\pi \cdot 10^{-7} \text{ Н/А}^2.$$

Измерения показывают, что при заполнении магнитного поля веществом, обладающим магнитными свойствами, силы, действующие на заряды и токи, увеличиваются в μ раз, поэтому формула Лоренца (3.14) и закон Ампера (3.15) принимают вид:

$$d\mathbf{F} = \mu_0 \mu I [d\mathbf{l} \cdot \mathbf{H}];$$

$$\mathbf{F} = \mu_0 \mu e [\mathbf{v} \cdot \mathbf{H}], \qquad (3.17)$$

а сила взаимодействия (3.16) для параллельных токов

$$F = \mu_0 \mu \frac{I_1 I_2}{2\pi r} l. \qquad (3.18)$$

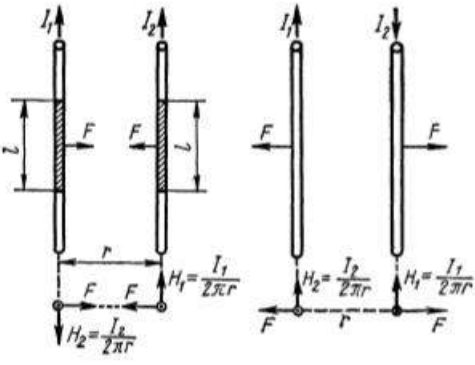


Рис. III.57

Величина μ_0 называется *магнитной постоянной*, а также магнитной проницаемостью вакуума. Произведение $\mu_0\mu$ называется *абсолютной магнитной проницаемостью* данной среды, а безразмерная величина μ показывающая, во сколько раз сила, действующая на движущиеся заряды и проводники с током в данной среде, больше, чем в вакууме, называется *относительной магнитной проницаемостью* данной среды (по отношению к вакууму); она приводится в соответствующих таблицах.

Формулы (3.17) можно переписать в другом виде, если для характеристики каждой точки магнитного поля ввести новый вектор

$$\mathbf{B} = \mu_0 \mu \mathbf{H}, \qquad (3.19)$$

называемый *магнитной индукцией*. Тогда

$$\mathbf{F} = e[\mathbf{vB}]; \quad d\mathbf{F} = I[d\mathbf{lB}]. \qquad (3.20)$$

Заметим, что, измеряя силу, действующую на движущийся заряд или на проводник с током, мы находим непосредственно не напря-

женность **H**, а произведение $\mu_0\mu\mathbf{H}$, т. е. индукцию поля. Более того, можно показать, что использование любых «пробных тел» для изучения магнитного поля (см. § 17) приводит к нахождению не напряженности **H**, а индукции **B**. Ввиду этого существует тенденция исключить понятие напряженности магнитного поля и характеризовать это поле только одной (измеряемой) величиной — индукцией **B**. По этому вопросу заметим следующее:

1) *напряженность* магнитного поля в каждой его точке, согласно определению, *рассчитывается* по закону Био—Савара—Лапласа с применением принципа суперпозиции. Для этого должны быть известны заряды и скорости их движения или форма и размеры проводников и протекающие по ним токи;

2) *индукция* же магнитного поля, согласно определению, должна быть *измерена* по воздействию этого поля на «пробные тела». Ввиду этого какие-либо расчетные формулы для индукции поля B могут быть получены только эмпирическим путем и будут содержать все неточности, с которыми связаны измерительные процедуры. Однако соотношение (3.19) позволяет рассчитывать индукцию магнитного поля B вокруг движущихся зарядов и проводников с токами по тем же формулам (например, (3.8) — (3.13)), которые были получены для напряженности поля. Например, умножив формулу (3.12) на $\mu_0\mu$, получим для индукции поля внутри длинного соленоида с током I, имеющего n одинаковых витков на метр длины, расчетную формулу

$$B = \mu_0\mu I n_1. \qquad (3.21)$$

Однако для расчета индукции поля по приведенным в § 18 формулам для напряженности необходимо *заранее знать* магнитную проницаемость среды в каждой точке, в которой вычисляется индукция. Величину μ для данного вещества можно найти по измеренным значениям индукции в вакууме (B_0) и внутри этого вещества (B):

$$\mu = B/B_0,$$

либо же путем сопоставления *измеренного* значения индукции B с *рассчитанным* значением напряженности H магнитного поля, действующего на это вещество:

$$\mu = B/\mu_0 H.$$

Второй способ имеет некоторые преимущества перед первым, поэтому понятие напряженности магнитного поля продолжает использоваться в физике и электротехнике.

По своим магнитным свойствам вещество может быть *изотропным* (если магнитная проницаемость μ одинакова по всем направлениям) или *анизотропным* (если μ различна в различных направлениях, например в кристаллах), *однородным* или *неоднородным*, в зависимости от того, одинаковы или различны значения μ в пределах объема данного тела.

Стандартной единицей напряженности магнитного поля принята напряженность в центре длинного соленоида с плотной равномерно распределенной обмоткой (из одинаковых витков), по которой пропущен

ток силой $1/n$ ампер, где n — число витков, приходящихся на один метр длины соленоида. Согласно формуле (3.12), этой единицей будет ампер на метр:

$$H = In_1; \quad [H] = \text{А/м}.$$

Единица измерения магнитной индукции — тесла — обозначается Т (стандартное определение тесла приведено в § 21). Тесла — это индукция однородного магнитного поля, которое действует на каждый метр прямолинейного проводника с током в один ампер, расположенного перпендикулярно к полю, с силой в один ньютон:

$$[B] = \frac{[F]}{[I]\,[l]}; \quad \text{Т} = \frac{\text{Н}}{\text{А} \cdot \text{м}}.$$

В § 18 указывалось на необходимость применения принципа суперпозиции для расчета напряженности магнитного поля; выясним, каким образом подтверждается этот принцип и применим ли он для вектора **B**. Допустим, что имеется несколько отдельных контуров, по которым можно пропускать токи I_1, I_2, ... Сначала пропустим ток только по одному контуру и измерим силу $\mathbf{F}_1$, с которой поле этого тока действует на выбранное нами «пробное тело», например на элемент тока $I_0 \Delta l$. Согласно закону Ампера,

$$\mathbf{F}_1 = I_0 [\Delta l \mathbf{B}_1] = \mu_0 \mu_1 I_0 [\Delta l \mathbf{H}_1].$$

Затем, выключив ток I_1, пропустим ток I_2 через второй контур и вновь измерим силу $\mathbf{F}_2$, действующую на «пробное тело». Полагая, что относительная магнитная проницаемость среды μ может зависеть от интенсивности вызванного в ней поля, вместо μ_1 напишем μ_2:

$$\mathbf{F}_2 = I_0 [\Delta l \mathbf{B}_2] = \mu_0 \mu_2 I_0 [\Delta l \mathbf{H}_2].$$

Повторим измерение для каждого из токов в отдельности (взаимное расположение контуров и пробного тела сохраняется неизменным) и затем пропустим эти токи одновременно по всем контурам; сила **F**, действующая на «пробное тело» со стороны суммарного магнитного поля, равна

$$\mathbf{F} = I_0 [\Delta l \mathbf{B}] = \mu_0 \mu I_0 [\Delta l \mathbf{H}].$$

Измерения показывают, что в вакууме ($\mu_1 = \mu_2 = \ldots = \mu = 1$)

$$\mathbf{F} = \mathbf{F}_1 + \mathbf{F}_2 + \ldots$$

Подставив значение сил, получим (после сокращения) принцип суперпозиции для вектора **H**:

$$\mathbf{H} = \mathbf{H}_1 + \mathbf{H}_2 + \ldots$$

Очевидно, что принцип суперпозиции будет соблюдаться и для вектора индукции в вакууме: если $\mathbf{B}_{01} = \mu_0 \mathbf{H}_1$ есть индукция поля, создаваемого в вакууме первым контуром в отсутствии остальных, $\mathbf{B}_{02} = \mu_0 \mathbf{H}_2$ — то же для второго контура и т. д., а $\mathbf{B}_0 = \mu_0 \mathbf{H}$ есть индукция суммарного магнитного поля, то

$$\mathbf{B}_0 = \mathbf{B}_{01} + \mathbf{B}_{02} + \ldots$$

Допустим теперь, что магнитное поле образовано в среде, в которой магнитная проницаемость μ не есть определенная величина, а зависит от H и B. Тогда

$$\mathbf{H}_1 = \mathbf{B}_1/\mu_0\mu_1; \quad \mathbf{H}_2 = \mathbf{B}_2/\mu_0\mu_2; \quad \ldots; \quad \mathbf{H} = \mathbf{B}/\mu_0\mu$$

и из принципа суперпозиции для **H** следует, что $\mathbf{B} \neq \mathbf{B}_1 + \mathbf{B}_2 + \ldots$

В *изотропной* среде векторы **H** и **B** имеют одинаковое направление; в *анизотропной* среде μ имеет различные значения в различных направлениях и поэтому направления **H** и **B** могут не совпадать (аналогичные рассуждения для E и D приведены в § 2).

Указанные выше и в § 18 расчетные формулы для индукции и напряженности магнитного поля даны в Международной системе

единиц (СИ). В физике употребляется также «абсолютная электромагнитная система единиц (СГС)», в которой магнитная проницаемость полагается безразмерной величиной, равной для вакуума единице, а расчетные формулы для H отличаются от приведенных в § 18 множителем 4π. В этой системе единица силы тока, так же как и в СИ, определяется по взаимодействию бесконечных параллельных прямолинейных проводников с токами: сила тока равна одной абсолютной электромагнитной единице — 1 СГС (I), если сила взаимодействия между этими проводниками в вакууме при расстоянии между ними в 2 см равна одной дине на каждый сантиметр их длины. Напряженность поля в этой системе измеряется в *эрстедах* (Э); это есть напряженность поля на расстоянии 2 см от бесконечного прямолинейного проводника, по которому течет ток, равный 1 СГС (I). Индукция поля измеряется в *гауссах* (Гс); это есть индукция магнитного поля, которая действует на прямолинейный проводник с током в 1 СГС (I), расположенный перпендикулярно полю, с силой в одну дину на каждый сантиметр длины.

Сопоставляя определения для I, H и B в обеих системах, получим:

$$1\ \text{СГС}\,(I) = 10\ \text{А};\quad 1\ \text{Э} = \frac{1}{4\pi}\cdot 10^3\ \text{А/м};$$
$$1\ \text{Гс} = 10^{-4}\ \text{Т}.$$

Ниже приводится сводка основных расчетных формул в СИ и СГС:

	СИ	СГС
Закон Био—Савара—Лапласа	$\mathrm{d}H = \dfrac{I\,\mathrm{d}l\sin\alpha}{4\pi r^2}$	$\mathrm{d}H = \dfrac{I\,\mathrm{d}l\sin\alpha}{r^2}$
Напряженность поля на расстоянии r от бесконечно длинного прямолинейного проводника с током	$H = \dfrac{I}{2\pi r}$	$H = \dfrac{2I}{r}$
Напряженность поля в центре кругового тока	$H = \dfrac{I}{2r}$	$H = \dfrac{2\pi I}{r}$
Напряженность поля внутри длинного соленоида, содержащего n витков на длине l	$H = \dfrac{In}{l} = In_1$	$H = \dfrac{4\pi In}{l} = 4\pi In_1$
Теорема о циркуляции (закон полного тока)	$\oint H\,\mathrm{d}l\cos\beta = \sum I_i$	$\oint H\,\mathrm{d}l\cos\beta = 4\pi\sum I_i$
Связь между индукцией и напряженностью магнитного поля	$\mathbf{B} = \mu_0\mu\mathbf{H}$	$\mathbf{B} = \mu\mathbf{H}$
Сила Лоренца	$\mathbf{F} = e\,[\mathbf{vB}]$	
Закон Ампера	$\mathrm{d}\mathbf{F} = I\,[\mathrm{d}\mathbf{l}\mathbf{B}]$	

Сила Лоренца и закон Ампера имеют в обеих системах одинаковый вид, поэтому все формулы, по которым рассчитываются силы, действующие в магнитном поле на заряды и токи, также будут в обеих системах одинаковыми. Например, сила взаимодействия между бесконечными прямолинейными проводниками с токами I_1 и I_2, приходящаяся на длину l, имеет в обеих системах вид (см. формулу (3.16))

$$F = I_1 B_2 l = I_2 B_1 l,$$

где B_1 — индукция магнитного поля первого проводника на том расстоянии, на котором находится второй проводник; B_2 — индукция поля второго проводника, действующая на первый. Однако если в этой формуле выразить индукцию поля B через напряженность H или через силу тока, создающего поле, то ввиду различия в выражениях для H сила F будет иметь в СИ и в СГС различный вид:

$$F = \mu_0 \mu \frac{I_1 I_2}{2\pi r} l; \quad F = \mu \frac{2 I_1 I_2}{r} l.$$

Во избежание такого различия следует силы, действующие со стороны магнитного поля на движущиеся заряды или на проводники с током, всегда выражать через индукцию этого поля.

§ 20. ДЕЙСТВИЕ МАГНИТНОГО ПОЛЯ НА КОНТУР С ТОКОМ. ДВИЖЕНИЕ ЗАРЯДА В МАГНИТНОМ ПОЛЕ

КОНТУР С ТОКОМ В МАГНИТНОМ ПОЛЕ

Рассмотрим действие магнитного поля на замкнутый проводник с током. Если проводник криволинейный и не лежит в одной плоскости, а также если магнитное поле неоднородное, то механическое действие поля на весь контур следует определить, вычислив и затем

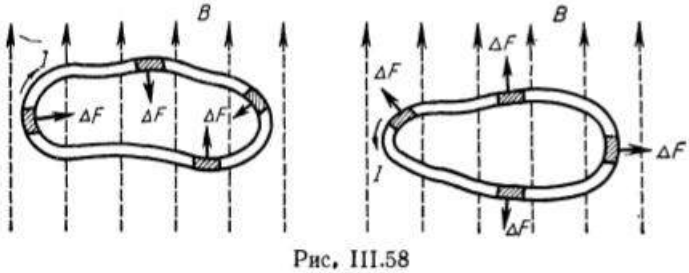


Рис. III.58

сложив силы, приложенные к отдельным коротким его участкам (которые можно полагать прямолинейными и находящимися в однородных областях поля). Расчет упрощается, если элементы контура лежат в одной плоскости. Допустим, что эта плоскость перпендикулярна направлению поля (рис. III. 58). Тогда элементарные силы ΔF, действующие на отдельные участки контура, будут лежать в этой же плоскости. Если поле всюду однородное, то эти силы только де-

формируют контур (сжимают или растягивают его в зависимости от направления тока), а равнодействующая этих сил, как нетрудно показать, равна нулю. Если же магнитное поле неоднородное, то равнодействующая элементарных сил не равна нулю; контур не только деформируется, но и перемещается в ту область, куда направлена равнодействующая.

Рассчитаем эту равнодействующую для прямоугольной рамки с током в неоднородном поле, в котором вектор **B** возрастает вдоль оси OX по линейному закону, так что (рис. III. 59)

$$B_2 - B_1 = \frac{\partial B}{\partial x} a,$$

где $\frac{\partial B}{\partial x}$ — скорость изменения B вдоль оси OX. Тогда $F_2 > F_1$ и разность

$$F = F_2 - F_1 = IbB_2 - IbB_1 = \frac{\partial B}{\partial x} Iab = \frac{\partial B}{\partial x} IS, \qquad (3.22)$$

где $S = ab$ — площадь рамки. Условимся изображать контур с током некоторым вектором $\mathbf{p}_м$, численно равным IS и направленным перпендикулярно плоскости контура в соответствии с правилом правого винта (рис. III. 59). Тогда можно

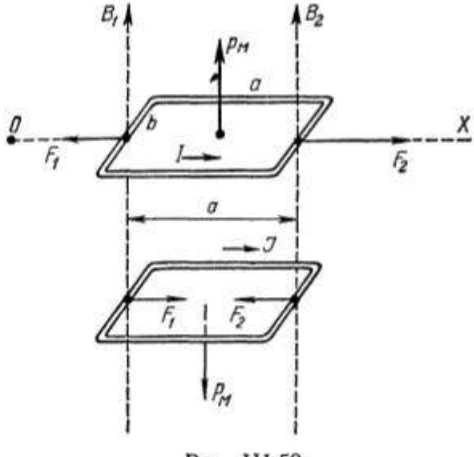


Рис. III.59

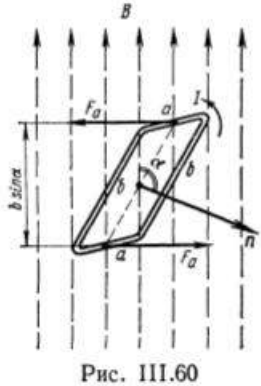


Рис. III.60

отметить, что при $\mathbf{p}_м \uparrow\uparrow \mathbf{B}$ на контур действует сила, направленная в сторону возрастания B, т. е. контур будет *втягиваться* в поле. Если же $\mathbf{p}_м \uparrow\downarrow \mathbf{B}$, то на контур действует сила, направленная в сторону убывания B, т. е. такой контур будет *выталкиваться* из поля.

Допустим теперь, что плоскость контура составляет некоторый угол α с направлением поля. Для простоты рассуждений предположим, что контур представляет собой рамку с размерами сторон a и b, причем стороны a перпендикулярны вектору B (рис. III. 60). Тогда на стороны a будут действовать силы $F_a = IBa$, составляющие пару с моментом

$$M = F_a b \sin\alpha = IBab \sin\alpha = BIS \sin\alpha, \qquad (3.23)$$

где $S = ab$ — площадь рамки. На стороны b действуют силы, лежащие в плоскости рамки и деформирующие ее. Таким образом, *на замкнутый контур с током, помещенный в однородное магнитное поле, действует вращающий момент,* максимальное значение которого равно произведению силы тока, индукции поля и площади, охватываемой контуром. Этот момент стремится повернуть контур с током так, чтобы плоскость контура была перпендикулярна направлению индукции поля. В более общем случае, когда поле неоднородное, а контур с током имеет произвольную форму, магнитное поле деформирует контур, поворачивает его и перемещает к областям с большей индукцией **B**. Если контур с током представляет собой катушку или соленоид с n одинаковыми витками, то вращающий момент (см. формулу (3.23)) следует умножить на n.

Работа, совершаемая моментом M при повороте рамки (или контура) с током на угол $\mathrm{d}\alpha$,

$$\mathrm{d}A = M\,\mathrm{d}\alpha = IBS\sin\mathrm{d}\alpha = I\mathrm{d}(BS\cos\alpha).$$

Произведение

$$\Phi = BS\cos\alpha, \tag{3.24}$$

где α — угол между направлением вектора **B** и нормали **n** к площадке S, называется *магнитным потоком* через площадку S (эту величину часто называют «потоком магнитной индукции»). Если магнитное поле графически изображать силовыми линиями вектора B, то Φ есть число силовых линий индукции, проходящих (или, лучше, проведенных) через площадку S. Для неоднородного поля

$$\mathrm{d}\Phi = B\cos\alpha\,\mathrm{d}S; \quad \Phi = \int\limits_S B\cos\alpha\,\mathrm{d}S.$$

Таким образом, работа, совершаемая моментом M, действующим на контур с током, при повороте на угол $\Delta\alpha$ может быть выражена через изменение магнитного потока:

$$A = \int\limits_{\alpha}^{\alpha+\Delta\alpha} M\,\mathrm{d}\alpha = \int\limits_{\Phi_1}^{\Phi_2} I\,\mathrm{d}\Phi. \tag{3.25}$$

Если $I = \mathrm{const}$, то $A = I\Delta\Phi$, где $\Delta\Phi$ — изменение магнитного потока, охватываемого рамкой (контуром). В этой формуле работа выражается в джоулях, сила тока — в амперах, а для магнитного потока введена единица измерения — *вебер* (Вб); если в однородном магнитном поле при повороте одиночного замкнутого контура с током в один ампер совершается работа в один джоуль, то изменение магнитного потока, охватываемого контуром, равно одному веберу: $1\ \text{Вб} = 1\ \text{Дж/А} = 1\ \text{Н}\cdot\text{м/А}$.

Сопоставляя с определением единицы магнитной индукции — тесла (см. § 19), получаем $1\ \text{Вб} = 1\ \text{Т}\cdot\text{м}^2$. В системе СГС магнитный поток выражается в максвеллах:

$$1\ \text{максвелл} = 1\ \text{Гс}\cdot 1\ \text{см}^2 = 10^{-8}\ \text{Вб}.$$

Произведение IS или для соленоида с n одинаковыми витками

$$ISn = p_м \tag{3.26}$$

называют *магнитным моментом замкнутого контура* или катушки (соленоида) с током. Представим этот момент в виде вектора $\mathbf{p}_м$, ориентированного перпендикулярно площадке S (в направлении, куда двигался бы правый винт, вращаемый по направлению тока); тогда формула (3.23) может быть записана в виде векторного произведения:

$$\mathbf{M} = [\mathbf{p}_м \mathbf{B}]. \tag{3.27}$$

Магнитный момент выражается в амперах, умноженных на квадратный метр (А · м²), или в джоулях, деленных на тесла (Дж/Т).

Формула (3.27) используется для определения магнитной индукции и единицы ее измерения (см. также § 19): магнитная индукция есть величина, равная отношению максимального механического момента $M_{макс}$, действующего на контур с током, к магнитному моменту $p_м$ этого контура:

$$B = M_{макс}/p_м.$$

Так как механический момент выражается в Н·м, а магнитный момент контура — в А · м², то единица магнитной индукции — тесла — будет выражаться в Н/(А · м).

Допустим, что свободная рама с магнитным моментом $p_м = ISn$ (т. е. система из миниатюрного источника тока, замкнутого плоской катушкой с n витками) вносится в магнитное поле B, причем первоначальное значение угла α (см. формулу (3.25)) не равно нулю: $\alpha = \alpha_0$. Под действием механического момента M рамка будет поворачиваться таким образом, чтобы угол α уменьшался до нуля. Однако совершаемая при этом работа A может пойти только на сообщение рамке энергии вращения, поэтому в момент, когда $\alpha = 0$, $M = 0$, рамка будет иметь некоторую кинетическую энергию вращения, равную $W = \int_{\alpha}^{0} M \, d\alpha$.

Вследствие этого движение рамки не прекратится и вектор магнитного момента $p_м$ начнет отклоняться от направления $\mathbf{B}$ в другую сторону, до достижения угла α_0. Затем начнется новое уменьшение α и процесс будет повторяться с некоторой периодичностью T или частотой $\nu = 1/T$. Таким образом, свободный контур с магнитным моментом $p_м$ приобретает в внешнем магнитном поле $\mathbf{B}$ некоторую энергию W, зависящую от угла α_0 между векторами $\mathbf{B}$ и $\mathbf{p}_м$; максимальное значение этой энергии (при $\alpha_0 = \pi/2$) равно

$$W = \int_{\pi/2}^{0} IBSn \sin\alpha \, d\alpha = ISnB = p_м B.$$

МАГНИТНЫЙ МОМЕНТ ЭЛЕКТРОНА

Эти рассуждения применимы и для любого тела, обладающего магнитным моментом. Таким телом может быть, например, электрон, вращающийся по какой-либо орбите (например, внутри атомов и мо-

лекул). Для расчета этого момента предположим, что орбита есть окружность радиуса r, по которой электрон движется с постоянной по величине скоростью v. Сила тока, соответствующая этому движению, равна заряду электрона e, умноженному на число оборотов в единицу времени $n = v/2\pi r$, следовательно,

$$p_м = IS = en\pi r^2 = \frac{1}{2} evr. \tag{3.28}$$

Умножив и разделив это выражение на массу электрона, получим, что *магнитный момент электронной орбиты*

$$p_м = \frac{e}{2m} mvr = \frac{e}{2m} p_{мех}$$

прямо пропорционален механическому моменту импульса электрона на этой орбите ($p_{мех} = mvr$).

Установлено, что кроме магнитного момента, обусловленного вращением по орбите, электрон обладает собственным магнитным моментом, равным $\mu_e = 9{,}28 \cdot 10^{-24}$ Дж/Т. Этот момент связывается с собственным (опять-таки не связанным с вращением по орбите) механическим моментом $p_{мех}$, причем оказалось, что отношение

$$\frac{\mu_e}{p_{мех}} = \frac{e}{m}$$

вдвое больше, чем для электронной орбиты. В дальнейшем выяснилось, что собственные магнитные и механические моменты существуют также у протона, нейтрона, атомных ядер и у других элементарных частиц.

Собственный механический момент элементарной частицы (не обусловленной движением по какой-либо траектории) называется *спином*, а связанный с ним собственный магнитный момент — *спиновым магнитным моментом*. Предположение о том, что спин электрона (или другой элементарной частицы) может быть обусловлен его вращением вокруг своей оси (подобно моменту импульса вращающегося волчка), а магнитный момент — замкнутыми токами, образованными зарядами частицы при этом вращении, оказалось неприемлемым. Установлено, что спин и спиновой магнитный момент не есть «классическое» свойство элементарных частиц, которое можно было бы изменять, действуя на них различными силами и вращающими моментами, а есть их фундаментальное свойство, которое, подобно электрическому заряду, остается без изменений в любых условиях и при любых воздействиях.

Выше было показано, что рамка с током, а следовательно, и любое тело, обладающее магнитным моментом $p_м$, приобретают в внешнем магнитном поле B дополнительную энергию, равную $p_м B$. Этот результат относится также и к физическим системам, состоящим из элементарных частиц (атом, атомные ядра, молекулы). Частота, с которой происходят колебания таких тел, например атомных ядер в внешнем магнитном поле, используется для измерения их магнитных моментов.

ДВИЖЕНИЕ ЗАРЯДА В МАГНИТНОМ ПОЛЕ

Рассмотрим движение заряда в магнитном поле. Так как сила Лоренца (см. рис. III. 56) F всегда перпендикулярна скорости движения заряда, то при движении заряда в магнитном поле эта сила работы не совершает и, следовательно, скорость движения заряда по величине не изменяется. Сила F будет изменять только направление скорости, заставляя заряд описывать криволинейную траекторию. Приравнивая силу Лоренца $F = \mu_0 evH \sin\alpha = evB \sin\alpha$ центростремительной силе $F = mv^2/R$, можно рассчитать радиус кривизны R траектории заряда в том месте, где существует магнитное поле B:

$$evB \sin\alpha = mv^2/R; \quad R = mv/eB \sin\alpha,$$

где m — масса движущегося заряда. Если магнитное поле однородное и постоянное, то радиус кривизны траектории — постоянная величина. Если составляющая скорости вдоль направления поля $v \cos\alpha$ отлична от нуля, то заряд будет описывать винтовую линию вокруг оси, ориентированной по направлению поля. При $\alpha = \pi/2$, $\sin\alpha = 1$, $\cos\alpha = 0$ заряд будет описывать окружность, плоскость которой перпендикулярна направлению поля. Частота обращения вокруг направления магнитного поля

$$\omega = \frac{v}{R} = \frac{e}{m} B$$

одинакова для электронов, имеющих различные скорости.

ЭФФЕКТ ХОЛЛА

Отклонение электронов (и других заряженных частиц) в магнитных полях широко используется в технике (ускорители заряженных частиц, электронные микроскопы, измерительные приборы для изу-

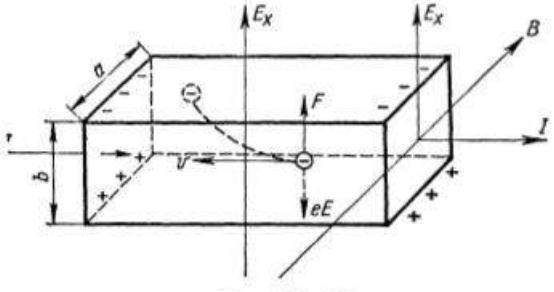


Рис. III. 61

чения колебательных и других быстропротекающих процессов, телевизоры, магнитные ловушки для заряженных частиц и т. п.).

Магнитное поле действует также и на те электроны внутри проводников, упорядоченное движение которых образует электрический ток. Допустим, что проводник (рис. III.61) с током I находится

в однородном магнитном поле с индукцией B. На каждый электрон, имеющий скорость упорядоченного движения v, действует сила Лоренца $F = evB$, перпендикулярная v и B. Вследствие этого электроны будут описывать криволинейные траектории, показанные пунктиром; на верхней поверхности проводника появится избыточное количество электронов и она зарядится отрицательно, а между нижней и верхней поверхностями образуется некоторая разность потенциалов (эффект Холла). Напряженность электрического поля между этими поверхностями E_x перпендикулярна v, поэтому на электрон действует сила eE_x, направленная против силы Лоренца. В установившемся состоянии должно соблюдаться равенство $eE_x = evB$. Так как сила тока через проводник равна $I = nevS$ (n — число перемещающихся электронов в единице объема проводника, $S = ab$ — его сечение), то

$$E_x = \frac{1}{ne} B \frac{I}{S},$$

а разность потенциалов между верхней и нижней поверхностями (считая поле E_x однородным) равна

$$\varphi_1 - \varphi_2 = E_x b = \frac{1}{ne} \frac{IB}{a}.$$

Величина $1/ne = R$ называется *коэффициентом Холла*. Точный расчет, учитывающий взаимодействие электронов с кристаллической решеткой проводника, дает поправочный множитель A:

$$R = A \frac{1}{ne},$$

где $A \approx 1$ для металлов; у полупроводников в зависимости от структуры решетки A имеет различные значения (1,11—1,93). У некоторых металлов и у полупроводников с дырочной проводимостью наблюдается противоположный знак разности потенциалов: верхняя поверхность на рис. III.61 заряжается положительно, нижняя — отрицательно (аномальный эффект Холла). Поэтому измерение коэффициента Холла позволяет определить характер проводимости (электродный или дырочный), а также концентрацию и подвижность носителей тока.

Эффект Холла широко используется для измерительных целей. Миниатюрные датчики холл-эффекта сделаны из небольшой полупроводниковой пластинки или из тонкого (0,1—0,01 мм) слоя полупроводника на слюдяной пластинке: два электрода служат для подводки тока, два — для измерения поперечной разности потенциалов. При помощи этих датчиков можно измерить любые величины, влияющие на холловскую разность потенциалов (сила тока через датчик, индукция и напряженность внешнего магнитного поля, ориентировка датчика относительно этого поля и т. д.). Кроме того, холл-эффект используется во многих электро- и радиотехнических установках (преобразование токов, модуляция электрических колебаний, запись звуков, усиление постоянного и переменного токов и т. д.).

§ 21. РАБОТА ПЕРЕМЕЩЕНИЯ ПРОВОДНИКА С ТОКОМ В МАГНИТНОМ ПОЛЕ

Воздействие магнитного поля на проводники с током используется, например, в электромоторах для превращения электрической энергии в механическую. Рассчитаем работу, которая совершается при перемещении проводника с током в магнитном поле. Для простоты допустим, что вдоль неподвижных проводников, расположенных в плоскости, перпендикулярной вектору индукции **B** магнитного поля, может перемещаться контактирующий с ними прямолинейный проводник длиной l (рис. III.62). Ток I через этот проводник вызывается источником, присоединенным к неподвижным проводникам. Обозначим через v_1 среднюю скорость упорядоченного движения электронов в проводниках. Допустим сначала, что рассматриваемый проводник неподвижен. На него будет действовать сила Ампера $F = BIl$ (т. е. равнодействующая всех сил Лоренца, действующих на электроны проводимости, находящиеся в данный момент в пределах объема проводника). Так как силы Лоренца и Ампера перпендикулярны направлению скорости v_1, то они работы не совершают и поэтому не могут изменить величину скорости v_1.

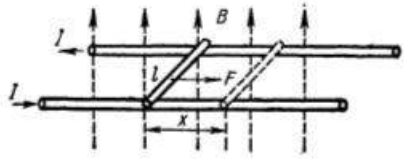


Рис. III.62

Допустим теперь, что проводник движется со скоростью v_2 и за некоторое время t проходит расстояние x (рис. III. 62). В этом случае скорость упорядоченного движения электронов внутри проводника будет равна уже не v_1, а векторной сумме скоростей $\mathbf{v}_1$ и $\mathbf{v}_2$. Вследствие этого у движущегося проводника направления сил Лоренца и Ампера будут иными, чем у неподвижного. Значение этого важного обстоятельства будет обсуждаться в дальнейшем. Пока же мы допустим, что сила $F = BIl$, действующая на проводник в направлении его движения, сохраняется постоянной. Работа этой силы будет равна

$$W = Fx = BIlx = IB\,\Delta S = I\,\Delta\Phi, \qquad (3.29)$$

где $\Delta\Phi = B\Delta S$ — магнитный поток через площадь $\Delta S = lx$, описанную проводником. Таким образом, *работа, совершаемая амперовой силой при перемещении проводника с током в магнитном поле, равна силе тока, умноженной на поток магнитной индукции через площадь, описанную этим проводником.*

Этот вывод сохраняется и в общем случае, когда проводник имеет сложную форму. Допустим, что его отдельные элементарные части $\mathrm{d}l$ ориентированы под различными углами (α) к направлению вектора индукции **B**; возможно также, что для них углы (β) между направлением перемещения $\mathrm{d}x$ и направлением силы Ампера $\mathrm{d}F = BI\mathrm{d}l\sin\alpha$ будут отличны от прямого. В этом случае работа, совершаемая при перемещении элементарной части проводника, будет равна

$$\mathrm{d}W = \mathrm{d}F\,\mathrm{d}x\cos\beta = BI\,\mathrm{d}l\sin\alpha\,\mathrm{d}x\cos\beta = IB\sin\alpha\,\mathrm{d}S = I\,\mathrm{d}\Phi, \qquad (3.30)$$

где $dS = dl\,dx\cos\beta$ — площадь, описанная рассматриваемым элементом тока, а $B\,dS\sin\alpha = d\Phi$ — магнитный поток через эту площадь. Суммируя эти элементарные работы, получим для всего проводника независимо от его формы и характера движения формулу

$$W = I\,\Delta\Phi. \tag{3.31}$$

Формула (3.31) может применяться также и для вычисления работы, которая совершается в магнитном поле при перемещениях замкнутых контуров. Можно показать, что для замкнутых контуров $\Delta\Phi$ означает изменение магнитного потока через площадь, охваченную этим контуром. При этом поле может быть неоднородным, поворот — неравномерным. Если плоский контур поворачивается в магнитном поле так, что угол α между направлениями вектора **B** и нормали к плоскости контура изменяется от α_1 до α_2, то

$$W = I\Phi_0(\cos\alpha_2 - \cos\alpha_1),$$

где Φ_0 — наибольшее значение магнитного потока, когда плоскость контура перпендикулярна полю. Если контур представляет собой катушку (рамку) с n одинаковыми витками, то $W = nI\Delta\Phi$.

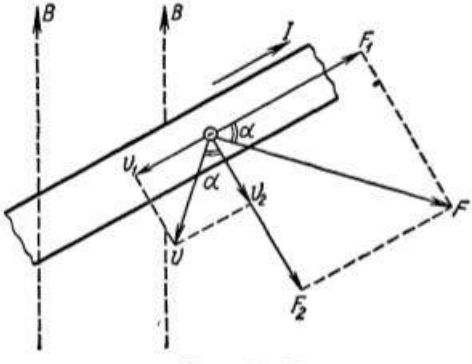


Рис. III.63

Покажем теперь, что работа перемещения W совершается не магнитным полем, а электрическим током, проходящим через проводник. На рис. III.63 показаны: $\mathbf{v}_1$ — скорость упорядоченного движения электронов в проводнике, зависящая от величины тока I, $\mathbf{v}_2$ — скорость движения самого проводника в магнитном поле, $\mathbf{v}$ — векторная сумма этих скоростей. Силы Лоренца (а следовательно, и их равнодействующая — сила Ампера) будут перпендикулярны вектору результирующей скорости $\mathbf{v}$ и поэтому направлены под углом α к проводнику. Разложим каждую из сил Лоренца на составляющие $\mathbf{F}_1$ и $\mathbf{F}_2$ (рис. III.63). При движении электронов внутри проводника сила Лоренца, перпендикулярная их скорости $\mathbf{v}$, работы не совершает, но составляющие $\mathbf{F}_1$ и $\mathbf{F}_2$ совершают равные и противоположные по знаку работы; за время dt

$$dW_1 = -F_1 v_1\,dt = -Fv\sin\alpha\cos\alpha\,dt,$$

$$dW_2 = F_2 v_2\,dt = Fv\sin\alpha\cos\alpha\,dt.$$

Очевидно, что такой же результат мы получим и для амперовой силы, действующей на прямолинейный элемент тока (Il). Заметим, что составляющая амперовой силы $\mathbf{F}_2$, хотя и приложена к электронам проводимости, передается ионному каркасу проводника и поэтому может рассматриваться как механическая сила, приложенная ко

всему проводнику. По величине сила $\mathbf{F}_2$ равна BIl, т. е. равна той амперовой силе, с которой магнитное поле действовало бы на неподвижный проводник. Таким образом, при движении проводника с током в магнитном поле работа силы $F_2 = BIl$, направленной вдоль скорости проводника v_2, должна компенсироваться равной и противоположной по знаку работой силы F_1, направленной против скорости v_1. Это означает, что работа сил F_2 совершается не за счет магнитного поля, а за счет энергии упорядоченного движения электронов проводимости.

Рассмотрим частный случай, когда проводник движется в магнитном поле с постоянной скоростью v_2 и ток I в нем поддерживается постоянным. Это возможно, во-первых, если к проводнику будет приложена внешняя сила $F_{\text{внеш}}$, равная F_2 и направленная в противоположную сторону, и, во-вторых, если источник тока будет поддерживать постоянной скорость упорядоченного движения электронов v_1. Благодаря первому условию сумма сил $\mathbf{F}_{\text{внеш}} + \mathbf{F}_2$ равна нулю и поэтому скорость v_2 будет постоянной. Второе условие означает, что источник тока должен затрачивать энергию, в точности равную работе сил $\mathbf{F}_1$, тормозящих движение электронов проводимости. Согласно закону сохранения энергии, для любого промежутка времени $\mathrm{d}t$ должно соблюдаться соотношение (верное как для каждого отдельного электрона проводимости, так и, следовательно, для всех электронов проводимости, содержащихся в каком-нибудь объеме проводника с током)

$$\mathbf{F}_{\text{внеш}}\mathbf{v}_2\,\mathrm{d}t + \mathbf{F}_2\mathbf{v}_2\,\mathrm{d}t + \mathbf{F}_1 v_1\,\mathrm{d}t + \mathrm{d}W_{\text{ист}} = 0,$$

где $\mathrm{d}W_{\text{ист}}$ — работа, совершаемая источником тока. Так как суммарная работа сил $\mathbf{F}_1$ и $\mathbf{F}_2$ (т. е. работа силы $\mathbf{F}$) равна нулю, то из этого соотношения следует, что

$$\mathbf{F}_{\text{внеш}}\mathbf{v}_2\,\mathrm{d}t + \mathrm{d}W_{\text{ист}} = 0,$$

т. е. в магнитном поле работа внешней силы, приложенной к проводнику с током, совершается за счет энергии источника тока. Заметим, что в движущемся проводнике для поддержания одной и той же силы тока напряженность электрического поля (E) должна быть больше, чем в неподвижном проводнике (E_0), на величину F_1/e (e — заряд электрона, F_1 — составляющая силы Лоренца внутри проводника):

$$E = E_0 + F_1/e = E_0 + vB\sin\alpha. \qquad (3.32)$$

§ 22. ЭЛЕКТРОМАГНИТНАЯ ИНДУКЦИЯ; ЗАКОНЫ ФАРАДЕЯ И ЛЕНЦА. ВРАЩАЮЩИЙСЯ ВИТОК В МАГНИТНОМ ПОЛЕ

ДВИЖУЩИЙСЯ ПРОВОДНИК В МАГНИТНОМ ПОЛЕ

Допустим, что в магнитном поле $\mathbf{B}$ находится металлический проводник; на рис. III.64 показан небольшой его участок длиной Δl. Свободные электроны, имеющиеся в этом проводнике, участвуют в тепловом движении, и в данный момент времени скорости их v_i раз-

личны и беспорядочно ориентированы в пространстве. Магнитное поле действует на каждый электрон с силой $\mathbf{f}_i$, направление которой перпендикулярно скорости $\mathbf{v}_i$ и полю **B**. Так как скорости теплового движения электронов ориентированы беспорядочно, то и силы $\mathbf{f}_i$, действующие на различные электроны, а также силы, действующие на один электрон в различные моменты времени, имеют всевозможные направления. Таким образом, магнитное поле не может «навести порядок» в тепловом движении зарядов, т. е. создать их упорядоченное одностороннее перемещение. Это означает, что при рассмотрении вопроса о воздействии магнитного поля на свободные электроны можно отвлечься от их теплового движения и полагать, что все они покоятся в объеме металла.

Если проводник движется в магнитном поле с некоторой скоростью **v**, то эта скорость накладывается на скорости беспорядочного теплового движения электронов. Условившись исключить из

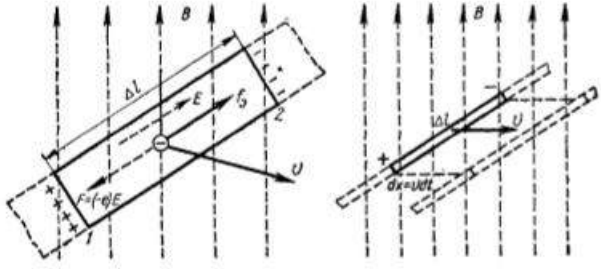


Рис. III.64

рассуждений тепловое движение электронов, предположим, что все они вместе с проводником движутся в магнитном поле с одной и той же скоростью **v**. Тогда на каждый электрон со стороны магнитного поля действует одна и та же сила $f_э$:

$$f_э = (-e)\,vB\sin\alpha; \quad \mathbf{f}_э = (-e)\,[\mathbf{vB}], \tag{3.33}$$

где $(-e)$ — заряд электрона; α — угол между векторами **v** и **B**.

Смещение электронов в направлении действия силы $f_э$ вызывает их скопление в сечении *2* и уменьшает их концентрацию в сечении *1*. Это перераспределение зарядов в пределах объема проводника Δl сопровождается появлением внутри него электрического поля, напряженность которого **E** направлена от сечения *1* к сечению *2*; это поле действует на отрицательные электроны с силой $\mathbf{F} = (-e)\mathbf{E}$, направленной против **E**. Рассмотрим два случая:

1) если в магнитном поле движется незамкнутый проводник длиной Δl, то электрическое поле **E** со временем приостановит движение зарядов в направлении действия силы $\mathbf{f}_э$. Равновесное распределение зарядов в проводнике наступит (при постоянстве скорости движения зарядов **v** и индукции магнитного поля **B**) при условии

$$\mathbf{f}_э = -\mathbf{F} = -(-e)\,\mathbf{E} = e\mathbf{E}. \tag{3.34}$$

Из постоянства **v** и **B** следует постоянство $\mathbf{f}_э$ и **E**. Электрическое поле внутри участка Δl однородное, поэтому разность потенциалов между сечениями *1* и *2*

$$\varphi_1 - \varphi_2 = E\,\Delta l; \qquad (3.35)$$

2) если проводник замыкается за пределами магнитного поля (полагаем, что поле **B** действует на электроны, находящиеся только в пределах участка Δl), поле **E** вызовет движение электронов через внешние участки цепи; по проводнику потечет электрический ток.

Таким образом, участки проводника, перемещающиеся в магнитном поле, можно рассматривать как «источники тока», обладающие особой электродвижущей силой, которую называют *электродвижущей силой индукции*. Э. д. с. индукции некоторого участка проводника Δl можно найти по разности потенциалов на его концах, если проводник не

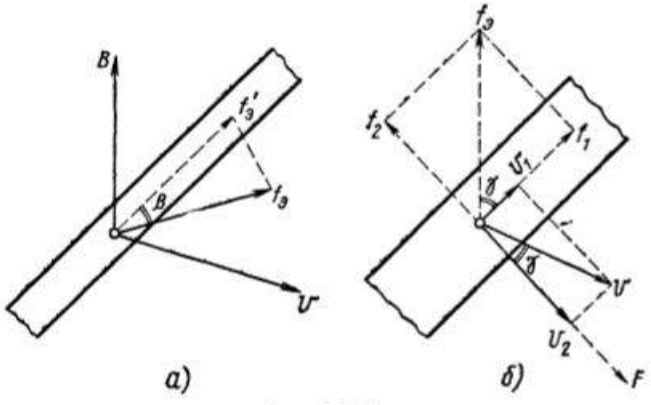


Рис. III.65

замкнут. Из соотношения (3.33) и формулы (3.34) найдем напряженность электрического поля в проводнике E в зависимости от v и B:

$$(-e)\,vB\sin\alpha = eE; \quad E = -vB\sin\alpha. \qquad (3.36)$$

Тогда разность потенциалов на концах участка проводника

$$\mathscr{E}_{инд} = E\,\Delta l = -Bv\,\Delta l\sin\alpha. \qquad (3.37)$$

В этой формуле учтен только угол α между векторами скорости проводника **v** и магнитного поля **B**. Однако в общем случае сам проводник может быть не перпендикулярен **v** и **B**, как это предполагалось выше и показано на рис. III.64, а ориентирован произвольным образом. Тогда следует учесть, что сила $\mathbf{f}_э$, которая всегда перпендикулярна **v** и **B**, не будет уже направлена вдоль проводника, а составляет с ним некоторый угол β (рис. III.65, *а*). Очевидно, в этом случае движение электронов в проводнике вызывается не всей силой $f_э$, а только ее проекцией $f_э' = f_э\cos\beta$. Для равновесного состояния вместо формулы (3.34) следует писать $f_э' = eE$ и тогда

$$E = -Bv\sin\alpha\cos\beta;$$

$$\mathscr{E}_{инд} = -Bv\,\Delta l\sin\alpha\cos\beta. \qquad (3.38)$$

Однако для упрощения рассуждений в дальнейшем будем всегда полагать, что $\beta = 0$ и $\cos\beta = 1$, т. е. что проводник перпендикулярен плоскости, в которой лежат векторы $\mathbf{v}$ и $\mathbf{B}$.

Формула (3.37) для э. д. с. индукции записывается и в другом, более общем виде. Сделаем еще одно упрощение: допустим, что векторы $\mathbf{v}$ и $\mathbf{B}$ перпендикулярны между собой, т. е. $\alpha = 90°$, $\sin\alpha = 1$, $\mathscr{E}_{\text{инд}} = -Bv\Delta l$. Так как $v = \mathrm{d}x/\mathrm{d}t$, $\Delta l\mathrm{d}x = \mathrm{d}S$ — площадь, описываемая проводником при его движении за время $\mathrm{d}t$ (см. рис. III.64), $B\mathrm{d}S = \mathrm{d}\Phi$ — магнитный поток, проходящий через эту площадь, то

$$\mathscr{E}_{\text{инд}} = -\frac{\mathrm{d}\Phi}{\mathrm{d}t}, \tag{3.39}$$

т. е. *электродвижущая сила, возникающая при движении проводника в магнитном поле, численно равна магнитному потоку через площадь, которую этот проводник описывает в единицу времени.* Этот вывод можно получить и в общем случае, когда $\alpha \neq 90°$ и проводник ориентирован в магнитном поле любым образом. В частности, если прямолинейный проводник параллелен вектору $\mathbf{B}$, то магнитный поток через описываемую им площадь равен нулю; в этом случае на концах проводника разности потенциалов не возникает. Заметим, что в приведенных выше формулах для E и $\mathscr{E}$ знак минус вовсе не связан с тем, что рассматриваются силы, действующие на отрицательно заряженные электроны. Повторив рассуждения для свободных положительных зарядов, можно получить те же самые формулы (3.37) и (3.39).

Если проводник имеет сложную форму и движется в неоднородном магнитном поле, то его можно разделить на такие элементарные участки Δl_i и рассматривать такие малые перемещения Δx_i, чтобы для каждого участка в пределах описываемых ими площадок ΔS_i можно было магнитное поле считать однородным. Затем можно произвести суммирование элементарных э. д. с.:

$$\mathscr{E}_{\text{инд}} = -\sum B_i v_i \Delta l_i \sin\alpha_i \cos\beta_i = -\sum \frac{\Delta\Phi_i}{\Delta t};\ \mathscr{E}_{\text{инд}} = -\frac{\Delta\Phi}{\Delta t}.$$

В пределе, когда участки Δl_i бесконечно малые и $\Delta t \to 0$, мы получим точную формулу (3.39).

Формулы (3.37) и (3.39) можно применять и для *замкнутых контуров.* Однако как только по проводнику потечет ток, т. е. начнется упорядоченное движение электронов вдоль проводника, скорость перемещения электронов относительно магнитного поля уже не будет равна скорости проводника. Обозначим скорость упорядоченного движения электронов вдоль проводника через v_1, а скорость их перемещения вместе с проводником — через v_2 (рис. III.65, *б*). В этом случае сила $f_э$, действующая на электрон со стороны магнитного поля, будет составлять с проводником угол γ.

Разложим силу $\mathbf{f}_э$ на составляющие: f_1, ориентированную вдоль проводника по направлению движения электронов, и f_2, направленную против движения проводника. Сила f_1 совершает положительную работу, сообщая электронам энергию, которую они расходуют при

своем движении по контуру. Сила f_2 препятствует движению проводника и совершает отрицательную работу.

Сумма работ, совершаемых этими силами за какой-нибудь отрезок времени, должна быть равна работе силы $\mathbf{f}_э$ за то же время, т. е. должна равняться нулю, так как сила $f_э$ перпендикулярна скорости движения электронов $\mathbf{v}$ и поэтому работы не совершает. Таким образом, положительная работа силы f_1 по величине равна отрицательной работе силы f_2.

Если к проводнику не прилагать внешней силы, то силы f_2 со временем остановят его (как только будет израсходована кинетическая энергия этого проводника). Для того чтобы поддержать движение проводника (со скоростью v_2), к нему необходимо приложить внешнюю силу $\mathbf{F}$, равную сумме всех сил $\mathbf{f}_2$ и направленную в сторону движения. Если в единице объема проводника имеется n электронов, участвующих в токе, то на участке длиной Δl и сечением S их будет $S\Delta ln$. Сила $f_2 = f_э \sin\gamma = evB \sin\gamma = ev_1B$. Тогда сумма всех сил f_2 равна

$$\sum f_2 = S\,\Delta l n e v_1 B = n e v_1 S\,\Delta l B = I B\,\Delta l$$

(так как nev_1S — сила тока в проводнике I). Заметим, что Σf_2 равна амперовой силе, действующей на проводник с током со стороны магнитного поля.

Таким образом, для того чтобы вызвать в проводнике электрический ток, перемещая его в магнитном поле, необходимо преодолеть амперову силу, действующую на появившийся ток со стороны того же магнитного поля. Положительная работа приложенной внешней силы $\mathbf{F}$ равна положительной работе сторонних сил $\mathbf{f}_1$, вызывающих движение электронов в проводнике. Следовательно, *источником энергии индукционных токов* (т. е. токов, возбуждаемых э. д. с. индукции) *является работа внешних сил,* приложенных к движущемуся проводнику. Если же проводник не замкнут, то упорядоченное движение электронов в проводнике существует только в течение малого отрезка времени, пока устанавливается равновесное распределение зарядов; после наступления равновесия $v_1 = 0$, $f_2 = ev_1B = 0$ и необходимости во внешней силе F нет.

РАМКА В МАГНИТНОМ ПОЛЕ

Большое практическое значение имеет возбуждение э. д. с. индукции в проводниках, имеющих форму одного или многих витков, обычно намотанных на жесткую рамку. Рассмотрим плоский виток прямоугольной формы и допустим, что он движется вдоль своей плоскости (как это показано на рис. III.66, *а*). Силы $f_э$ направлены вдоль длины проводников только для участков *1—2* и *3—4*; для участков *2—3* и *1—4* эти силы перпендикулярны оси проводников и поэтому упорядоченного движения зарядов в витке вызвать не могут. На участке *1—2* э. д. с. индукции равна $\mathscr{E}_1 = -B_1va$ ($\sin\alpha = 1$), где B_1 — индукция магнитного поля в том месте, где находится проводник *1—2*. На участке *3—4* имеем $\mathscr{E}_2 = -B_2va$. Так как при движении зарядов в витке силы $f_э$ действуют в противоположных направ-

лениях, то суммарная э. д. с. в витке равна разности:

$$\mathscr{E}_{\text{инд}} = \mathscr{E}_1 - \mathscr{E}_2 = (B_2 - B_1)\, va. \tag{3.40}$$

Следовательно, чтобы при таком движении витка вызвать в нем индукционный ток, магнитное поле должно быть неоднородным (в однородном поле $B_1 = B_2$, $\mathscr{E} = 0$). Заметим, что $v = \mathrm{d}x/\mathrm{d}t$; $va = a(\mathrm{d}x/\mathrm{d}t) = \mathrm{d}S/\mathrm{d}t$, где $\mathrm{d}S$ — площадь, описанная проводниками *1—2* и *3—4* за время $\mathrm{d}t$. Далее,

$$B_1 va = B_1 \frac{\mathrm{d}S}{\mathrm{d}t} = \frac{\mathrm{d}\Phi_1}{\mathrm{d}t}; \quad B_2 va = \frac{\mathrm{d}\Phi_2}{\mathrm{d}t},$$

где $\mathrm{d}\Phi_1$ — магнитный поток, входящий в пределы площади витка, а $\mathrm{d}\Phi_2$ — магнитный поток, выходящий из витка за время $\mathrm{d}t$. Тогда

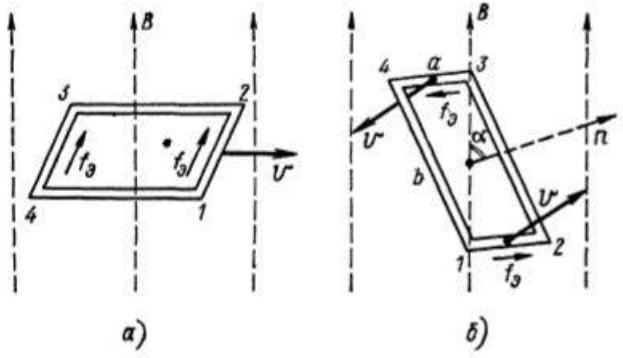


Рис. III.66

приращение магнитного потока через площадь витка равно $\mathrm{d}\Phi = \mathrm{d}\Phi_1 - \mathrm{d}\Phi_2$ и формула (3.40) примет уже знакомый нам вид

$$\mathscr{E}_{\text{инд}} = -\frac{\mathrm{d}\Phi}{\mathrm{d}t}, \tag{3.41}$$

где Φ означает магнитный поток через площадь витка (т. е. «связанный» с данным витком).

Таким образом, *э. д. с. индукции в витке равна скорости изменения магнитного потока через площадь этого витка.* Если виток имеет сложную форму, то его можно разбить на элементарные участки и тогда, повторив приведенные выше рассуждения, мы снова придем к тому же утверждению. Заметим, что индукционные токи в замкнутых проводниках можно вызвать и путем деформации контура, когда изменяется площадь, охватываемая контуром. Большое применение имеет возбуждение э. д. с. индукции при *вращении* витка; в этом случае силы $f_{\text{э}}$ на участках *1—2* и *3—4* (рис. III.66, *б*) действуют в одном и том же направлении, поэтому индукционный ток может быть получен и в однородном магнитном поле.

В катушке, содержащей n одинаковых витков (т. е. n одинаковых последовательно соединенных контуров), э. д. с. индукции, возбуждае-

мые в витках, суммируются и тогда

$$\mathscr{E}_{\text{инд}} = -\frac{d\Phi}{dt} n; \tag{3.42}$$

можно сказать, что с этой катушкой «связан» магнитный поток, в n раз больший, чем с одним витком.

Для одного витка, вращающегося в однородном магнитном поле, $\Phi = BS\cos\alpha$, где $S = ab$ — площадь витка; α — угол между вектором площадки S (нормали **n** к площади витка) и вектором **B**. Обозначим максимальное значение магнитного потока через площадь витка $BS = \Phi_0$; допустим, что виток вращается с постоянной угловой скоростью ω, и будем отсчитывать время с момента, когда направления S и B совпадают. Тогда угол $\alpha = \omega t$ и

$$\Phi = BS\cos\omega t = \Phi_0\cos\omega t, \tag{3.43}$$

$$\mathscr{E}_{\text{инд}} = -\frac{d\Phi}{dt} = \Phi_0\omega\sin\omega t = BS\omega\sin\omega t = \mathscr{E}_0\sin\omega t. \tag{3.44}$$

($\mathscr{E}_0 = BS\omega$); при этих условиях э. д. с. индукции и индукционный ток синусоидальные. Если вращаются n одинаковых последовательно соединенных витков, то $\mathscr{E}_0 = BS\omega n$.

В некоторых случаях э. д. с. индукции и индукционные токи возбуждаются в течение очень малых промежутков времени и изменяются со временем по очень сложным законам. Для таких кратковременных индукционных токов можно определить количество электричества, прошедшее через сечение контура за время существования токов. Если кроме э. д. с. индукции других источников электродвижущей силы нет, то сила тока в контуре, по закону Ома,

$$I = \frac{\mathscr{E}_{\text{инд}}}{R} = -\frac{1}{R}\frac{d\Phi}{dt}, \tag{3.45}$$

где R — полное сопротивление контура (о применимости закона Ома к переменным токам будет сказано в следующем параграфе). Тогда за время dt через сечение контура пройдет количество электричества

$$dq = I\,dt = -\frac{1}{R}d\Phi.$$

За время прохождения индукционного тока, когда магнитный поток через площадь контура изменяется на $\Delta\Phi = \Phi_1 - \Phi_2$, через сечение контура пройдет заряд

$$q = \int I\,dt = \int_{\Phi_1}^{\Phi_2}\left(-\frac{1}{R}d\Phi\right) = \frac{\Phi_1 - \Phi_2}{R} = \frac{\Delta\Phi}{R} \tag{3.46}$$

(если сопротивление контура R сохраняется постоянным). Этой формулой пользуются для измерения магнитного потока Φ, вектора магнитной индукции B или магнитной проницаемости вещества μ. Количество электричества q измеряется при помощи так называемого «баллистического гальванометра» (при условии, что время прохождения тока через рамку этого гальванометра достаточно мало по сравне-

нию с периодом ее колебаний). Если к такому гальванометру присоединить контур, находящийся в измеряемом магнитном поле, то при быстром повороте этого контура охватываемый им магнитный поток будет изменяться; для удобства можно осуществить поворот от положения, при котором $\Phi_1 = 0$, до положения $\Phi_2 = BS$; тогда $\Delta\Phi = BS$. Измеряя q и R, можно рассчитать $\Delta\Phi$ и B (а также вычислить μ, если опыт произвести один раз в вакууме, а затем в данной среде).

Формула (3.46) используется для выбора единицы измерения магнитного потока. По принятым в настоящее время стандартам физических величин, *вебер* равен магнитному потоку, при убывании которого до нуля в сцепленной с ним электрической цепи (т. е. в контуре, охватывающем этот поток), имеющей сопротивление 1 Ом, через поперечное сечение проводника проходит количество электричества, равное 1 Кл.

При расчете силы тока в контуре по формуле (3.45) предполагалось, что в контуре действует только одна электродвижущая сила, а именно $\mathscr{E}_{\text{инд}} = -(d\Phi/dt)$. Однако возможны сложные цепи, в которых токи возбуждаются не только электромагнитной индукцией, но одновременно и другими источниками тока, включенными в эту цепь (например, аккумуляторными батареями). Обозначим э. д. с. других источников через $\mathscr{E}$, полное сопротивление контура (включая сопротивление самих источников тока) — через R. Тогда

$$I = \frac{\mathscr{E} + \mathscr{E}_{\text{инд}}}{R} = \frac{\mathscr{E} - \frac{d\Phi}{dt}}{R}. \qquad (3.47)$$

ЗАКОНЫ ФАРАДЕЯ И ЛЕНЦА

Наблюдения показывают, что э. д. с. индукции возбуждается не только при движении проводников в магнитном поле, но и в неподвижных проводниках, если только они находятся в *переменном* магнитном поле. В частности, в неподвижных витках вторичной обмотки трансформатора возбуждается э. д. с. индукции благодаря тому, что эти витки находятся в переменном магнитном поле первичной обмотки. В этих явлениях наблюдается воздействие переменного магнитного поля на неподвижные заряды, которое будет рассматриваться в § 24. Измерения показали, что при электромагнитной индукции, вызванной переменным магнитным полем в замкнутых контурах, соблюдаются те же формулы (3.41) и (3.42). Таким образом, для электромагнитной индукции можно сформулировать весьма общий закон:

если магнитный поток Φ *через площадь контура изменяется со временем* (вследствие движения контура или изменения магнитного поля, в котором находится контур), *то в контуре возбуждается электродвижущая сила, равная*

$$\mathscr{E}_{\text{инд}} = -\frac{d\Phi}{dt}. \qquad (3.48)$$

Это утверждение есть закон электромагнитной индукции Фарадея. В случае, если контур имеет сложную форму (рис. III. 67), необходимо спроецировать его на плоскость, перпендикулярную направлению поля; для каждого полученного в этой проекции замкнутого контура (*1*, *2* и т. д.) следует рассчитать э. д. с. индукции по формуле (3.48) и затем сложить их. У катушки, содержащей n витков с одинаковыми $d\Phi/dt$, э. д. с. индукции будет в n раз больше, чем у одного витка.

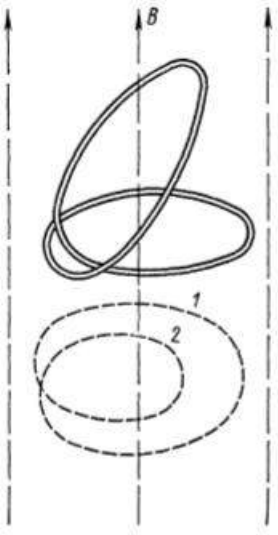


Рис. III.67

В явлениях электромагнитной индукции действует также и другой важный закон (закон Ленца):

индукционные токи всегда имеют такое направление, при котором оказывается противодействие причинам, вызвавшим эти токи.

Например, при движении проводника в магнитном поле в нем появляется индукционный ток такого направления, что амперова сила, действующая на этот ток со стороны магнитного поля, противодействует движению проводника. При вращении контура в магнитном поле в нем индуцируется ток такого направления, что механический момент, действующий на этот контур с током со стороны магнитного поля, противодействует вращению контура. Если неподвижный контур находится в переменном магнитном поле, то в нем индуцируется ток такого направления, что собственное магнитное поле этого тока противодействует изменению внешнего магнитного поля (если магнитный поток Φ через площадь контура увеличивается, то собственное магнитное поле индукционного тока уменьшает его, т. е. $\mathbf{B}_{инд}$ направлено против $\mathbf{B}_{внеш}$ внешнего поля, и, наоборот, если Φ уменьшается, то собственное магнитное поле индукционного тока одного направления с $B_{внеш}$). Во всех перечисленных примерах индукционные токи противодействуют тем причинам (внешним воздействиям на контур), которые их вызвали.

§ 23. ЯВЛЕНИЯ САМОИНДУКЦИИ И ВЗАИМОИНДУКЦИИ. ЭНЕРГИЯ МАГНИТНОГО ПОЛЯ. ВИХРЕВЫЕ ТОКИ

Электромагнитную индукцию в неподвижных проводниках можно вызвать при помощи переменного магнитного поля, создаваемого переменным электрическим током. Если э. д. с. индукции в данном проводнике вызывается переменным током, протекающим в других, расположенных по соседству проводниках, то это явление называют *взаимоиндукцией*; если же э. д. с. индукции возбуждается в том же

самом проводнике, по которому течет переменный ток, то это явление называется *самоиндукцией*.

Допустим, что через проводник, который на рис. III. 68, *а* изображен в виде катушки с n витками, проходит ток силой I. Этот ток создает вокруг проводника магнитное поле, так что каждый элемент проводника Δl_i находится в магнитном поле $\mathbf{H}_i$, созданном в этом месте остальными элементами проводника; значение вектора $\mathbf{H}_i$ зависит от силы тока и от расположения остальных элементов проводника, т. е. от его формы. Если ток I меняется со временем, то изменяется и $\mathbf{H}_i$, вследствие чего на концах участков Δl_i возбуждаются

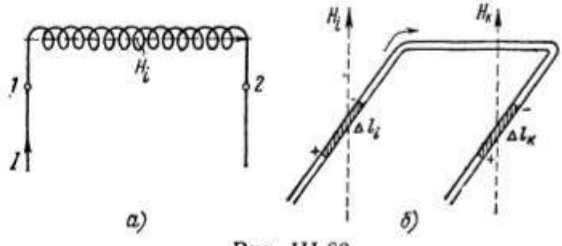


Рис. III.68

э. д. с. индукции $\Delta \mathscr{E}_i$. Так как все элементы Δl_i соединены последовательно, то общая э. д. с. индукции, возбуждаемая на концах *1* и *2* проводника, равна сумме: $\mathscr{E}_{инд} = \Sigma \Delta \mathscr{E}_i$. Однако в этой сумме могут содержаться и отрицательные члены, так как знак $\Delta \mathscr{E}_i$ зависит от ориентации участков Δl_i относительно вектора $\mathbf{H}_i$. Если, например, для участка Δl_i (рис. III. 68, *б*) знак $\Delta \mathscr{E}_i$, согласно указанному стрелкой направлению обхода, положителен, то для последовательно соединенного с ним участка $\Delta l_к$ того же проводника знак $\Delta \mathscr{E}_к$ отрицателен. У длинных катушек (соленоидов), внутри которых угол между элементами витков Δl_i и направлением магнитного поля везде одинаков, все $\mathscr{E}_{инд}$ имеют одинаковый знак (см. вывод формулы (3.41)).

ИНДУКТИВНОСТЬ СОЛЕНОИДА

Расчет э. д. с. индукции для каждого участка проводника в зависимости от формы всего проводника и быстроты изменения силы тока представляет в общем случае большие трудности. Ограничимся рассмотрением только длинного соленоида. Магнитное поле такого соленоида сосредоточено внутри его объема, причем это поле можно считать однородным (см. § 18). Допустим, что весь объем магнитного поля соленоида заполнен однородным веществом с относительной магнитной проницаемостью μ. Тогда магнитный поток Φ_1, «пронизывающий» каждый виток с сечением S, равен

$$\Phi_1 = BS = \mu_0 \mu HS = \mu_0 \mu \frac{In}{l} S, \qquad (3.49)$$

где l — длина соленоида; n — число витков. Если сила тока в соленоиде изменяется со временем, то изменяется и магнитный поток Φ_1

охватываемый каждым витком, и тогда, по закону электромагнитной индукции Фарадея, в каждом витке появится э. д. с., равная

$$\Delta\mathscr{E}_1 = -\frac{d\Phi_i}{dt} = -\frac{d}{dt}\left(\mu_0\mu\frac{In}{l}S\right).$$

В n последовательно соединенных витках, составляющих соленоид, возбуждается суммарная э. д. с. самоиндукции; если витки одинаковые, то

$$\mathscr{E}_{\text{самоинд}} = \sum\Delta\mathscr{E}_1 = n\,\Delta\mathscr{E}_1 = -\frac{d}{dt}\left(\mu_0\mu\frac{In^2}{l}S\right). \qquad (3.50)$$

Так как между витками соленоида по закону Ампера действуют механические силы, зависящие от тока I, то при изменении этого тока в свободном соленоиде его размеры (в частности, длина) могут несколько измениться. Однако это изменение незначительно и обычно игнорируется. Допустим, что силы, деформирующие соленоид, очень малы, поэтому l и S остаются постоянными. Магнитная проницаемость вещества, находящегося внутри соленоида, зависит от напряженности намагничивающего поля, создаваемого током внутри соленоида. У ферромагнитных тел (железо, магнитные сплавы и др.) эта зависимость очень сильная, поэтому если внутри соленоида находится ферромагнитный сердечник, то магнитный поток Φ_1 изменяется не только вследствие изменения силы тока I, но и от изменения μ. Обозначим для простоты записи

$$L = \mu_0\mu\frac{n^2}{l}S. \qquad (3.51)$$

Тогда

$$\mathscr{E}_{\text{самоинд}} = -\frac{d}{dt}(LI) = -L\frac{dI}{dt} - I\frac{dL}{dt}. \qquad (3.52)$$

Если L остается постоянным (например, у соленоида без сердечника), то *э. д. с. самоиндукции прямо пропорциональна скорости изменения силы тока в проводнике*:

$$\mathscr{E}_{\text{самоинд}} = -L\frac{dI}{dt}. \qquad (3.53)$$

Из формулы (3.52) следует, что э. д. с. самоиндукции может возникнуть и при постоянной силе тока, если только изменяется L.

Вернемся к формуле (3.50); имеем

$$\mathscr{E}_{\text{самоинд}} = -n\frac{d\Phi_i}{dt} = -\frac{d}{dt}(n\Phi_1) = -\frac{d\Phi}{dt}. \qquad (3.54)$$

Это означает, что для расчета э. д. с. самоиндукции по формуле Фарадея соленоид можно представить либо как совокупность последовательно соединенных витков, охватывающих один и тот же магнитный поток Φ_1, либо же как некоторый замкнутый проводник сложной формы, с которым «связан» магнитный поток

$$\Phi = n\Phi_1 = \mu_0\mu\frac{In^2}{l}S = LI. \qquad (3.55)$$

Однако нельзя забывать, что внутри соленоида в действительности существует не поток Φ, а только поток Φ_1, определяемый формулой (3.49). Если, например, отрезать от соленоида часть его витков (но чтобы соленоид оставался по-прежнему «длинным») и вновь пропустить через него тот же ток I, то Φ_1 не изменится, тогда как Φ уменьшится пропорционально числу удаленных витков. Умножение на число витков в формуле (3.50) производится потому, что изменение магнитного потока Φ_1 возбуждает э. д. с. самоиндукции одновременно в n последовательно соединенных витках.

Выше рассматривалось явление самоиндукции в длинном соленоиде с однородным сердечником внутри. В общем случае для проводников любой формы также можно воспользоваться формулами (3.52) и (3.53), если найти магнитный поток.

$$\Phi = LI, \tag{3.56}$$

«связанный» с этим проводником при силе тока I. Пропорциональность между Φ и I очевидна, так как, согласно изложенному в § 18, напряженности магнитного поля в различных местах вокруг проводника любой формы прямо пропорциональны силе тока через проводник. Величина L учитывает форму и размеры проводника, а также магнитные свойства тел, находящихся в магнитном поле этого проводника; ее называют *индуктивностью* или *коэффициентом самоиндукции* данного проводника.

РАСЧЕТ ИНДУКЦИОННЫХ ТОКОВ

Таким образом, в проводнике, по которому течет переменный ток, действуют две электродвижущие силы:

1) э. д. с. источника тока, создающая ток I, и

2) э. д. с. самоиндукции, пропорциональная скорости изменения этого тока.

Если ток I в контуре возрастает, то производная $\mathrm{d}I/\mathrm{d}t$ имеет положительный знак, а э. д. с. самоиндукции, согласно формуле (3.53), — отрицательный знак. Это означает, что при движении зарядов в контуре э. д. с. самоиндукции совершает отрицательную работу, т. е. тормозит движение зарядов и тем уменьшает силу тока в контуре. Явление происходит таким образом, как будто э. д. с. самоиндукции возбуждает в контуре индукционный ток, направленный против основного тока или возбуждает электрическое поле $\mathbf{E}_{\text{самоинд}}$, направленное против поля $\mathbf{E}$, создаваемого источником тока. Если же ток в проводнике убывает, то $\mathrm{d}I/\mathrm{d}t$ имеет отрицательный, а э. д. с. самоиндукции — положительный знаки. Следовательно, работа этой э. д. с. будет положительной, а возбуждаемые ею индукционный ток и электрическое поле $\mathbf{E}_{\text{самоинд}}$ будут иметь одинаковые направления с основным током и направлением $\mathbf{E}$. Таким образом, и *при возрастании и при убывании тока э. д. с. самоиндукции противодействует той причине, которая вызвала ее,* а именно препятствует изменению основного тока; э. д. с. самоиндукции (или вызываемые ею индукционные токи) является

противодействующей реакцией контура с током на внешнее воздействие, изменяющее ток в контуре (закон Ленца, см. § 22).

Расчет переменных токов в контурах в общем случае представляет сложную задачу. Изменение силы тока, вызванное в данный момент времени в каком-нибудь месте контура (например, возле полюсов источника тока), распространяется по проводникам с некоторой *конечной скоростью*; поэтому если эти изменения происходят очень быстро и размеры контура очень велики, то сила тока в контуре не успевает быстро выравниваться и в различных местах контура будет различной.

Будем интересоваться только контурами малых размеров и относительно медленными изменениями токов; тогда силу тока в данный момент времени можно считать во всех местах контура одинаковой (такие токи называются *квазистационарными*). При этих условиях можно применить законы Ома и Кирхгофа, учитывая существование э. д. с. самоиндукции. Для простого (неразветвленного) контура с сопротивлением R полное падение напряжения IR в каждый момент времени равно сумме э. д. с., действующих в этой цепи:

$$IR = \mathscr{E} + \mathscr{E}_{\text{самоинд}}; \quad I = \frac{\mathscr{E} - \frac{\mathrm{d}}{\mathrm{d}t}(LI)}{R}. \tag{3.57}$$

В частном случае, когда $L = \text{const}$,

$$I = \frac{\mathscr{E} - L\frac{\mathrm{d}I}{\mathrm{d}t}}{R}. \tag{3.58}$$

Таким образом, ток I состоит из двух частей: тока $I_и = \mathscr{E}/R$, вызываемого источником, и индукционного тока $I_с = -(L/R)\cdot(\mathrm{d}I/\mathrm{d}t)$, создаваемого э. д. с. самоиндукции.

В некоторых случаях э. д. с. самоиндукции может быть значительно больше, чем электродвижущая сила основного источника тока. Например, при размыкании токов $\mathrm{d}I/\mathrm{d}t$ приобретает очень большое значение, вследствие чего даже при малых значениях индуктивности L в контуре возбуждаются большие э. д. с. самоиндукции. Вызываемые при замыкании или размыкании индукционные токи иногда называются *экстратоками*.

Индуктивность проводников L зависит от их формы и размеров. Расчетная формула для длинного соленоида с плотной обмоткой и однородным сердечником имеет вид выражения (3.51); перепишем ее в виде

$$L = k\mu\frac{4\pi n^2}{l}S, \tag{3.59}$$

где коэффициенту k будем придавать различное значение в зависимости от выбранной системы единиц. В Международной системе

$$k = \mu_0/4\pi \tag{3.60}$$

и формула (3.59) совпадает с выражением (3.51). Единицей индуктивности является генри (Г), который можно определить, например, на основании формулы (3.53): г е н р и — *это индуктивность такого*

проводника, на концах которого возбуждается э. д.с. самоиндукции в один вольт, если сила тока в нем равномерно изменяется со скоростью, равной одному амперу в секунду. Согласно формуле (3.55), генри есть индуктивность проводника, с которым при силе тока в один ампер «связан» магнитный поток в один вебер. Таким образом,

$$1 \text{ генри} = \frac{1 \text{ вольт}}{1 \text{ ампер/секунду}} = 1 \frac{\text{вебер}}{\text{ампер}} \text{ или } 1 \text{ Г} = 1 \text{ Вб/А}.$$

В связи с этим магнитная проницаемость вакуума μ_0 выражается в единицах генри, деленного на метр:

$$\mu_0 = 4\pi \cdot 10^{-7} \text{ Г/м}; \quad k = \frac{\mu_0}{4\pi} = 10^{-7} \text{ Г/м}.$$

Приведем некоторые расчетные формулы для индуктивности:

1) прямолинейный одиночный провод круглого сечения длиной l

$$L = k \cdot 2\mu l \ln \frac{2l}{r} \tag{3.61}$$

(r — радиус сечения);

2) соленоид длиной l, имеющий n витков радиуса R,

$$L = k\mu \frac{4\pi n^2}{l^2} S \left(\sqrt{l^2 + R^2} - R\right) \tag{3.62}$$

($S = \pi R^2$ — площадь витка). При $l \gg R$ эта формула переходит в выражение (3.59);

3) проводник круглого сечения (радиуса r), имеющий форму окружности радиуса R,

$$L = k\mu 4\pi R \sqrt{\ln \frac{8R}{r} - 2}.$$

В Международной системе единиц (СИ) k имеет значение (3.60), в системе СГС полагается $k = 1$.

ВЗАИМНАЯ ИНДУКЦИЯ

Рассмотрим процесс *взаимной индукции* между двумя проводниками, изображенными на рис. III.69 в виде двух катушек с витками n_1 и n_2, намотанных на общий сердечник с магнитной проницаемостью μ. Допустим, что по первой катушке пропускается переменный ток I_1; создаваемое им переменное магнитное поле будет действовать на свободные заряды второй катушки и приведет их в движение. Если второй проводник разомкнут, то движению зарядов в нем будет препятствовать электрическое поле, возникающее при перераспределении зарядов в проводнике. Если на одном конце скапливаются электроны, то в противоположном конце «оголяются» положительные заряды ионной решетки металла. Достаточно незначительного перераспределения зарядов в проводнике, чтобы внутри них появились электрические силы eE, приостанавливающие движение электронов; поэтому в разомкнутом контуре возбуждается весьма слабый переменный ток.

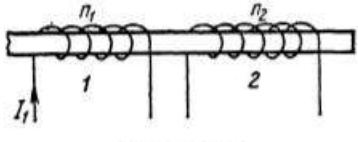


Рис. III.69

Разность потенциалов на концах этого контура будет изменяться со временем вместе с изменением тока в первом проводнике.

Если же вторая катушка замкнута, то переменная э. д. с. вызовет в ней переменный ток. Э. д. с. индукции, действующая в этой катушке, может быть определена по закону Фарадея (см. формулу (3.48)); для этого необходимо найти магнитный поток Φ_2, охватываемый витками катушки, и определить скорость его изменения $d\Phi_2/dt$.

Очевидно, Φ_2 зависит:

1) от напряженностей магнитного поля, которые создаются первой катушкой в тех местах, где находятся элементы второй катушки. Эти напряженности определяются силой тока I_1 в первой катушке, формой и размерами обеих катушек и их относительным расположением; если вторая катушка состоит из n_2 одинаковых витков, охватывающих один и тот же магнитный поток Φ, то $\Phi_2 = n_2\Phi$;

2) от магнитных свойств окружающих тел (или среды), так как магнитный поток Φ_2 определяется не напряженностью поля H, создаваемой электрическими токами, а индукцией B.

Все указанные выше факторы, кроме силы тока I_1, учитывают одной величиной L_{12}, которую называют *взаимной индуктивностью* или *коэффициентом взаимной индукции*. Тогда магнитный поток, связанный со вторым контуром, можно представить в виде, аналогичном выражению (3.56):

$$\Phi_2 = L_{12}I_1. \tag{3.63}$$

По закону электромагнитной индукции Фарадея, э. д. с. взаимной индукции, возбуждаемая во втором контуре переменным магнитным полем первого контура, равна

$$\mathscr{E}_2 = -\frac{d\Phi_2}{dt} = -\frac{d}{dt}(L_{12}I_1). \tag{3.64}$$

В частном случае, если форма, размеры и относительное расположение проводников, а также магнитные проницаемости тел (например, сердечника, на который намотаны проводники) остаются постоянными, т. е. $L_{12} = \text{const}$,

$$\mathscr{E}_2 = -L_{12}\frac{dI_1}{dt}. \tag{3.65}$$

Обозначим сопротивление второго контура через R_2; тогда сила тока в этом контуре, вызванная э. д. с. взаимной индукции $\mathscr{E}_2$, будет равна

$$I_2 = \frac{\mathscr{E}_2}{R_2} = -\frac{1}{R_2}\frac{d}{dt}(L_{12}I_1).$$

Переменный ток I_2, появившийся во втором контуре, создает свое собственное магнитное поле и поэтому может в свою очередь вызвать э. д. с. индукции в первом контуре. Введем коэффициент L_{21}, учитывающий форму, размеры и относительное расположение контуров, а также магнитные свойства окружающих тел. Тогда магнитный поток Φ_1 и электродвижущая сила $\mathscr{E}_1$, вызываемые в первом контуре переменным током второго контура, равны:

$$\Phi_1 = L_{21}I_2; \quad \mathscr{E}_1 = -\frac{d\Phi_1}{dt} = -\frac{d}{dt}(L_{21}I_2). \tag{3.66}$$

Таким образом, между контурами имеется взаимная «электромагнитная» связь. Всякие изменения, происходящие в одном контуре, в той или иной степени отражаются на другом контуре. Аналитически эту связь можно записать в виде системы уравнений, выражающих правила Кирхгофа для взаимодействующих контуров. С учетом самоиндукции и взаимоиндукции, полагая что токи являются квазистационарными, можно написать правило Кирхгофа для первого контура:

$$I_1R_1 = \mathscr{E} - \frac{\mathrm{d}}{\mathrm{d}t}(L_1I_1) - \frac{\mathrm{d}}{\mathrm{d}t}(L_{21}I_2). \qquad (3.67)$$

Если во втором контуре источника тока нет, то в ней действуют только э. д. с. самоиндукции и взаимоиндукции:

$$I_2R_2 = -\frac{\mathrm{d}}{\mathrm{d}t}(L_{21}I_2) - \frac{\mathrm{d}}{\mathrm{d}t}(L_{12}I). \qquad (3.68)$$

Коэффициенты взаимной индукции L_{12} и L_{21} являются коэффициентами пропорциональности между магнитным потоком и силой тока, поэтому они должны выражаться в тех же единицах, что и коэффициент самоиндукции, т. е. в генри (система СИ) или в сантиметрах (система СГС).

Допустим, что катушки, изображенные на рис. III.69, представляют собой «длинные соленоиды». Если сердечник сделан из немагнитного материала ($\mu \approx 1$), то магнитные поля катушек сосредоточены преимущественно в их объемах; тогда взаимодействие между катушками будет очень слабым. Связь между ними можно было усилить, поместив одну катушку внутри другой.

Если же сердечник сделан из ферромагнитного вещества, то в осуществлении связи между катушками будет принимать участие также и сильное магнитное поле намагниченного сердечника. Изменение магнитного поля одной катушки вызывает изменение намагниченности стержня по всей его длине, следовательно, и в том месте, где находится другая катушка. Если пренебрегать рассеянием магнитного потока вдоль сердечника (см. § 26), то можно считать, что через сечения обеих катушек проходит один и тот же магнитный поток Ф, который «охватывается» первой катушкой n_1 раз, а второй катушкой n_2 раз, поэтому

$$\Phi_1 = n_1\Phi; \quad \Phi_2 = n_2\Phi. \qquad (3.69)$$

Электродвижущие силы индукции, возбуждаемые в катушках вследствие изменения Ф со временем, равны:

$$\mathscr{E}_1 = -n_1\frac{\mathrm{d}\Phi}{\mathrm{d}t}; \quad \mathscr{E}_2 = -n_2\frac{\mathrm{d}\Phi}{\mathrm{d}t}; \quad \frac{\mathscr{E}_1}{\mathscr{E}_2} = \frac{n_1}{n_2}. \qquad (3.70)$$

Так как в намагничивании сердечника, т. е. в создании магнитного потока Ф, участвуют своими токами I_1 и I_2 обе катушки, то $\mathscr{E}_1$ и $\mathscr{E}_2$ являются здесь суммарными э. д. с. самоиндукции и взаимоиндукции. Тогда формулы (3.67) и (3.68) можно переписать так:

$$I_1R_1 = \mathscr{E} - \mathscr{E}_1; \quad I_2R_2 = -\mathscr{E}_2. \qquad (3.71)$$

Здесь все токи и э. д. с. являются функциями времени, поэтому эти соотношения соблюдаются для каждого определенного момента времени. Если падение напряжения в первой катушке мало, т. е. $I_1R_1 \approx 0$, то $\mathscr{E}_1 \approx \mathscr{E}$ и тогда из формулы (3.70) следует, что отношение э. д. с. источника тока, действующего в первой катушке, к э. д. с. индукции, действующей во второй катушке, равно отношению соответствующих чисел витков. Таким образом, система из двух катушек, соединенных ферромагнитным сердечником, представляет собой *трансформатор*, при помощи которого можно повышать или понижать напряжение переменного тока. Трансформатор с замкнутым сердечником изображен на рис. III.70.

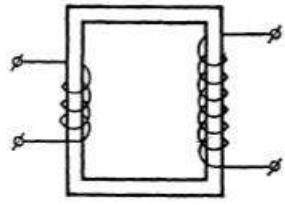


Рис. III.70

ЭНЕРГИЯ МАГНИТНОГО ПОЛЯ

В явлениях взаимоиндукции во втором контуре наблюдается появление энергии: при замыкании этого контура на сопротивление R_2 в нем за время $\mathrm{d}t$ при силе тока I_2 выделяется теплота $\mathrm{d}Q = I_2^2R_2\mathrm{d}t$. Согласно закону сохранения энергии, эта энергия может быть получена только извне (в самом контуре не происходят изменения, которые могли бы служить источником энергии). Наблюдения показывают, что при замыкании второго контура, т. е. при появлении, а также при изменении силы тока в нем изменяется расход энергии электрического тока в первом контуре. Аналитически это видно из соотношения (3.67). Таким образом, можно утверждать, что источником энергии для второго контура является энергия электрического тока, протекающего в первом контуре. Так как взаимодействие контуров осуществляется только благодаря магнитному полю, то можно предполагать, что это поле является «носителем» энергии и что от первого контура ко второму переходит энергия магнитного поля.

Энергия магнитного поля должна зависеть от значения напряженностей H или индукции $\mathbf{B}$ в различных местах этого поля. Для магнитного поля эти величины пропорциональны силе тока, поэтому всякое увеличение или уменьшение силы тока в проводнике должно означать соответствующее изменение энергии магнитного поля, связанного с этим током. Для расчета энергии магнитного поля рассмотрим одиночный контур, сопротивление которого равно нулю, а индуктивность равна L. Допустим, что в некоторый момент времени этот контур замыкается на источник тока. При возрастании тока в контуре от нуля до I источник тока совершает некоторую работу против э. д. с. самоиндукции, которая, по закону Ленца, препятствует возрастанию тока. Согласно закону сохранения энергии, эта работа должна вызвать в контуре эквивалентное изменение энергии. Так как сопротивление контура предположено равным нулю и в самом контуре никаких физических процессов, кроме появления магнитного поля, не происходит, то работу источника тока придется приравнять энергии появившегося

магнитного поля. За время dt по контуру будет перенесен заряд $dq = I dt$ и совершена работа

$$dW = -\mathcal{E}_{\text{самоинд}}\, dq = L \frac{dI}{dt} I\, dt = LI\, dI. \tag{3.72}$$

Полная работа по увеличению силы тока от нуля до I будет равна

$$W = \int_0^I LI\, dI = \frac{LI^2}{2}. \tag{3.73}$$

(При этом предполагается, что индуктивность контура не изменяется с изменением силы тока.)

Заметим, что эта работа не зависит ни от времени, в течение которого ток в проводнике изменялся от нуля до I, ни от того, как возрастал этот ток и сохранялась ли постоянной электродвижущая сила источника тока. Если ток I, текущий в проводнике в каком-нибудь определенном направлении, убывает до нуля, то dI/dt имеет отрицательный знак, а электродвижущая сила самоиндукции совершает положительную работу, равную выражению (3.73). При всяком изменении силы тока в проводнике W увеличивается ($I_2 > I_1$) или уменьшается ($I_2 < I_1$) на величину

$$\Delta W = \frac{LI_2^2}{2} - \frac{LI_1^2}{2}.$$

Таким образом $W = LI^2/2$ есть *энергия магнитного поля*, окружающего проводник с индуктивностью L, по которому течет ток I. Увеличение этой энергии возможно за счет положительной работы источника тока, уменьшение же этой энергии означает, что магнитное поле тока или электродвижущая сила самоиндукции совершает положительную внешнюю работу.

Энергия магнитного поля (как и само поле) распределяется по всему пространству, окружающему проводник с током. Поэтому можно найти связь между энергией, содержащейся в единице объема поля, и напряженностью магнитного поля. Эту связь можно установить, рассмотрев любой проводник с током, для которого рассчитывается L и напряженность поля. Наиболее удобен для этой цели тороид или же очень длинный соленоид с плотной обмоткой, у которого магнитное поле, а следовательно, и энергия поля сосредоточены в пределах объема Sl (S —площадь витка, l — длина соленоида), а напряженность поля в пределах этого объема с удовлетворительной точностью можно считать везде одинаковой. Тогда на основании выражений (3.51) и (3.12) для L и H имеем

$$W = \frac{1}{2} \mu_0 \mu \frac{n^2}{l} S \frac{H^2 l^2}{n^2} = \frac{\mu_0 \mu H^2}{2} Sl.$$

Следовательно, в единице объема магнитного поля содержится энергия

$$w = \frac{\mu_0 \mu H^2}{2} = \frac{B^2}{2\mu_0 \mu} = \frac{HB}{2}. \tag{3.74}$$

Это выражение для *плотности энергии магнитного поля* применимо и к полю, созданному проводниками любой формы, а также к полю постоянных магнитов.

ТОКИ ФУКО

Индукционные токи могут быть возбуждены не только в тонких проводах, обычно применяемых в электро- и радиотехнике, но и в массивных проводниках больших размеров: в кусках металла, в барабанах и дисках приборов, сердечниках катушек и т. д. Токи внутри таких сплошных тел представляют собой вихревое движение свободных зарядов (электронов); поэтому индукционные токи в массивных телах (токи Фуко) иногда называются *вихревыми*.

Для таких токов сопротивление проводника очень мало, так как линии токов коротки, а «поперечные сечения» проводника велики. Поэтому даже при возбуждении слабых э. д. с. индукции вихревые токи достигают больших значений и могут вызвать сильное нагревание тела. Особенно большие значения вихревых токов можно получить при помещении массивных тел в быстро меняющееся магнитное поле, вызываемое токами высокой частоты; в этих случаях $\mathrm{d}I/\mathrm{d}t$, а следовательно, и э. д. с. индукции получаются большими. Вихревые токи используются, например, для нагревания или плавления проводников, особенно в вакууме.

Взаимодействие между вихревыми токами в телах и внешним магнитным полем можно использовать для того, чтобы привести в движение тела или тормозить их, а также для «успокоения» колебаний. Например, если намагниченный стержень (или стрелка) колеблется вблизи массивного проводника, то в этом проводнике индуцируются токи, взаимодействие которых с колеблющимся стержнем, по закону Ленца, противодействует этим колебаниям. Однако мы не имели бы еще затухания колебаний, если бы при этом энергия индукционных токов не превращалась бы в тепло.

§ 24. ТОК СМЕЩЕНИЯ И ЕГО МАГНИТНОЕ ПОЛЕ; ТЕОРИЯ МАКСВЕЛЛА

В § 23 указывалось, что в переменном магнитном поле индукционные токи могут быть вызваны и в разомкнутом контуре. Допустим, что участок $a - b$ проводника (рис. III.71) находится в переменном магнитном поле $B(t)$. Возбуждаемая в этом проводнике электродвижущая сила индукции перемещает электроны или к одному (*1*) или к другому (*2*) концу проводника в зависимости от того, возрастает или убывает B. Обозначим потенциалы этих концов φ_1 и φ_2. В каждый данный момент времени разность потенциалов $\varphi_1 - \varphi_2$ равна э. д. с. индукции. Количество приводимого в движение электричества q, а следовательно, и значения силы тока в различные моменты времени зависят от емкости контура: $q = C(\varphi_1 - \varphi_2)$; $I = \mathrm{d}q/\mathrm{d}t$. Если к концам *1* и *2* проводника присоединить обкладки конденсатора большой емкости (чтобы в расчетах можно было пренебрегать емкостью самого проводника), то при данной $\mathscr{E}_{\text{инд}}$ можно получить в этом разомкнутом

контуре значительные токи, сила которых

$$I = \frac{dq}{dt} = \frac{d}{dt}[C(\varphi_1 - \varphi_2)] = \frac{d}{dt} C\mathcal{E}_{инд} = C\frac{d\mathcal{E}_{инд}}{dt}$$

прямо пропорциональна емкости конденсатора и быстроте изменения э. д. с. индукции. Если $\mathcal{E}_{инд}$ постоянна, то на обкладках конденсатора поддерживается постоянная разность потенциалов, и тока в проводнике не будет.

В рассматриваемой системе упорядоченное движение зарядов (электронов) существует только в проводнике, соединяющем обкладки. Движущиеся в этом проводнике заряды создают свое собственное магнитное поле; величины, характеризующие это поле, изменяются вместе с изменением величины и направления индукционного тока I. Однако на обкладках конденсатора упорядоченное движение зарядов обрывается; возникает вопрос, обрывается ли там и магнитное поле, которое всегда связано с токами и является их важнейшим признаком. Максвелл сделал предположение (которое было впоследствии подтверждено опытом), что магнитное поле существует и между обкладками конденсатора, но только благодаря тому, что электрическое поле в этом месте изменяется с течением времени. *В пространстве, где существует переменное электрическое поле, возбуждается магнитное поле*, напряженность которого H в каждой точке по величине прямо пропорциональна скорости изменения вектора электрической индукции dD/dt в этой точке, а по направлению — лежит в плоскости, перпендикулярной к этому вектору (здесь и в дальнейшем dD/dt означает производную по времени при постоянных координатах x, y и z). Таким образом,

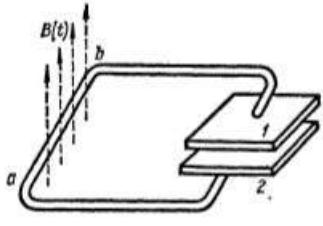


Рис. III.71

если напряженность электрического поля изменяется со временем, то это поле уже не является «чисто» электрическим полем; оно содержит в себе еще и магнитное поле, неразрывно с ним связанное и им самим вызванное. Такое поле называют электромагнитным.

Магнитное поле считается основным признаком электрического тока; это поле всегда существует вокруг отдельных движущихся зарядов, потоков заряженных частиц, в частности, вокруг тока проводимости в металлах. Там, где имеется электрический ток, всегда существует образованное им магнитное поле. Однако можно сделать и обратное утверждение: если есть магнитное поле, то оно обязательно вызвано каким-нибудь электрическим током. В частности, можно предположить, что указанное выше магнитное поле между обкладками конденсатора вызвано особым током; Максвелл назвал его *током смещения*. По его идее этот «ток смещения» не связан с каким-нибудь упорядоченным перемещением зарядов в пространстве между обклад-

ками; он может существовать и в вакууме, когда между обкладками нет никакого вещества. Важно только, чтобы имелось переменное электрическое поле, которое вызывает и «ток смещения», и связанное с ним магнитное поле.

Найдем формулу для расчета «тока смещения» в зависимости от скорости изменения электрического поля. Для нашего контура предположим, что расстояние между обкладками очень мало, поэтому электрическое поле однородное и сосредоточено только в объеме конденсатора (электрическим полем за пределами этого объема пренебрегаем). Тогда вектор индукции электрического поля в пространстве между обкладками $D = \sigma = q/S$, где q — заряд на одной обкладке; S — площадь обкладки. Так как D и q изменяются со временем, то найдем их производные:

$$\frac{dq}{dt} = \frac{dD}{dt} S.$$

Однако dq/dt — количество электричества, поступающее или покидающее обкладки конденсатора в единицу времени — есть сила тока в проводнике в данный момент времени (считаем, что она одинакова в любом сечении контура). Согласно предположению Максвелла, в разомкнутых контурах «токи смещения» замыкают собою токи проводимости и по величине равны им. Следовательно, $I = dq/dt$ есть также и сила тока смещения, а $I/S = j_{см}$ есть плотность этого тока в пространстве между обкладками; таким образом,

$$j_{см} = \frac{dD}{dt}. \qquad (3.75)$$

Эта формула применяется во всех случаях, где имеется переменное электрическое поле. Таким образом, любое изменение электрического поля в вакууме или в веществе создает «ток смещения» и связанное с ним магнитное поле.

Однако упорядоченное движение зарядов может существовать не только в проводнике, но и в диэлектрике. Если между обкладками конденсатора находится диэлектрик, то $D = \varepsilon_0 E + P$, где E — напряженность электрического поля; P — вектор поляризации (см. § 6). Тогда

$$j_{см} = \frac{dD}{dt} = \varepsilon_0 \frac{dE}{dt} + \frac{dP}{dt}.$$

Изменение вектора поляризации со временем dP/dt характеризует некоторое упорядоченное движение связанных зарядов в диэлектрике, которое называют *током поляризации*; поэтому «чистым» током смещения, который был введен Максвеллом, остается $\varepsilon_0 dE/dt$; *этот ток не выделяет тепла, но создает магнитное поле.*

С другой стороны, «ток смещения» (хотя и очень слабый), может существовать не только в вакууме или в диэлектриках, но и в проводниках, если внутри них имеется *переменное* электрическое поле. Допустим, что по проводнику течет синусоидальный ток, плотность которого $j = j_0 \sin\omega t$ (ω — угловая частота). Напряженность электрического поля E связана с плотностью тока в проводнике законом Ома:

$j = E/\rho$ (ρ — удельное сопротивление), следовательно, $E = \rho j = \rho j_0 \sin\omega t = E_0 \sin\omega t$. Плотность чистого тока смещения равна:

$$j_{см} = \varepsilon_0 \frac{dE}{dt} = \varepsilon_0 \omega E_0 \cos\omega t = \varepsilon_0 j_0 \rho\omega \cos\omega t.$$

Для металлических проводников даже при больших частотах $\rho\omega$ представляет весьма малую величину, поэтому отношение амплитуд $j_{см}/j \ll 1$.

Максвелл ввел также понятие *полного тока*, равного сумме тока проводимости и тока смещения:

$$I_{полн} = I + I_{см}. \qquad (3.76)$$

В разомкнутом электрическом контуре полный ток всегда замкнут, т. е. линии этого тока в некоторой части связаны с упорядоченным движением зарядов, в другой части — с переменным электрическим полем. В вакууме ($\rho = \infty$) токи проводимости и поляризации отсутствуют; существует только ток смещения.

Рассмотрим теперь вопрос о том, какова природа тех сил, с которыми переменное магнитное поле B действует на свободные заряды проводника (участок a — b на рис. 3.71) и вызывает в контуре индукционные токи. Обозначим силу, действующую на заряд e, через f. Отношение f/e можно назвать *напряженностью поля индукции*; необходимо выяснить, что представляет собой это «поле индукции», является ли оно электрическим полем или это есть какое-то новое поле, действующее на заряды сторонними (не электрическими) силами.

Максвелл предположил, что «поле индукции» $E_{инд} = f/e$ есть электрическое поле. В этом случае явление электромагнитной индукции заключается в том, что переменное магнитное поле возбуждает в проводниках электрическое поле, которое приводит в движение свободные заряды и тем самым вызывает индукционные токи. Однако Максвелл обобщил этот закон, предположив, что электрическое поле $E_{инд}$ возбуждается не только в веществах (где имеются заряды), но и в вакууме; по Максвеллу, в *пространстве, где существует переменное магнитное поле, возбуждается электрическое поле*, напряженность которого $E_{инд}$ в каждой точке по величине прямо пропорциональна скорости измерения вектора индукции магнитного поля dB/dt и лежит в плоскости, перпендикулярной этому вектору. Следовательно,

---

меняющееся со временем магнитное поле уже не является «чистым» магнитным полем; оно содержит в себе еще и электрическое поле, неразрывно с ним связанное и им самим созданное. Это поле также является электромагнитным.

Запишем приведенные выше основные положения теории электромагнитных явлений Максвелла в виде уравнений. Так как магнитное поле создается не только движущимися зарядами (токами проводимости), но и переменным электрическим полем (токами смещения), то циркуляция магнитного поля $\oint \mathbf{H}d\mathbf{l}$ должна быть приравнена «пол-

ному току» (см. формулу (3.76)):

$$\oint \mathbf{H}\, d\mathbf{l} = I_{\text{полн}} = I + I_{\text{см}}. \qquad (3.77)$$

Это есть обобщенный закон полного тока.

Формулу, аналогичную выражению (3.77), можно получить и для электрического поля. Для этого возьмем различные замкнутые линии и рассчитаем для них циркуляцию напряженности поля $\oint \mathbf{E}d\mathbf{l}$, т. е. найдем сумму произведений $E\, dl \cos\alpha$, где α — угол между $\mathbf{E}$ и $d\mathbf{l}$ ($d\mathbf{l}$ — участок линии, ориентированный по направлению обхода). Если электрическое поле создано только неподвижными зарядами, то эта циркуляция оказывается всегда равной нулю (при любом расположении зарядов и для любой замкнутой линии обхода). Вокруг и внутри контуров с токами эта циркуляция отлична от нуля.

Если линией обхода является замкнутая линия, вдоль которой происходит упорядоченное движение зарядов в проводниках контура, то рассчитанная вдоль этой линии циркуляция $\oint \mathbf{E}d\mathbf{l}$ равна суммарной э. д. с. источников тока, включенных в этот контур. Действительно, $\mathbf{F} = e\mathbf{E}$ есть сила, приложенная к заряду e (электрона), $\mathbf{F}\, d\mathbf{l} = e\mathbf{E}d\mathbf{l}$ — работа переноса этого заряда на расстояние dl, а $\oint \mathbf{F}d\mathbf{l} = e\oint \mathbf{E}d\mathbf{l}$ — работа переноса по всему замкнутому контуру. Отношение этой работы к переносимому заряду есть, по определению, суммарная (полная) э. д. с., действующая в этом контуре. Из этой полной э. д. с. можно выделить э. д. с. самоиндукции ($\mathscr{E}_{\text{полн}} = \mathscr{E} + \mathscr{E}_{\text{инд}}$) и тогда

$$\oint \mathbf{E}\, d\mathbf{l} = \mathscr{E} + \mathscr{E}_{\text{инд}} = \mathscr{E} - \frac{d\Phi}{dt}. \qquad (3.78)$$

По теории Максвелла эта циркуляция отлична от нуля и для любой замкнутой линии, проведенной в вакууме, если только эта линия находится в переменном магнитном поле. Для электрического поля, вызванного в вакууме переменным магнитным полем ($\mathscr{E} = 0$; $\mathscr{E}_{\text{инд}} = -d\Phi/dt$), имеем

$$\oint \mathbf{E}\, d\mathbf{l} = -\frac{d\Phi}{dt}, \qquad (3.79)$$

где $d\Phi/dt$ — скорость изменения потока магнитной индукции через площадь, охватываемую линией обхода.

Написанные выше интегральные уравнения (3.77), (3.78) и (3.79) используются в различных расчетах. Важным результатом теории Максвелла является предсказание им существования электромагнитных волн в пространстве, сделанное при математической обработке уравнений (3.77) и (3.78). При отсутствии проводников с токами эти уравнения выражают определенную связь между напряженностями электрического и магнитного полей:

$$\oint \mathbf{H}\, d\mathbf{l} = \frac{d(\varepsilon_0\varepsilon\mathbf{E})}{dt}\mathbf{S}; \quad \oint \mathbf{E}\, d\mathbf{l} = -\frac{d(\mu_0\mu\mathbf{H})}{dt}\mathbf{S}, \qquad (3.80)$$

где $\mathbf{S} = S\mathbf{n}$ ($\mathbf{n}$ — единичный вектор нормали к площадке S).

Опыты Г. Герца и изобретение радио А. С. Поповым подтвердили теоретическое предсказание Максвелла.

В современной физике электромагнитное поле рассматривается как особый вид материи, к которой применимы важнейшие понятия физики: энергия, импульс, масса. Электромагнитное поле, так же как и другие разновидности материи, можно представить себе состоящим из «частиц», каждая из которых обладает определенной энергией и массой. Представление об этих частицах (получивших название *фотонов*) оказалось особенно плодотворным при объяснении ряда явлений, в которых участвует высокочастотное электромагнитное поле — световое, рентгеновское и γ-излучение. Эти вопросы будут рассмотрены в четвертой части курса.

§ 25. МАГНИТНОЕ ПОЛЕ В ВЕЩЕСТВЕ. ДИАМАГНЕТИЗМ; ПАРАМАГНИТНЫЕ И ФЕРРОМАГНИТНЫЕ ВЕЩЕСТВА

В атомах и молекулах вещества имеются движущиеся заряды, на которые магнитное поле действует с силой Лоренца. Общее воздействие магнитного поля на тело определяется характером движения этих зарядов, их скоростями, формой и расположением траекторий их движения в объеме тела. Поэтому все вещества обладают магнитными свойствами; эти свойства обнаруживаются при введении тел в магнитное поле. При этом наблюдаются следующие явления:

1) *поле действует на тела с некоторыми силами* или механическими моментами. В неоднородном поле одни тела (*парамагнитные*) перемещаются в сторону возрастания напряженности поля, т. е. втягиваются в поле; другие тела (*диамагнитные*) перемещаются в сторону убывания напряженности поля, т. е. выталкиваются из поля. В однородном поле, на тела, имеющие продолговатую форму (стержни, стрелки), действуют механические моменты, которые устанавливают парамагнитные тела вдоль направления поля, а диамагнитные — перпендикулярно к нему;

2) *поле намагничивает вещество*, вследствие чего тела создают свое собственное магнитное поле, которое складывается с основным (намагничивающим) полем. У диамагнитных тел собственное поле направлено против внешнего поля, у парамагнитных тел собственное и внешнее поля имеют одинаковое направление.

Однако собственное магнитное поле намагниченного тела как внутри него, так и в окружающем пространстве, зависит не только от вещества, но и от размеров этого тела и от его формы (шар, стержень, замкнутое кольцо и т. д.). Поэтому при изучении магнитных свойств вещества желательно выбрать такие методы, в которых не было бы необходимости учитывать форму и размеры тел.

Задача упрощается, если исследуемое вещество полностью заполняет однородное магнитное поле; такое поле существует, например, внутри очень длинного соленоида с плотной обмоткой (поле за пределами его объема практически отсутствует). В этом случае вектор $\mathbf{B}$ магнитного поля в среде будет слагаться из векторов $\mathbf{B}_0 = \mu_0\mathbf{H}$ намагничивающего поля тока и вектора $\mathbf{B}'$ собственного магнитного поля намагнитившегося вещества:

$$\mathbf{B} = \mathbf{B}_0 + \mathbf{B}', \tag{3.81}$$

причем $\mathbf{B}'$ будет определяться только магнитными свойствами вещества (форма и размеры намагниченного тела практически не учитываются).

Вектор $\mathbf{B}'$ собственного магнитного поля среды можно представить в зависимости от вектора $\mathbf{B}_0$ намагничивающего поля:

$$\mathbf{B}'(x, y, z) = k_m \mathbf{B}_0(x, y, z). \tag{3.82}$$

Величина k_m называется *магнитной восприимчивостью* и является важной характеристикой магнитных свойств вещества. Тогда

$$\mathbf{B} = \mathbf{B}_0 + k_m \mathbf{B}_0 = (1 + k_m)\mathbf{B}_0 = \mu \mathbf{B}_0 = \mu_0 \mu \mathbf{H}. \tag{3.83}$$

Таким образом, относительная магнитная проницаемость среды μ связана с магнитной восприимчивостью соотношением

$$\mu = 1 + k_m. \tag{3.84}$$

У диамагнитных веществ k_m — отрицательная величина ($k_m < 0$, $\mu < 1$) и вектор $\mathbf{B}'$ направлен против вектора $\mathbf{B}_0$; у парамагнитных веществ k_m — положительная величина ($k_m > 0$, $\mu > 1$) и вектор $\mathbf{B}'$ имеет одинаковое направление с $\mathbf{B}_0$. Приведем значения магнитной восприимчивости для некоторых веществ:

Вещество	Магнитная восприимчивость	Вещество	Магнитная восприимчивость
Висмут	$-16{,}7 \cdot 10^{-5}$	Платина	$+2{,}93 \cdot 10^{-4}$
Серебро	$-2{,}64 \cdot 10^{-5}$	Алюминий	$+2{,}14 \cdot 10^{-5}$
Вода	$-0{,}88 \cdot 10^{-5}$	Жидкий кислород	$+3{,}46 \cdot 10^{-3}$
Водород	$-0{,}208 \cdot 10^{-8}$	Воздух	$+3{,}65 \cdot 10^{-7}$

НАМАГНИЧЕННОСТЬ

Намагниченность вещества оценивается по магнитному моменту $\mathbf{J}$, которым обладает единица объема вещества; это суммарный магнитный момент всех молекул, атомов и электронов, имеющихся в этом объеме. Вектор $\mathbf{J}$ связан с напряженностью $\mathbf{H}'$ или индукцией $\mathbf{B}'$ собственного магнитного поля внутри вещества. Для нахождения этой связи воспользуемся несколько неточными, но простыми рассуждениями, приводящими, однако, к правильному результату. Допустим, что намагниченность вещества обусловлена упорядоченным расположением электронных орбит в атомах относительно направления внешнего поля. Выделим внутри вещества цилиндрический объем с основанием S, равным площади орбиты, и длиной l, охватывающий большое число n орбит (витков). Этот цилиндр эквивалентен длинному соленоиду, поэтому среднюю напряженность H' магнитного поля внутри цилиндра можно рассчитать по формуле

$$H' = \frac{In}{l} = \frac{InS}{lS} = IS\frac{n}{lS} = p_m N,$$

где $N = n/lS$ — число орбит в единице объема вещества, а $p_m = IS$ — магнитный момент одной орбиты. Таким образом, $H' = p_m N$; $\mathbf{H}' = \mathbf{J}$.

Вектор **J** называется *вектором намагничения* или *намагниченностью* вещества.

Вектор магнитной индукции собственного магнитного поля вещества $\mathbf{B}'$, согласно формуле (3.82), равен $k_m\mathbf{B}_0'$; с другой стороны, $\mathbf{B}' = \mu_0\mathbf{H}' = \mu_0\mathbf{J}$, а $\mathbf{B}_0 = \mu_0\mathbf{H}$, поэтому вектор намагничения можно представить в зависимости от напряженности **H** внешнего намагничивающего поля:

$$\mathbf{J} = k_m\mathbf{H}. \tag{3.85}$$

Следовательно, если вещество находится во внешнем магнитном поле $\mathbf{H}(x, y, z)$, то внутри него индукция поля будет равна

$$\mathbf{B} = \mathbf{B}_0 + \mathbf{B}' = \mu_0\mathbf{H} + \mu_0\mathbf{J} = \mu_0(\mathbf{H} + \mathbf{J}). \tag{3.86}$$

Из этого уравнения можно определить H или J по измеренным значениям B.

Иногда намагниченность и магнитную восприимчивость определяют для единицы массы вещества; обозначим их $J_м$ и $k_{m_м}$. Очевидно, $J_м = \frac{1}{\rho}J$, $k_{m_м} = \frac{1}{\rho}k_m$, где ρ — плотность вещества.

Заметим, что в формулах (3.82) и (3.85) вовсе не предполагается постоянство магнитной восприимчивости; у некоторых сред она оказывается сложной функцией от B_0 (или H). В этом случае связь между собственным магнитным полем вещества B' и внешним намагничивающим полем B_0 по-прежнему запишется в виде формулы (3.82), но эта связь уже не будет линейной. Кроме того, k_m, а следовательно, и μ могут иметь различные значения в различных направлениях (например, в кристалле); среда будет неизотропной и векторы **B**, **H** и **J** могут иметь различные направления. Нелинейность и неизотропность намагничивающихся тел требует более детального исследования их магнитных свойств. Если для данного тела (вещества) при всевозможных значениях H в любом месте и направлении k_m и μ являются постоянными величинами, то тело (вещество) называется *линейным*, *однородным* и *изотропным*.

Важным для научных и технических применений являются ферромагнитные вещества, которые относятся к парамагнитным, но имеют очень большие значения магнитной восприимчивости и проницаемости. К ним относятся железо ($\mu = 5000 - 15\,000$), никель ($\mu \approx 300$), кобальт ($\mu \approx 180$) и различные магнитные сплавы, которые могут и не содержать железо; например, сплав из 61,5% меди, 23,5% марганца и 15% алюминия имеет $\mu = 235$. Таким образом, у ферромагнитных веществ значение вектора $\mathbf{B}'$ собственного магнитного поля во много раз превосходит значение вектора $\mathbf{B}_0$ внешнего намагничивающего поля, тогда как у диамагнитных и парамагнитных веществ собственное магнитное поле значительно слабее внешнего поля.

МАГНИТНЫЕ СВОЙСТВА АТОМОВ И МОЛЕКУЛ

Для объяснения магнитных свойств различных веществ необходимо рассмотреть действие магнитного поля на движущиеся заряды в молекулах и атомах вещества, в частности на электроны, вращающиеся по некоторым орбитам. Оси вращения электронов могут быть ориентированы под различными углами к направлению внеш-

него поля. Упростим рассуждения, заменив вращение ω по каждой орбите на два вращения: ω_1 — относительно оси, совпадающей с направлением поля, и ω_2 — относительно перпендикулярной оси (рис. III.72).

На орбиту с осью вращения ω_2 со стороны поля действует только поворачивающий момент. Суммируя эти моменты по всем электронным орбитам данного тела, можно найти результирующий вращающий момент, действующий на тело. Если тело не намагничено и в ориентировках электронных орбит нет никакого порядка, то этот результирующий момент будет равен нулю. Для намагниченного тела, имеющего некоторый порядок в расположении электронных орбит, вращающий момент отличен от нуля.

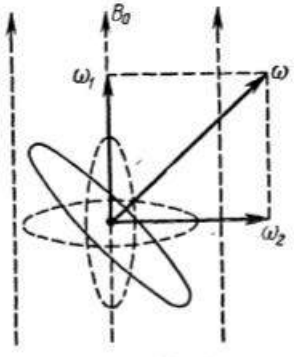


Рис. III.72

Выясним теперь действие магнитного поля на вращение электрона в плоскости, перпендикулярной полю. Допустим, что в однородном магнитном поле электрон вращается вокруг заряда q по окружности радиуса r со скоростью v_0. Кулоновская сила $F_к = keq/r^2$ является центростремительной, поэтому в отсутствие поля

$$\frac{mv_0^2}{r} = k\frac{eq}{r^2}; \quad v_0 = \sqrt{k\frac{eq}{rm}}. \qquad (3.87)$$

Магнитный момент орбиты, согласно формуле (3.28), равен

$$p_0 = \frac{1}{2}ev_0r. \qquad (3.88)$$

При наличии поля B_0 на электрон действует сила Лоренца $F_л = evB_0$; так как плоскость орбиты перпендикулярна B_0, то сила Лоренца, в зависимости от направления движения электрона по орбите, будет направлена либо от оси (рис. III.73, *а*), либо к оси вращения (рис. III.73, *б*). Тогда вместо формул (3.87) и (3.88) следует написать:

$$\frac{mv^2}{r} = k\frac{eq}{r^2} \pm evB_0; \quad p = \frac{1}{2}evr. \qquad (3.89)$$

Следовательно, скорость v движения электрона по орбите при наличии поля должна отличаться от скорости v_0 в отсутствие поля. Найдем разность этих скоростей из выражений (3.89) и (3.87):

$$\frac{mv^2}{r} = \frac{mv_0^2}{r} \pm evB_0;$$

$$\pm\frac{evB_0r}{m} = v^2 - v_0^2 \approx 2v(v - v_0); \quad v - v_0 \approx \pm\frac{erB_0}{2m}.$$

Изменение магнитного момента орбиты

$$\Delta p = p - p_0 = \frac{1}{2}er(v - v_0) \approx \frac{e^2r^2}{4m}B_0.$$

Следует подчеркнуть, что сила Лоренца, всегда перпендикулярная направлению движения электрона, не может изменить величину скорости электрона. Необходимое для соотношения (3.89) изменение скорости электрона осуществляется тем электрическим полем, которое возбуждается магнитным полем B_0. При «включении» этого поля (или при введении атома в это магнитное поле) значение B_0 для атома должно со временем изменяться от нуля до B_0. При этом, согласно теории электромагнитного поля (см. § 24), возбуждается электрическое поле, действующее на электрон и изменяющее его энергию. Можно показать, что изменение энергии электрона не зависит от быстроты установления поля, а определяется только конечным значением B_0. Суммарное действие магнитного и индуцированного им электрического поля приводит к увеличению скорости электрона при неизменном радиусе орбиты.

Докажем, что при любом направлении вращения электрона вектор Δp направлен *против* внешнего поля. У орбиты типа «а» силы $F_к$ и $F_л$ направлены в разные

стороны; в этом случае скорость движения электрона при наличии поля должна быть меньше, чем в отсутствие поля, т. е. $v < v_0$, и поэтому $p < p_0$; Δp направлено против p и, следовательно, против B_0. У орбиты типа «б» сила $F_л$ направлена к оси вращения,

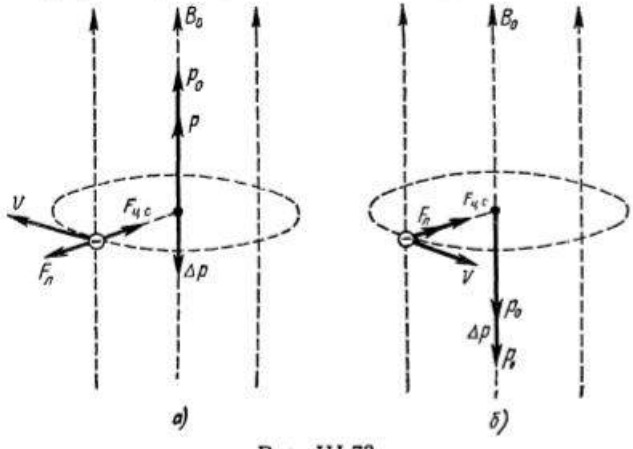


Рис. III.73

поэтому v должно быть больше v_0 и, следовательно, $p > p_0$; Δp направлено в ту же сторону, что и p, т. е. опять-таки против поля B_0. Таким образом, в векторной записи для каждой орбиты изменения магнитного момента

$$\Delta \mathbf{p} = -\frac{e^2 r^2}{4m} \mathbf{B}_0. \tag{3.90}$$

ДИАМАГНЕТИЗМ И ПАРАМАГНЕТИЗМ

Допустим, что суммарный магнитный момент электронных орбит каждого атома или молекулы данного вещества в отсутствие поля равен нулю. Это возможно, если в атомах имеются орбиты с противоположными направлениями вращения электронов, т. е. орбиты типов «а» и «б». Как было показано выше, при внесении любого вещества в магнитное поле каждая орбита независимо от направления движения электрона (т. е. независмо от ориентировки его магнитного момента $\mathbf{p}$ относительно $\mathbf{B}_0$) приобретает индуцированный магнитный момент $\Delta \mathbf{p}$, всегда направленный против поля. Таким образом, вещество окажется намагниченным; каждая единица объема вещества, содержащая N орбит, приобретает индуцированный магнитный момент, равный

$$J = \sum_N \Delta p_i = N\, \Delta p_{\text{сред}} = -N \frac{e^2}{4m} r^2_{\text{сред}} B_0. \tag{3.91}$$

Появление индуцированных магнитных моментов $\Delta \mathbf{p}_i$ и $\mathbf{J} = \Sigma \Delta \mathbf{p}_i$, направленных против внешнего поля, называется *диамагнитным эффектом*; им объясняются диамагнитные свойства вещества. Так как $J = k_m H$, $B_0 = \mu_0 H$, то диамагнитная восприимчивость вещества

$$k_m = -N \frac{e^2 r^2_{\text{сред}}}{4m} \mu_0. \tag{3.92}$$

Если вещество помещено в неоднородное магнитное поле, то на каждую орбиту (эквивалентную контуру тока, см. § 20) будет действовать некоторая сила. Орбита «а» имеет $\mathbf{p} \uparrow\uparrow \mathbf{B}$ и поэтому будет втягиваться в поле с силой $f_1 = (p_0 - \Delta p)\, dB/dx$; орбита «б» имеет $\mathbf{p} \uparrow\downarrow \mathbf{B}$ и будет выталкиваться из поля с силой $f_2 = (p_0 + \Delta p)\, dB/dx$. Так как $f_2 > f_1$, то на каждую пару одинаковых орбит с разными направлениями вращения действует сила $\Delta f = f_2 - f_1 = 2\Delta p\, dB/dx$, направленная в сторону убывания поля. Таким образом, *диамагнитный эффект приводит к появлению сил, выталкивающих тело из неоднородного магнитного поля.* Если в веществе число орбит типа «а» больше, чем орбит «б», то сумма сил f_1 больше суммы сил f_2 и вещество втягивается в магнитное поле, т. е. будет парамагнитным. Однако диамагнитный эффект существует и у парамагнитных тел (так как Δp появляются у всех орбит); оно выражается в некотором уменьшении той силы, которая втягивает парамагнитное тело в поле.

Для того чтобы внести диамагнитное тело в магнитное поле, необходимо совершить работу против выталкивающей силы. Эта работа совершается электродвижущей силой индукции, которая возбуждается в любом замкнутом контуре, а в данном случае — в электронных орбитах, при их движении в неоднородном магнитном поле. Для электронной орбиты э. д. с. индукции также равна $\mathcal{E}_{\text{инд}} = -d\Phi/dt$, где $d\Phi$ есть изменение потока вектора $\mathbf{B}$ через площадь орбиты. В результате действия этих э. д. с. увеличивается кинетическая энергия электронов в орбитах «б» и уменьшается — в орбитах «а». Если диамагнитное тело неподвижно и находится, например, внутри катушки, которая присоединяется к источнику тока, то работа диамагнитного намагничивания совершается источником тока, возбуждающим магнитное поле.

Итак:

1) *диамагнетизм есть свойство, присущее всем телам*, так как оно обусловлено действием внешнего магнитного поля на электронные орбиты в атомах и молекулах;

2) изменение скоростей движения электронов по орбитам сопровождается появлением магнитного поля, направленного против внешнего поля и ослабляющего его (закон Ленца); таким образом, *каждое вещество противодействует проникновению магнитного поля внутрь его объема*;

3) диамагнитный эффект не связан с появлением упорядоченности в расположении электронных орбит, поэтому *диамагнитная восприимчивость k_m не зависит от температуры*;

4) вещество является диамагнетиком, если только его атомы и молекулы не имеют собственного магнитного момента; тогда диамагнитный эффект является единственной реакцией вещества на воздействие внешнего магнитного поля. *В парамагнитных и ферромагнитных веществах намагниченность, обусловленная упорядоченным расположением элементарных магнитов (электронов, атомов и т. д.), значительно превосходит диамагнитный эффект.*

Парамагнитными свойствами обладают вещества, у которых атомы или молекулы в отсутствие внешнего поля имеют некоторый магнитный момент $\mathbf{p}$ (например, если они содержат нечетное число электронов).

Тепловое движение создает беспорядок в ориентировках их магнитных моментов, поэтому при отсутствии внешнего поля суммарная проекция векторов **p** на любое направление в веществе равна нулю. При наличии поля сумма проекций p на направление поля делается отличной от нуля, и тело намагничивается. Намагниченность **J** (суммарный магнитный момент единицы объема вещества) в данном случае будет зависеть не только от вектора $\mathbf{B}_0$ внешнего поля, но и от температуры. При очень низких температурах, когда дезориентирующее действие теплового движения весьма слабое, наблюдается *магнитное насыщение*, т. е. все элементарные магнитные моменты **p** ориентируются по полю, и дальнейшее усиление внешнего поля не вызывает увеличения намагниченности. При высоких температурах такое насыщение удается получить только в очень сильных магнитных полях. В слабых полях и при высоких температурах удельная магнитная восприимчивость парамагнитных веществ обратно пропорциональна температуре (закон Кюри):

$$k_m = C/T, \tag{3.93}$$

где C — постоянная, равная $Np^2/3k$ (N — число молекул в единице объема; k — постоянная Больцмана).

ФЕРРОМАГНЕТИКИ

В ферромагнитных веществах намагничение обусловлено не ориентацией электронных орбит по полю, а ориентацией собственного магнитного момента («спина») самих электронов $p_м = \mu_e$. В опытах А. Эйнштейна и Де-Гааза маленький ферромагнитный стержень, подвешенный на кварцевой нити, помещался внутрь соленоида, по которому пропускался переменный ток (рис. III.74). При перемагничивании стержень поворачивался то в одну, то в другую сторону. Этот поворот объясняется тем, что ось вращения каждого электрона при перемагничивании изменяется на 180°, поэтому механический момент импульса изменяется от $+p_{мех}$ до $-p_{мех}$, т. е. на $2p_{мех}$. Так как в стержне имеется много таких перевертывающихся электронов, то общее изменение их магнитных моментов будет равно $\Sigma 2p_м$, а механических моментов импульса $-\Sigma 2p_{мех}$. Измерение этих величин показало, что отношение $p_м/p_{мех}$ соответствует не электронной орбите (см. формулу (3.28)), а вдвое больше, т. е. соответствует спину электрона.

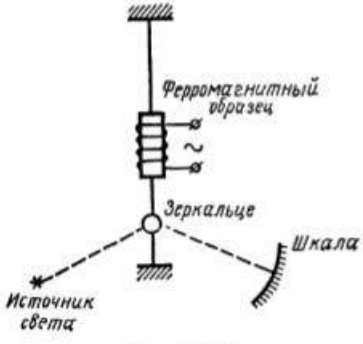


Рис. III.74

Известно следующее свойство ферромагнетиков: при температуре, называемой точкой Кюри, они скачком теряют свои ферромагнитные свойства и превращаются в парамагнитное вещество. В опытах А. Ф. Иоффе и П. Л. Капицы намагниченный сердечник

быстро нагревался до этой температуры; при размагничивании суммарный механический момент упорядоченно расположенных векторов $\mathbf{p}_{мех}$ изменялся от $\Sigma\mathbf{p}_{мех}$ до нуля. Согласно закону сохранения момента импульса весь стержень получал вращение. Измерения подтвердили, что отношение $p_м/p_{мех}$ соответствует спину электрона.

Отдельные атомы (например, железа) не имеют ферромагнитных свойств; эти свойства наблюдаются только у кристаллов ферромагнитных веществ при температурах ниже точки Кюри. Измерения показали, что работа, затрачиваемая на намагничение кристалла, различна в различных направлениях; существуют направления, вдоль которых эта работа имеет *минимальные* значения (анизотропия кристалла ферромагнетика оценивается разностью работ намагничения по различным направлениям); монокристалл кобальта имеет одну, железо — три, никель — четыре оси легчайшего намагничения. Каждый кристалл имеет большое число областей, намагниченных до насыщения в одном из направлений легчайшего намагничения. Эти области «спонтанного» (т. е. самопроизвольного) намагничения называются *доменами*. Определенная ориентация векторов $\mathbf{p}_0$ в пределах каждого домена обусловлена особыми силами. В двух соседних доменах вектора магнитных моментов $\mathbf{p}_0$ ориентированы различно (антипараллельно или под другими углами в зависимости от анизотропии кристалла). Между доменами существуют тонкие прослойки, в которых ориентация векторов $\mathbf{p}_0$ плавно меняется от направления намагниченности одного домена до направления намагниченности соседнего домена. Суммарный магнитный момент всех доменов в отсутствии поля равен нулю, что соответствует устойчивому состоянию кристалла (минимуму его энергии).

При наложении внешнего магнитного поля на электроны начинают действовать механические моменты, ориентирующие их по полю. По мере увеличения напряженности внешнего поля сначала происходит сравнительно легкий процесс поворота векторов $\mathbf{p}_0$ в прослойках между доменами. Это приводит к росту тех доменов, у которых угол между направлением намагничения и направлением внешнего поля оказался малым, за счет соседних доменов, у которых этот угол велик. Таким образом, постепенно происходит процесс разрушения доменной структуры, по окончании которого кристалл оказывается намагниченным только в одном направлении; это будет то направление легчайшего намагничения, которое составляет в данном случае наименьший угол с направлением поля. При дальнейшем увеличении напряженности внешнего поля происходит поворот суммарного вектора намагничения $\mathbf{J} = \Sigma\mathbf{p}_0$ (т. е. одновременный поворот всех $\mathbf{p}_0$) до совпадения с направлением внешнего поля. При обратном процессе, по мере ослабления внешнего поля, сначала происходит поворот вектора $\mathbf{J}$ до ближайшего направления легчайшего намагничения в кристалле, а затем происходит расслоение всего объема кристалла на домены, т. е. переход его в более устойчивое состояние.

Изучение зависимости намагничения от температуры привело к следующим представлениям. При температуре абсолютного нуля в пределах каждого домена все вектора спинов $\mathbf{p}_0$ ориентированы вдоль

той или иной оси легчайшего намагничения (такое состояние называется «абсолютным насыщением»). При повышении температуры некоторые электроны под действием теплового движения переходят в антипараллельное состояние, т. е. их вектор магнитного момента ориентируется против направления намагниченности домена, вследствие чего общая намагниченность каждого домена несколько уменьшается. С повышением температуры число таких электронов растет, особенно резко при приближении к температуре Кюри, при которой число параллельных и антипараллельных спинов становится равным и спонтанная намагниченность домена исчезает. Выше температуры Кюри $T_к$, когда ферромагнитное вещество становится парамагнитным, магнитная восприимчивость зависит от абсолютной температуры по формуле (закон Кюри — Вейса):

$$k_m = \frac{C}{T - T_к}, \tag{3.94}$$

где C — постоянная величина (для данного вещества). При температурах $T < T_к$ эта зависимость имеет сложный вид. Для железа $T_к = 768°$ С, для никеля — 358° С, для кобальта — 1120° С.

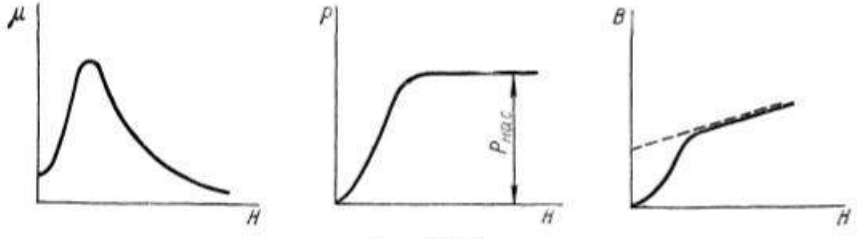


Рис. III.75

Перечислим важнейшие свойства ферромагнетиков:

1) ферромагнетики (в отличие от парамагнетиков) намагничиваются до насыщения уже в слабых полях;

2) магнитные проницаемости μ и восприимчивости k_m зависят от напряженности H внешнего намагничивающего поля. На рис. III.75 показаны зависимости μ, $J = k_m H$ и $B = \mu_0 \mu H$ от напряженности внешнего поля H;

3) ферромагнитные тела сохраняют состояние намагничения после удаления их из пределов намагничивающего поля (остаточный магнетизм, постоянные магниты);

4) при намагничивании и размагничивании ферромагнитные тела изменяют свои размеры; это явление называется *магнитострикцией*. Относительное изменение длины тела в направлении внешнего поля пропорционально квадрату намагниченности

$$\Delta l / l = \alpha J^2. \tag{3.95}$$

Магнитострикция (как и электрострикция) используется для получения и измерения ультразвука;

5) при перемагничивании обнаруживается «магнитный гистерезис», т. е. намагничивание происходит при одной зависимости B от H (кривые *1—2* и *4—5* на рис. III.76, *а*), а размагничивание — при другой зависимости (кривые *2—3—4* и *5—6—1*). Пунктирная линия *0—2* показывает намагничивание тела от начального состояния $B = 0$, когда тело полностью размагничено еще до введения его в магнитное поле. Увеличивая напряженность внешнего поля от $H = 0$ до некоторого значения, можно довести намагниченность тела до величины, соответствующей точке *2*. При уменьшении H (для размагничивания) индукция B уменьшается по кривой *2—3* и при $H = 0$ сохраняется остаточный магнетизм, измеряемый отрезком *0—3*. Для того чтобы размагнитить тело, необходимо приложить внешнее поле обратного направления с напряженностью H_0, соответствующей отрезку *0—4*; эта напряженность называется *коэрцитивной* (задерживающей) силой и характеризует магнитные свойства вещества. При дальнейшем

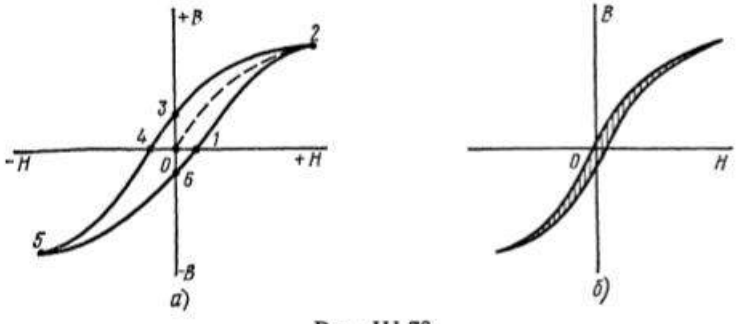


Рис. III.76

увеличении напряженности (обратного направления) происходит намагничивание по кривой *4—5*. Участок *5—6—1* снова соответствует размагничиванию. Таким образом, повторное намагничивание в первоначальном (положительном) направлении осуществляется уже не по кривой *0—2*, а по кривой *1—2*. В дальнейшем процесс перемагничивания происходит по замкнутой кривой, показанной на рис. III.76, *а* сплошной линией.

Площадь петли гистерезиса в масштабе чертежа означает *работу*, которую затрачивает внешнее поле для *одного перемагничивания* тела, преодолевая силы, препятствующие переориентировкам «областей намагничивания»; эта работа выделяется в виде тепла. Очевидно, для уменьшения потерь на это перемагничивание, например в сердечнике трансформатора, необходимо использовать «мягкие» ферромагнетики (рис. III.76, *б*), для которых работа перемагничивания мала. Наоборот, для изготовления постоянных магнитов подходят «жесткие» ферромагнетики с большим значением коэрцитивной силы.

Кроме ферромагнитного состояния, наблюдаются также *антиферромагнитные* и *ферритные* (или ферримагнитные) состояния. Схематическое расположение спиновых моментов в элементарных магнитных ячейках этих веществ при очень низких температурах ($T \ll T_к$)

указаны на рис. III.77. Для всех этих состояний характерна правильная ориентация спиновых магнитных моментов атомов, молекул или ионов данного вещества в ячейках его кристаллической решетки. У ферромагнитов спиновые моменты атомов в каждой элементарной ячейке кристалла направлены в одну сторону и создают некоторую намагниченность этой ячейки. У антиферромагнетиков такая намагниченность отсутствует, так как спиновые моменты соседних атомов равны по величине и направлены в противоположные стороны. У ферритов ориентация спиновых моментов соответствует антиферромагнитному, но магнитные моменты соседних атомов, направленные в противоположные стороны, не компенсируются.



Рис. III.77

К антиферромагнетикам относятся MnO, FeO, $FeCl_2$, CoO и др.; их магнитная восприимчивость с повышением температуры сначала увеличивается, достигает максимума при некоторой температуре (точка Кюри) и затем уменьшается. Ферриты имеют общую формулу $M \cdot O \cdot Fe_2O_3$, где M — ион двухвалентного металла (медный, никелевый, марганцевый и другие ферриты). Они обладают высокой магнитной восприимчивостью и проницаемостью (μ доходит до 6000), малой коэрцитивной силой, малыми потерями на перемагничивание и очень большим удельным электросопротивлением (до 10^4 Ом·м), благодаря чему они получили широкое применение в радиотехнике (потери на токи Фуко в ферритах, находящихся в высокочастотных электромагнитных полях, очень малы). Ферритовые сердечники применяются в электронно-счетных машинах, радиолокационных, телефонных, электроакустических и прочих аппаратурах. Точка Кюри у различных применяющихся ферритов лежит в пределах 300—600° С.

ПАРАМАГНИТНЫЙ РЕЗОНАНС

Для изучения магнитных свойств атомов и молекул вещества важное значение имеет явление *парамагнитного резонанса*, открытого и изученного Е. К. Завойским. Если парамагнитное вещество поместить в постоянное магнитное поле, то вектор магнитного момента молекулы будет совершать прецессию вокруг направления внешнего поля с некоторой частотой ω, зависящей от величины этого момента. Перпендикулярно к постоянному магнитному полю накладывается переменное магнитное поле высокой (радиотехнической) частоты Ω — от 10^7 до $3 \cdot 10^9$ герц и более. При помощи весьма чувствительного метода измеряется поглощение электромагнитной энергии веществом

при различных частотах переменного поля. Измерения показали, что кривая поглощения имеет резкий максимум при определенной частоте переменного поля; это резонансное поглощение соответствует точному равенству $\Omega = \omega$. Таким образом может быть определена частота прецессии ω, а затем вычислен магнитный момент молекулы. Этот метод «парамагнитного резонанса» получил широкое применение при измерении магнитных моментов ионов, атомов и атомных ядер, а также для изучения структуры и свойств различных твердых, жидких, органических веществ, живых тканей и т. д.

§ 26. МАГНИТНОЕ ПОЛЕ В ФЕРРОМАГНИТНЫХ ТЕЛАХ РАЗЛИЧНОЙ ФОРМЫ; МАГНИТОДВИЖУЩАЯ СИЛА И МАГНИТНОЕ СОПРОТИВЛЕНИЕ

Допустим, что n витков проводника с током намотаны на сердечник из магнитного вещества. Напряженность магнитного поля в каждой точке внутри сердечника будет состоять из напряженности поля тока $\mathbf{H}(x, y, z)$ и собственного поля намагнитившегося сердечника $\mathbf{H}'(x, y, z)$. Поле сердечника определяется магнитным моментом единицы объема, т. е. вектором намагниченности $\mathbf{J} = \mathbf{H}'$, который во

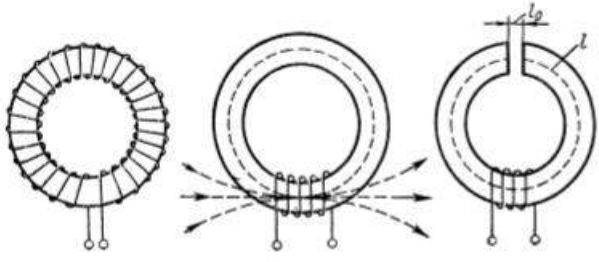


Рис. III 78

всех частях его объема должен быть известен. Однако этот вектор в каком-нибудь месте сердечника нельзя вычислять только по значению напряженности намагничивающего поля тока $\mathbf{H}$ в этом месте; соотношение (3.85) $\mathbf{J} = k_m\mathbf{H}$ применимо только в том случае, когда намагничивающее поле является однородным и целиком заполнено данным веществом. Например, из этой формулы следовало бы, что намагниченность J тороидального сердечника (рис. III.78, *б*) должна была быть большой внутри катушки (где H велика) и очень малой в тех частях сердечника, которые расположены вне катушки (где H мала); в действительности весь тороид оказывается намагниченным равномерно. Это объясняется тем, что намагниченные (внутри катушки) части сердечника своим собственным магнитным полем H' (значительно более сильным, чем поле катушки с током) намагничивают остальные участки сердечника, находящиеся вне поля катушки; намагниченность передается по объему сердечника от одних его участков к другим и весь сердечник оказывается намагниченным. Так как эле-

ментарные магниты (атомы, молекулы, домены) ориентируются благодаря взаимодействию между собой, то внутри сердечника формируются силовые линии собственного магнитного поля (вдоль которого направлены векторы **J** и **H′**), целиком лежащие внутри объема сердечника и поэтому подобные ему по форме (у тороида — окружности и т. д.). Таким образом, магнитное поле оказывается сосредоточенным в пределах объема сердечника.

Если, как это показано на рис. III.78, *а*, весь сердечник находится в однородном магнитном поле **H** соленоида с током, то в любом месте сердечника можно написать:

$$\mathbf{H}' = \mathbf{J} = k_m \mathbf{H}; \quad \mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{H}') = \mu_0 (\mathbf{H} + \mathbf{J}) = \mu_0 \mu \mathbf{H}. \qquad (3.96)$$

Если же, как это показано на рис. III.78, *б*, только часть сердечника находится внутри поля соленоида, то формула (3.96) может быть использована только для этой части сердечника. Однако и в этом случае намагничивающее поле катушки (приложенное только к некоторой части объема сердечника) фактически определяет намагниченность всего сердечника в целом, т. е. можно условно полагать, что поле катушки как бы действует по всему объему сердечника.

Для того чтобы формула (3.96) применялась во всех случаях, полагают, что во всем объеме сердечника напряженность намагничивающего поля равна

$$\mathbf{H} = \frac{\mathbf{B}}{\mu_0 \mu} = \frac{\mathbf{B}}{\mu_0} - \mathbf{J}. \qquad (3.97)$$

В той части сердечника, которая находится внутри катушки с током, рассчитанная по этой формуле напряженность поля будет равна напряженности магнитного поля самой катушки.

Если в тороиде имеется воздушный зазор l_0 (рис. III.78, *в*), то намагниченность сердечника уже не будет везде одинаковой. При очень малом зазоре один конец (полюс) сердечника будет оказывать достаточно сильное ориентирующее действие на элементарные магниты противоположного конца (полюса) и отклонение намагниченности концов от намагниченности остальных участков сердечника незначительно. По мере увеличения зазора взаимодействие полюсов через зазор ослабляется и ориентация элементарных магнитов в этих местах сердечника оказывается менее упорядоченной; магнитный момент единицы объема вещества (т. е. намагниченность **J**) будет на концах сердечника меньше, чем внутри катушки. При графическом изображении магнитного поля это обстоятельство выражается в том, что силовые линии суммарного поля, т. е. линии индукции $\mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{J})$, выходят за пределы сердечника через его боковую поверхность (рис. III.79). Кроме того, имеет место обратное влияние полюсов на намагниченность всего сердечника. При отсутствии зазора все элементарные магниты вещества будут (в состоянии насыщения) расположены упорядоченно вдоль определившихся силовых линий вектора **B** внутреннего магнитного поля. При наличии зазора некоторое нарушение упорядоченности на полюсах ослабляет ориентирующее дей-

ствие внутреннего поля во всем объеме сердечника; поэтому общая намагниченность сердечника уменьшается с увеличением зазора (или расстояния) между его полюсами.

Таким образом, намагниченность сердечника зависит не только от напряженности поля **H** и магнитных свойств веществ (μ, k_m), но и от формы и размеров тела, помещаемого в магнитное поле (шар, тороид сплошной или с зазором, прямой или подковообразный сердечник и т. д.). При этом появившиеся полюса намагниченного тела будут действовать (как и в случае зазора в тороиде) ослабляющим образом. Поэтому связь между напряженностью поля **H** и намагниченностью тела **J** вместо формулы (3.97) запишется в виде

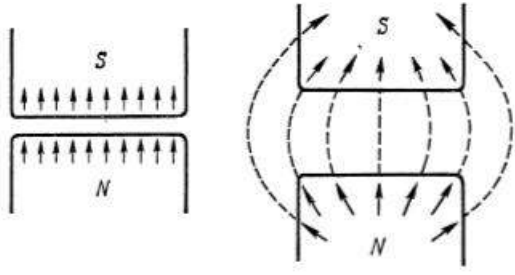


Рис. III.79

$$\mathbf{H} = \frac{\mathbf{B}}{\mu_0} - (1 - N)\,\mathbf{J} \quad \text{или} \quad \mathbf{B} = \mu_0 [\mathbf{H} + (1 - N)\,\mathbf{J}]. \qquad (3.98)$$

Величина N называется *размагничивающим фактором* данного тела; для шара $N = 1/3$; для цилиндра, намагничиваемого перпендикулярно оси, $N = 1/2$; для длинного стержня, намагничиваемого вдоль оси, $N \approx 0$. На рис. III.80 показана зависимость J от H: кривая a — для тонкого и длинного, b — для толстого и короткого стержней.

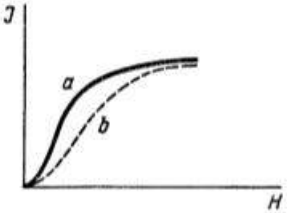


Рис. III.80

Рассчитаем магнитное поле внутри сердечника с воздушным зазором, полагая, что сердечник однородный, имеет форму кольца с одинаковым сечением S, а зазор l_0 настолько мал, что $N \approx 0$ (см. рис. III.78, *в*). Для применения закона полного тока (см. формулу (3.6)) выберем в качестве линии обхода L замкнутую силовую линию вектора напряженности поля **H**. Если длина этой линии внутри сердечника равна l, то циркуляция вектора **H** будет равна $\oint \mathbf{H}\,d\mathbf{l} = Hl + H_0 l_0$, где H_0 — напряженность поля в воздушном зазоре. Так как этот зазор очень мал, то индукция B суммарного поля внутри сердечника и в зазоре почти одинакова, поэтому $B \approx B_0$ и $H_0 = B_0/\mu_0 \approx B/\mu_0$; внутри же сердечника $H = B/\mu_0\mu$. Линия обхода L охватывает n витков с токами I, поэтому

$$\oint \mathbf{H}\,d\mathbf{l} = In; \quad Hl + H_0 l_0 = In. \qquad (3.99)$$

Заменив значения H и H_0 на индукцию поля, получим

$$B = \frac{In}{(l/\mu_0\mu) + (l_0/\mu_0)}, \quad \text{или} \quad \Phi = \frac{In}{(l/\mu_0\mu S) + (l_0/\mu_0 S)}, \qquad (3.100)$$

где $\Phi = BS$ — магнитный поток через сечение сердечника, а также и через воздушный зазор, так как магнитным полем за пределами зазора пренебрегаем.

Величина

$$F = \oint \mathbf{H}\, d\mathbf{l} = \sum_{i=1}^{n} I_i,$$

равная для соленоида с n витками произведению In, называется *магнитодвижущей силой* и выражается в амперах (или в ампер-витках). Величина $r_m = F/\Phi$, равная для однородного тела (или зазора) длиной l, сечением S и абсолютной магнитной проницаемостью $\mu_a = \mu_0\mu$

$$r_m = \frac{l}{\mu_0 \mu S},$$

называется *магнитным сопротивлением* этого тела (или зазора) и выражается в амперах на вебер (А/Вб), а также 1/Г, или 1/(Ом·с).

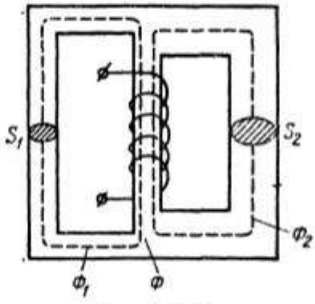


Рис. III.81

Если сердечник состоит из нескольких последовательных участков, имеющих различные l, S и μ, то магнитный поток

$$\Phi = \frac{F}{\sum r_m}; \quad \sum r_m = r_{1m} + r_{2m} + \ldots \quad (3.101)$$

Если сердечник замкнут несколькими телами, как это показано, например, на рис. III.81, то для каждого замыкающего тела можно выбрать отдельную линию обхода и тогда $F = \Phi_1 r_{1m} = \Phi_2 r_{2m}$; можно утверждать, что полный магнитный поток Φ, возбужденный магнитодвижущей силой $F = In$ катушки с током, разветвляется между «параллельно соединенными» магнитопроводами обратно пропорционально их магнитным сопротивлениям:

$$\Phi = \Phi_1 + \Phi_2; \quad \frac{\Phi_1}{\Phi_2} = \frac{r_{2m}}{r_{1m}}. \quad (3.102)$$

Если имеется несколько таких параллельно соединенных магнитопроводов, то при отсутствии рассеяния магнитного поля можно написать $\Phi = \Phi_1 + \Phi_2 + \ldots$; подставив сюда значения отдельных потоков в разветвлениях, получим

$$\Phi = \frac{F}{r_{1m}} + \frac{F}{r_{2m}} + \ldots = F\left(\frac{1}{r_{1m}} + \frac{1}{r_{2m}} + \ldots\right). \quad (3.103)$$

Полный магнитный поток Φ можно представить как F/r_m, где r_m — полное магнитное сопротивление разветвленной магнитной цепи. Тогда для «параллельного соединения» магнитных сопротивлений получим:

$$\frac{1}{r_m} = \frac{1}{r_{1m}} + \frac{1}{r_{2m}} + \ldots \quad (3.104)$$

Заметим, что полученные выше формулы для расчета магнитных систем аналогичны формулам, которые применяются для расчета электрических контуров.

Глава 4

ЭЛЕКТРОМАГНИТНЫЕ КОЛЕБАНИЯ И ВОЛНЫ

§ 27. КОЛЕБАТЕЛЬНЫЙ КОНТУР. НЕЗАТУХАЮЩИЕ И ЗАТУХАЮЩИЕ КОЛЕБАНИЯ

Электромагнитные колебания в контурах (периодические изменения силы тока и связанных с ней величин) обусловлены электрическими и магнитными свойствами этих контуров. Рассмотрим *электрический колебательный контур*, т. е. цепь, состоящую из катушки с индуктивностью L и конденсатора, обладающего емкостью C (рис. III.82); сопротивление контура обозначим через R. Если зарядить конденсатор до некоторой разности потенциалов U, сообщив его обкладками заряды $\pm q$, и затем замкнуть ключ K, то конденсатор начнет разряжаться через катушку и появившийся в цепи электрический ток при малом R будет очень быстро нарастать. Изменяющийся со временем электрический ток вызовет на концах катушки разность потенциалов, равную э. д. с. самоиндукции

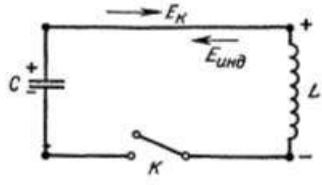


Рис. III.82

$$\mathscr{E}_{\text{самоинд}} = -L\frac{dI}{dt} \qquad (4.1)$$

и имеющую знак, противоположный знаку разности потенциалов на обкладках конденсатора. Это означает, что внутри проводников возбуждается напряженность электрического поля $E_{инд}$, которая направлена против напряженности $E_к$, создаваемой зарядами q конденсатора. Вследствие этого скорость возрастания тока в проводниках будет постепенно ослабляться, т. е. dI/dt (имеющая положительный знак) уменьшается до тех пор, пока конденсатор не разрядится совсем. В этот момент сила тока в цепи достигает наибольшего значения. В дальнейшем движущиеся заряды (электроны) начнут скапливаться на другой обкладке конденсатора, что вызовет появление в проводниках напряженности $E_к$, направленной против движения зарядов и уменьшающей силу тока в цепи. Ослабление тока приведет к появлению э. д. с. самоиндукции и индуцированного поля $E_{инд}$, направленного против $E_к$ и поддерживающего ток в цепи. Отрицательное dI/dt существует до тех пор, пока сила тока в цепи не станет равной нулю; тогда снова начнется разрядка конденсатора через катушку, и процесс повторится с обратным направлением тока.

Электрический ток (энергия этого тока) распространяется по проводникам с некоторой конечной скоростью. Поэтому движение зарядов, которое началось у обкладок конденсатора в какой-нибудь

момент времени, дойдет до катушки, спустя время Δt, зависящее от скорости распространения тока и длины проводников. Это означает, что сила тока в различных местах цепи должна быть различной. Изменения напряжения на обкладках конденсатора воспринимаются катушкой не в тот же момент, а с некоторым запозданием; точно так же э. д. с. самоиндукции, возбужденная в катушке в данный момент, окажет воздействие на конденсатор через некоторый промежуток времени. Однако, если размеры контура (длина соединительных проводов и длина обмотки катушки) малы, а быстрота изменения силы тока невелика, то можно допустить, что сила тока в данный момент времени одинакова в любом месте контура.

При этих условиях можно использовать второе правило Кирхгофа для замкнутого контура: алгебраическая сумма всех э. д. с. и падений напряжения в любой момент времени равна нулю. Это правило является точным только для постоянных токов; если же сила тока в контуре изменяется, то применение этого правила даст тем большую ошибку, чем быстрее изменяется ток и чем больше размеры контура.

Так как э. д. с. самоиндукции имеет знак, противоположный знаку разности потенциалов на пластинках конденсатора, то правило Кирхгофа запишется в виде

$$U - \mathscr{E}_{\text{инд}} + IR = 0; \quad \frac{q}{C} + L\frac{dI}{dt} + IR = 0. \tag{4.2}$$

Допустим сначала, что омическое сопротивление проводников, из которых состоит контур, исчезающе мало, т. е. $R \approx 0$, и выделением тепла можно пренебрегать. Кроме того, предположим, что в контуре не происходит вообще никаких потерь электрической энергии, например: нет потерь энергии в диэлектрике, находящемся между обкладками конденсатора, при его периодической поляризации; нет потерь в сердечнике катушки при его перемагничивании и, наконец, нет потерь на излучение электромагнитной энергии в окружающее пространство. При условии, что IR значительно меньше q/C и $L\,dI/dt$, формула (4.2) может быть переписана в виде ($I = dq/dt$; $dI/dt = d^2q/dt^2$):

$$\frac{d^2q}{dt^2} = -\frac{1}{LC}q. \tag{4.3}$$

Это уравнение при постоянных L и C аналогично связи между ускорением колеблющегося тела и смещением x от положения равновесия при гармоническом колебательном движении:

$$\frac{d^2x}{dt^2} = -\omega^2 x.$$

Решая дифференциальное уравнение (4.3), получим следующий закон изменения зарядов на обкладках конденсатора:

$$q = q_0 \sin \omega t, \tag{4.4}$$

причем угловая частота и период колебаний равны:

$$\omega = 1/\sqrt{LC}; \quad T = 2\pi/\omega = 2\pi\sqrt{LC}. \tag{4.5}$$

Сила тока в контуре, разность потенциалов U и напряженность электрического поля E между обкладками конденсатора, напряженность магнитного поля H внутри катушки также будут изменяться со временем по гармоническому закону:

$$I = \frac{dq}{dt} = q_0\omega \cos \omega t = I_0 \cos \omega t;$$

$$U = \frac{q}{C} = \frac{q_0}{C} \sin \omega t = U_0 \sin \omega t;$$

$$E = \frac{q}{\varepsilon_0 \varepsilon S} = E_0 \sin \omega t; \quad H = \frac{In}{l} = H_0 \cos \omega t.$$

Таким образом, при отсутствии омического сопротивления и каких-либо потерь энергии, электромагнитные колебания в контуре будут гармоническими с периодом, зависящим только от параметров контура L и C. Тогда энергия колебаний контура, равная сумме энергии электрического поля конденсатора и магнитного поля катушки

$$W = \frac{CU^2}{2} + \frac{LI^2}{2}, \tag{4.6}$$

с течением времени не будет изменяться. Колебания в контуре, как и в механических системах, означают периодическое превращение одного вида энергии в другой и обратно.

Если омическим сопротивлением контура R пренебречь нельзя, то следует решить уравнение (4.2), которое перепишем в виде

$$\frac{d^2q}{dt^2} + \frac{R}{L}\frac{dq}{dt} + \frac{1}{L\mathcal{C}} q = 0. \tag{4.7}$$

Чем меньше сопротивление R, тем меньше теряется энергия колебаний за период и тем ближе колебания к гармоническим, тем меньше отличие уравнения (4.7) от уравнения (4.3). Наоборот, если R очень велико, то возможно, что за время, даже меньшее одного периода, почти вся энергия W превратится в теплоту и колебаний вообще не будет.

Заметим, что в уравнении (4.7) вместе с потерей энергии на омическом сопротивлении, следовало бы учесть также и излучение контуром электромагнитной энергии в окружающее пространство; однако, для контуров, у которых электрическое и магнитное поля локализованы в очень малом объеме, это излучение чрезвычайно слабое. Кроме того, должны быть учтены потери в диэлектрике конденсатора и в сердечнике катушки. Предположим, что они или отсутствуют, или чрезвычайно слабые.

Допустим, параметры контура R, L и C остаются во время колебаний постоянными (R, например, могло бы изменяться от нагревания проводников при прохождении тока). Тогда период колебания, коэффициент и декремент затухания (см. ч. 1, § 18) также остаются постоянными; решая уравнение (4.7), можно показать, что заряд на обкладках конденсатора изменяется по закону

$$q = q_0 e^{-\delta t} \sin \omega t. \tag{4.8}$$

Из уравнений (4.7) и (4.8) можно получить следующие значения для коэффициента затухания δ, угловой частоты и периоды колебаний:

$$\delta = \frac{R}{2L}; \quad \omega = \sqrt{\frac{1}{LC} - \left(\frac{R}{2L}\right)^2}; \quad T = \frac{2\pi}{\sqrt{\frac{1}{LC} - \left(\frac{R}{2L}\right)^2}}. \qquad (4.9)$$

Таким образом, наличие омического сопротивления увеличивает период колебаний в контуре. При $1/LC = R^2/4L^2$, $R^2 = 4L/C$ период колебаний делается бесконечно большим; при $R^2 > 4L/C$ колебания отсутствуют; происходит так называемый *апериодический разряд*, т. е. постепенная (без колебаний) разрядка конденсатора через катушку. При $R \to 0$ период колебаний приближается к значению (4.5).

В рассмотренном выше колебательном контуре предполагалось, что вся емкость C системы сосредоточена в конденсаторе, а вся индуктивность L — в катушке; емкостью и индуктивностью соединительных проводов, а также емкостью обмотки катушки при этом пренебрегают; при этих условиях частота собственных колебаний определяется формулой $\omega^2 = 1/LC$ тем точнее, чем меньше емкость за пределами конденсатора и индуктивность за пределами катушки.

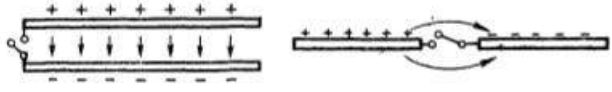


Рис. III.83

Для получения больших частот собственных колебаний необходимо взять контур с малыми значениями L и C; тогда емкость и индуктивность соединительных проводов может иметь заметное значение. Для очень больших частот надобность в конденсаторах и катушках отпадает, так как емкость и индуктивность линейных проводников оказываются достаточными. Например, системы из двух прямолинейных проводников, изображенные на рис. III.83, являются электрическими колебательными контурами с большой (в зависимости от длины проводников) частотой собственных колебаний. Если зарядить проводники разноименным электричеством и затем замкнуть ключ, то в системе начнется движение зарядов, т. е. появится электрический ток и связанное с ним магнитное поле. Так как проводники обладают индуктивностью, то в них возбуждается также и э. д. с. индукции. В контуре происходят электромагнитные колебания, т. е. периодические превращения энергии электрического поля между проводниками в энергию магнитного поля вокруг проводников и обратно.

Заметим, что такие открытые контуры, у которых электрическое и магнитное поля занимают большой объем пространства, излучают гораздо больше энергии, чем закрытые контуры, у которых поля локализованы в очень малых объемах внутри конденсаторов и катушек.

Если на колебательный контур подействовать извне периодической э. д. с., то в контуре можно вызвать *вынужденные* электромагнитные колебания. Внешнюю периодическую э. д. с. можно приложить различными способами, например, путем индуктивной электромагнитной связи контура с источником э. д. с. или путем непосредственного присоединения источника тока к контуру (рис. III.84, *а*, *б*, *в*).

Рассмотрим колебания в контуре, изображенном на рис. III.84, *б*. Допустим, что внешняя э. д. с. изменяется со временем по синусоидальному закону

$$\mathscr{E} = \mathscr{E}_0 \sin \omega t. \tag{4.10}$$

Присоединенный к контуру внешний источник тока совершает положительную работу и, следовательно, увеличивает энергию контура только в том случае, когда ток в контуре течет в направлении

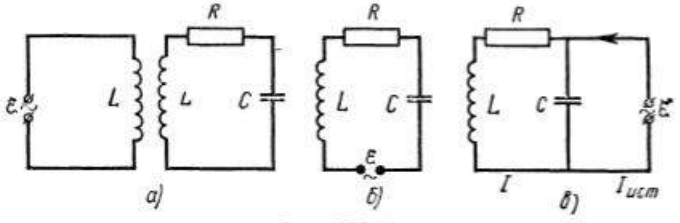


Рис. III.84

электрического поля E, вызванного этим источником тока. Наоборот, внешняя э. д. с. производит отрицательную работу и уменьшает энергию контура, если ток течет в направлении, противоположном E. Если при наличии внешней э. д. с. в контуре с сопротивлением установились незатухающие колебания, то это означает, что результирующая работа внешнего источника за один период колебаний является положительной и в точности равна потерям энергии в контуре за тот же промежуток времени (причем подкачка энергии извне производится также непрерывно, как она и расходуется на различные потери).

Найдем амплитуду, частоту и фазу силы тока при вынужденных колебаниях в контуре с заданными неизменными значениями параметров L, C и R под действием внешней э. д. с. (см. формулу (4.10)). Если частота колебаний небольшая, а длина проводников контура невелика, то можно полагать, что сила тока во всех сечениях контура в каждый момент времени одна и та же; следовательно, можно применить второе правило Кирхгофа для замкнутых контуров; сумма падений напряжения в контуре $IR + U$ равна сумме э. д. с., действующих в этом контуре:

$$IR + U = \mathscr{E} + \mathscr{E}_{\text{инд}}; \quad \frac{q}{C} + IR + L\frac{\mathrm{d}I}{\mathrm{d}t} = \mathscr{E}_0 \sin \omega t. \tag{4.11}$$

Первое уравнение должно соблюдаться в любой момент времени; поэтому, если внешняя э. д. с. изменяется по какому-нибудь закону,

то вслед за ней, для сохранения этого соотношения, вынуждены одновременно изменяться I, U и $\mathscr{E}_{\text{инд}}$. Это означает, что изменения I, U и $\mathscr{E}_{\text{инд}}$ со временем должны происходить с такой же частотой ω, с какой изменяется внешняя э. д. с.; однако фаза колебаний этих величин может отличаться от фазы $\mathscr{E}$.

Продифференцируем уравнение (4.11) по времени; получим дифференциальное уравнение для силы тока:

$$\frac{d^2I}{dt^2} + \frac{R}{L}\frac{dI}{dt} + \frac{1}{LC}I = \frac{\mathscr{E}_0}{L}\omega\cos\omega t. \qquad (4.12)$$

Для установившегося режима колебаний решение этого уравнения имеет вид

$$I = I_0 \sin(\omega t - \varphi). \qquad (4.13)$$

Расчет приводит к следующим значениям для амплитуды тока и разности фаз между током и внешней э. д. с.:

$$I_0 = \frac{\mathscr{E}_0}{\sqrt{R^2 + \left(L\omega - \frac{1}{C\omega}\right)^2}}; \quad \operatorname{tg}\varphi = \frac{L\omega - (1/C\omega)}{R}. \qquad (4.14)$$

Таким образом, амплитуда тока в контуре зависит от R и от соотношения между L, C и ω. При постоянном R можно получить максимальную амплитуду тока, если

$$L\omega = \frac{1}{C\omega} \quad \text{или} \quad \omega = \frac{1}{\sqrt{LC}}. \quad \text{Тогда } I_0 = \frac{\mathscr{E}_0}{R}; \quad \operatorname{tg}\varphi = 0; \quad \varphi = 0. \qquad (4.15)$$

В данном случае сила тока изменяется в одной фазе ($\varphi = 0$) с внешней э. д. с., вследствие чего работа этой э. д. с. в течение всего периода колебаний положительна (при $\varphi \neq 0$ внешняя э. д. с. в течение части периода совершает отрицательную работу, когда $\mathscr{E}$ и I имеют различные знаки). Условие $\omega = 1/\sqrt{LC}$ означает, что частота внешней силы равна частоте собственных колебаний (см. формулу (4.5)). При этом условии энергия, поступающая от внешнего источника к контуру за один период колебания, достигает максимума, т. е. *источник тока развивает наибольшую полезную мощность*. Накопление энергии в контуре, а следовательно, увеличение силы тока происходит до тех пор, пока сумма потерь за период не станет равной притоку энергии от внешнего источника за то же время. Очевидно, при малых R для этого потребуются большие токи (большие значения амплитуды тока I_0, при малых сопротивлениях R получаются из соотношения $I_0 = \mathscr{E}_0/R$. Следует учесть, что большие токи могут вызвать сильное нагревание проводников, изменение их сопротивления и даже расплавление).

Равенство частоты внешней э. д. с. и частоты собственных колебаний контура есть условие *электрического резонанса*. При этом не только амплитуда силы тока, но и амплитудные значения всех электрических и магнитных величин, изменяющихся при колебаниях, достигают наибольших значений. Отметим, что при резонансе величина $L\omega$ равна $1/C\omega$, и амплитуда падения напряжения на конденсаторе

равна амплитуде напряжения на катушке:

$$U_0 = \frac{q_0}{C} = \frac{I_0}{C\omega} = \frac{\mathscr{E}_0}{R}\sqrt{\frac{L}{C}}; \quad \mathscr{E}_{\text{инд}} = L\omega I_0 = \frac{\mathscr{E}_0}{R}\sqrt{\frac{L}{C}}. \qquad (4.16)$$

Чем сильнее отличается частота внешней э. д. с. от частоты собственных колебаний контура, тем, согласно формуле (4.14), меньше амплитудные значения силы тока и связанных с нею величин.

Аналогичные рассуждения можно провести и для других способов воздействия внешней периодической э. д. с. на колебательный контур (рис. III.84, *а*, *в*). Как в схеме с индуктивной связью (*а*), так и при параллельном соединении внешнего источника к емкости и индуктивности (*в*), в контуре можно получить незатухающие колебания, так как в обоих случаях удается осуществить непрерывную компенсацию потерь энергии в контуре. Для этого необходимо, чтобы внешний источник тока совершал положительную работу. В обеих схемах можно доказать, что если частота ω внешней э. д. с. сильно отличается от частоты собственных колебаний контура, т. е. от $1/\sqrt{LC}$, то у внешнего источника тока между э. д. с. и силой тока существует разность фаз, вследствие чего за одну часть периода колебаний она совершает положительную, а за остальную часть — отрицательную работу. При резонансе ток через внешний источник находится в фазе с его э. д. с., и тогда в течение всего периода совершается только положительная работа. Амплитуда силы тока через внешний источник, а также амплитуда силы тока в контуре достигают при этом максимальных значений.

Для схемы (*в*) при малом сопротивлении R амплитуда тока в контуре I_0 оказывается при резонансе значительно больше амплитуды тока $I_{\text{ист}}$, протекающего через внешний источник. В контуре происходит перекачка большого количества энергии от конденсатора к катушке и обратно, тогда как от внешнего источника в контур поступает лишь небольшая мощность, необходимая только для компенсации потерь.

Возбуждение колебаний в электрическом контуре возможно и без участия внешней переменной электродвижущей силы; подкачку энергии в контур можно произвести при помощи периодического м е х а н и ч е с к о г о воздействия на параметры контура — емкость и индуктивность («параметрическое возбуждение»). Допустим, что в начальный момент времени, когда ток в контуре равен нулю, на обкладках конденсатора имеются очень малые («затравочные») заряды $\pm q_0$; соответствующая им энергия будет равна $q_0^2/2C_0$. Произведем достаточно быстрое уменьшение емкости конденсатора (например, раздвигая его обкладки или удаляя диэлектрик). При этом придется совершить некоторую работу, которая пойдет на увеличение энергии электрического поля внутри конденсатора. Разность потенциалов на обкладках конденсатора $U = q_0/C$ увеличивается ($C < C_0$), и поэтому максимальный ток, вызванный в контуре при последующей разрядке конденсатора, будет несколько больше, чем это было бы без внешнего воздействия. В момент, когда конденсатор разрядится, можно произвести увеличение емкости (сближение обкладок или введение диэлектрика) без затраты энергии. Когда закончится один период колебания,

заряды на обкладках конденсатора окажутся больше q_0 (ввиду того что были большими токи через индуктивность). В этот момент можно повторить внешнее воздействие — затратить механическую работу и уменьшить емкость конденсатора; разность потенциалов между обкладками конденсатора опять увеличится, и в течение второго периода колебаний ток через индуктивность будет еще больше. Следовательно, периодически повторяя процесс уменьшения емкости конденсатора в его заряженном состоянии и увеличения — в разряженном, можно получить в контуре электрические колебания с очень большими значениями зарядов и токов. Необходимым условием для этого является равенство между частотой внешнего механического воздействия и частотой собственных колебаний контура.

Аналогичное параметрическое возбуждение колебаний в электрическом контуре может быть получено, если механическое воздействие оказывать не на конденсатор, а на индуктивность контура. Можно, например, вкладывать железный сердечник в катушку, когда ток через нее равен нулю, и вытаскивать из нее, когда ток равен максимуму; последняя операция потребует затраты механической работы, которая пойдет на увеличение энергии магнитного поля в катушке $LI^2/2$ (L увеличивается при почти постоянном I). При резонансе между частотой таких воздействий и частотой собственных колебаний контура можно довести амплитуды токов и зарядов до больших значений. Заметим, что в обоих процессах параметрического возбуждения затрачивается внешняя механическая энергия и увеличивается электромагнитная энергия контура, т. е. электрический контур работает как *преобразователь механической энергии в электромагнитную.*

§ 29. ЭЛЕКТРОМАГНИТНЫЕ ВОЛНЫ

РАСПРОСТРАНЕНИЕ ВОЛНЫ ПО ПРОВОДНИКАМ

Рассмотрим систему из двух параллельных проводов (рис. III.85), в которой внешний источник тока вызывает вынужденные электромагнитные колебания. Э. д. с., возбужденная на участке AB внешним источником энергии, приведет в движение электроны; допустим, что в начальный момент времени электроны стали перемещаться от B к A. Это движение распространяется по верхнему проводу AC с некоторой конечной скоростью, пока не дойдет до конца провода C, где временно происходит скопление электронов. По мере приближения к точке C скорость электронов будет уменьшаться и поэтому силы токов I_1 и I_2 на расстояниях x и $x + \Delta x$ от точки A будут различны.

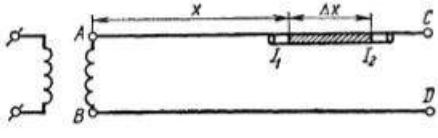


Рис. III.85

Одновременно придут в упорядоченное движение и электроны нижнего провода, подхваченные создавшимся внутри этого провода электрическим полем, направленным к положительно зарядившейся точке B. Таким образом, верхний провод заряжается отрицательным,

а нижний — положительным электричеством и между ними появляется электрическое поле; провода AC и BD играют роль обкладок линейного конденсатора. Скопление электронов в верхнем проводе и «оголение» положительных зарядов в нижнем проводе препятствуют дальнейшему движению электронов от B к A.

Пока в проводах происходит движение зарядов, вокруг них существует магнитное поле (возрастающее или убывающее). Это переменное магнитное поле возбуждает в самих проводах э. д. с. самоиндукции, которая, по закону Ленца, препятствует как возрастанию, так и убыванию тока в проводах.

Таким образом, благодаря совместному действию поля зарядов и э. д. с. самоиндукции движение электронов от нижнего провода к верхнему со временем прекратится и затем начнется обратное движение, когда внешняя э. д. с. изменит свой знак.

Если частота изменения внешней э. д. с. велика, то возможно, что движение электронов, вызванное этой э. д. с. за первую половину периода, не успеет еще распространиться до конца проводов, как уже вследствие перемены знака внешней э. д. с. на участке $A — B$ начнется движение электронов в противоположном направлении. Тогда точка A и прилегающие к ней участки верхнего провода будут заряжаться положительно, в то время как более удаленные участки пока еще остаются заряженными отрицательно; в нижнем проводе вследствие обратного движения электронов точка B и ближайшие к ней участки заряжаются отрицательно, тогда как в остальной части пока еще существует избыток положительных зарядов. При очень большой длине проводов и большой частоте внешней э. д. с. вдоль проводов образуется некоторое *распределение* положительных и отрицательных зарядов; это распределение меняется со временем по такому же закону, по которому изменяется внешняя э. д. с. Если внешняя э. д. с. — синусоидальная, то вдоль проводов будет существовать синусоидальное распределение зарядов, токов и потенциалов (рис. III.86).

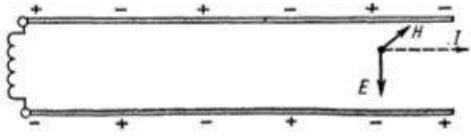


Рис. III.86

Итак, между проводами появляются электрические и магнитные поля, меняющиеся со временем. Переменное электрическое поле возбуждает магнитное поле, и, обратно, переменное магнитное поле создает электрическое поле, поэтому окружающее нашу систему поле является электромагнитным полем. В каждой точке этого поля будут какие-то напряженности электрического **E** и магнитного **H** полей, связанные между собой и меняющиеся со временем. Существенно, что изменение напряженности электрического поля в какой-нибудь определенной точке пространства вызывает появление магнитного поля не только в этой точке, но и во всех поблизости расположенных точках. Возбужденные в этих точках переменные **E** и **H** вызывают в свою очередь электрические и магнитные напряженности в других, более удаленных от нашей системы точках и т. д. Следовательно, электро-

магнитное поле не может быть локализовано в определенном объеме пространства; оно распространяется от одной точки пространства к другим с некоторой скоростью.

Распространяющееся электромагнитное поле, в котором напряженности электрического и магнитного полей изменяются по какому-нибудь периодическому закону, называются *электромагнитной волной*. Очевидно, источником электромагнитной волны может быть любой электрический колебательный контур или даже любой проводник, по которому течет переменный ток. «Излучающая способность» источника электромагнитной волны зависит от его формы и размеров, а также от частоты колебаний (прямо пропорциональна ω^4, поэтому провода с током малой технической частоты излучают ничтожную энергию).

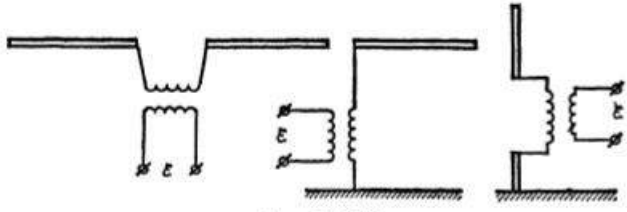


Рис. III.87

Для увеличения излучающей способности данного контура необходимо увеличить тот объем пространства, в котором этот контур создает электромагнитное поле. Например, обычный колебательный контур с сосредоточенными параметрами создает электромагнитное поле только в небольшом объеме внутри катушки индуктивности и между обкладками конденсатора и поэтому почти не излучает. Двухпроводная система с распределенными параметрами является более «открытым» контуром и поэтому обладает большей излучающей способностью; ее можно увеличить, если раздвинуть провода и получить так называемый линейный вибратор, изображенный на рис. III.87; там же показаны горизонтальная и вертикальная антенны (вторым «проводом» является поверхность Земли).

Источник энергии должен возбудить в излучающей антенне стоячие волны, и, чтобы амплитуда этих волн была максимальной, необходимо условие резонанса — равенство частоты внешнего воздействия и частоты собственных колебаний излучателя.

УРАВНЕНИЕ ЭЛЕКТРОМАГНИТНОЙ ВОЛНЫ

Выведем уравнение плоской электромагнитной волны, воспользовавшись уравнениями Максвелла (см. формулу (3.80)):

$$\oint \mathbf{H}d\mathbf{l} = \frac{d\mathbf{D}}{dt}\mathbf{S}; \quad \oint \mathbf{E}d\mathbf{l} = -\frac{d\mathbf{B}}{dt}\mathbf{S}. \qquad (4.17)$$

Напомним, что при расчете контурных интегралов направление обхода и направления векторов **D** и **B** должны удовлетворять правилу правого винта.

Для упрощения рассуждений допустим, что в том месте, где возбуждается электромагнитное поле, вектор **E** все время остается параллельным координатной оси OZ, а вектор **H** — оси OY. Так как уравнения (4.17) применимы для контуров любой формы, размеров и ориентировки в поле, то для использования первого из них выберем элементарный контур $Oabc$, лежащий в плоскости XOY, а для второго — контур $Oaef$, лежащий в плоскости XOZ (рис. III.88). Векторы **E** и **H** являются функциями координат, поэтому их значения в различных местах указанных контуров будут различными. Например, если в точке O вектор **E** имеет данное значение, то в точке a с координатой dx его значение будет равно $E + dx\,(\partial E/\partial x)$, где частная производная

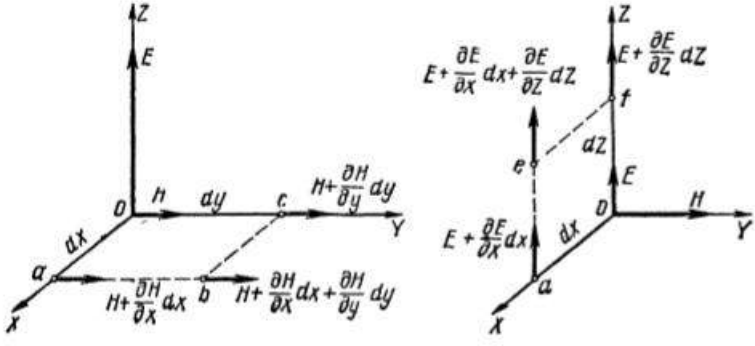


Рис. III.88

$\partial E/\partial x$ показывает быстроту изменения E в направлении оси OX. На рис. III.88 указаны значения E и H на концах элементарных участков.

Используем первое уравнение (4.17) для контура $Oabc$. На участках Oa и bc произведение **Hdl** будет равно нулю, так как вектор **H** перпендикулярен Oa и bc. Для участков ab и cO умножим длину каждого из этих участков dy на средние значения вектора **H** в пределах этих участков; так как на участке cO вектор **H** направлен против обхода, то получим:

$$dy \cdot \frac{1}{2}\left[\left(H + \frac{\partial H}{\partial x}\,dx\right) + \left(H + \frac{\partial H}{\partial x}\,dx + \frac{\partial H}{\partial y}\,dy\right)\right] - \\ - dy \cdot \frac{1}{2}\left[\left(H + \frac{\partial H}{\partial y}\,dy\right) + H\right] = \frac{\partial D}{\partial t}\,dx\,dy,$$

где $dxdy$ — площадь, охватываемая контуром обхода. Здесь вместо dD/dt написано $\partial D/\partial t$, так как для расчета потока вектора **D** через площадку $dxdy$ можно взять значение вектора **D** в одной точке, например в центре этой площадки. После сокращений получаем

$$\frac{\partial H}{\partial x} = \frac{\partial D}{\partial t} = \varepsilon_0 \varepsilon \frac{\partial E}{\partial t}. \qquad (4.18)$$

Аналогичный расчет для второго уравнения (4.17) и контура *Oaef* дает следующий результат:

$$\frac{\partial E}{\partial x} = \frac{\partial B}{\partial t} = \mu_0 \mu \frac{\partial H}{\partial t}. \tag{4.19}$$

Вычислим частные производные по времени от выражения (4.18) и по координате от выражения (4.19), полагая ε и μ постоянными величинами:

$$\frac{\partial^2 H}{\partial x \partial t} = \varepsilon_0 \varepsilon \frac{\partial^2 E}{\partial t^2}; \quad \frac{\partial^2 E}{\partial x^2} = \mu_0 \mu \frac{\partial^2 H}{\partial t \partial x}.$$

Из этих уравнений следует

$$\frac{\partial^2 E}{\partial t^2} = \frac{1}{\varepsilon_0 \varepsilon \mu_0 \mu} \frac{\partial^2 E}{\partial x^2}. \tag{4.20}$$

Мы получили *дифференциальное уравнение волны*. Этому уравнению удовлетворяет, в частности, простейшая, синусоидальная, волна

$$E = E_0 \sin \omega \left(t - \frac{x}{c} \right), \tag{4.21}$$

в которой вектор **E** распространяется вдоль оси OX со скоростью

$$c = \frac{1}{\sqrt{\varepsilon_0 \varepsilon \mu_0 \mu}} = \frac{3 \cdot 10^8}{\sqrt{\varepsilon \mu}} \text{ м/с}. \tag{4.22}$$

Аналогичными рассуждениями можно получить уравнение волны и для вектора **H**; соотношение между значениями E и H в волне имеет вид

$$\sqrt{\varepsilon_0 \varepsilon}\, E = \sqrt{\mu_0 \mu}\, H. \tag{4.23}$$

Следовательно, в каждой точке распространяющейся волны векторы **E** и **H** пропорциональны друг другу. Кроме того, они перпендикулярны направлению распространения (оси OX), ввиду чего электромагнитную волну называют поперечной волной. Исходя из уравнений (4.17), можно также доказать, что векторы **E** и **H** в волне перпендикулярны друг другу.

ЭНЕРГИЯ ЭЛЕКТРОМАГНИТНОЙ ВОЛНЫ

Электромагнитная волна в направлении своего распространения переносит некоторое количество энергии; она содержится в электрическом и магнитном полях и пропорциональна квадратам напряженностей E и H этих полей. Через 1 м^2 площадки, ориентированной перпендикулярно направлению распространения волны, за единицу времени пройдет вся та энергия, которая содержится в объеме параллелепипеда с основанием 1 м^2 и высотой, равной скорости распространения c; эта энергия является важной характеристикой электромагнитного излучения (*плотностью потока энергии*) и обозначается через I. Так как в единице объема электрического и магнитного полей содержится энергия

$$w = \frac{\varepsilon_0 \varepsilon E^2}{2} + \frac{\mu_0 \mu H^2}{2}, \tag{4.24}$$

то, имея в виду выражения (4.22), получим

$$I = wc = \frac{1}{2}\sqrt{\frac{\varepsilon_0\varepsilon}{\mu_0\mu}}E^2 + \frac{1}{2}\sqrt{\frac{\mu_0\mu}{\varepsilon_0\varepsilon}}H^2. \qquad (4.25)$$

В векторной записи с учетом равенства (4.23) выражение для плотности потока энергии в волне имеет простой вид

$$\mathbf{S} = [\mathbf{EH}]. \qquad (4.26)$$

Вектор **S**, ориентированный по направлению распространения волны, называется *вектором Пойнтинга* (см. также рис. III.86).

Электромагнитные волны имеют весьма широкий диапазон частот $\nu = \omega/2\pi$. Обычно их различают по длинам волн $\lambda = c/\nu = 2\pi c/\omega$. В радиотехнике используются электромагнитные волны длиной от нескольких километров до нескольких сантиметров. В лабораторных условиях получено излучение, длина волны которого измеряется миллиметрами. Более короткие электромагнитные волны, лежащие в пределах от 1 до 10^{-9} см, называются *оптическим излучением*. Видимое излучение, вызывающее у человеческого глаза световое ощущение различных цветов, является частью оптического излучения и имеет длины волн от 0,77 до 0,38 мкм. Оптическое излучение $\lambda > 0{,}77$ мкм называется *инфракрасным*, а $\lambda < 0{,}38$ мкм — *ультрафиолетовым* излучением. Рентгеновские лучи и γ-излучение радиоактивных веществ имеют более короткие длины волны, измеряемые несколькими нанометрами и менее.

ИЗЛУЧЕНИЕ ВОЛНЫ ЭЛЕКТРОНОМ

Излучение электромагнитных волн проводниками, по которым течет переменный ток, в конечном счете означает излучение этих волн электронами, совершающими колебательное движение внутри проводников. Свободный электрон, колеблющийся в вакууме, также будет излучать электромагнитные волны. Так как это излучение происходит непрерывно, то оно должно существовать и в течение любой части периода колебаний. При этом важно, что колебательное движение есть движение с ускорением, так как равномерно движущийся электрон не должен излучать ввиду постоянства его энергии.

В электродинамике выведена формула, позволяющая рассчитать мощность излучения электрона, движущегося с ускорением a:

$$P = \frac{e^2a^2}{6\pi\varepsilon_0c^3} = \frac{2}{3}k\frac{e^2}{c^3}a^2,$$

где e — заряд электрона, c — скорость света, $\varepsilon_0 = 1/4\pi k$ — электрическая постоянная. Например, если электрон совершает колебательное движение с радиотехнической частотой $\omega = 2\pi\nu = 10^8$ рад/с и имеет амплитуду $x_0 = 0{,}01$ м, то максимальная мощность, соответствующая амплитуде ускорения $a_0 = \omega^2x_0 = 10^{14}$ м/с^2, будет равна

$$P = \frac{2\cdot 9}{3}\cdot 10^9\,\frac{(1{,}6\cdot 10^{-19})^2}{(3\cdot 10^8)^3}\cdot 10^{28} = 0{,}53\cdot 10^{-23}\ \text{Дж/с}.$$

Для того чтобы получить среднюю мощность излучения около одного ватта, такие колебания должны совершать 10^{23} электронов.

В атомах вещества электрон совершает орбитальное движение, соответствующее двум колебаниям во взаимно перпендикулярных направлениях; этим колебаниям должно соответствовать некоторое излучение. Таким образом, излучение должно существовать не только при наличии тангенциальных ускорений, но и при равномерном движении по окружности, когда имеется только центростремительное ускорение. Например, мощность излучения электрона, вращающегося в атоме водорода по орбите радиуса $r = 0{,}53 \cdot 10^{-10}$ м со скоростью $v = 2{,}18 \cdot 10^6$ м/с и ускорением $a = v^2/r \approx 9 \cdot 10^{22}$ м/с^2, оказывается равной $5{,}27 \cdot 10^{-9}$ Дж/с $= 3{,}3 \cdot 10^{10}$ эВ/с. На этой орбите электрон имеет энергию порядка 13 эВ, поэтому при такой мощности излучения она будет израсходована за $4 \cdot 10^{-10}$ с. Таким образом, исходя из классических представлений продолжительное вращение электрона по орбите оказывается невозможным. Этот вывод имел большое значение в развитии современных представлений об атоме.

ЧЕРЕНКОВСКОЕ ИЗЛУЧЕНИЕ

Однако при определенных условиях излучение электрона будет зависеть не от ускорения, а от скорости движения, т. е. возможно и при р а в н о м е р н о м движении. Такое излучение было открыто С. И. Вавиловым и П. А. Черенковым (1934) и появляется, когда электрон движется в среде со скоростью v, превышающей скорость c распространения света в этой среде. На рис. III.89 показана схема образования фронта волны, испускаемой электроном. Перемещение электрона по траектории OX сопровождается изменением электрического поля в окружающем пространстве, что вызывает появление индуцированного магнитного поля; возникшее электромагнитное возмущение распространяется в виде волны. Для нахождения фронта этой волны допустим, что каждая точка A на прямой OX в момент t пребывания электрона в этой точке становится центром сферической волны (см. ч. I, § 31). За время Δt перемещения электрона на расстояние $AA_1 = v\Delta t$ волна a, возникшая в точке A, будет иметь вид сферы радиуса $r = c\Delta t$. Такие поверхности (b, c, ...) можно построить и для точек B, C, ... Согласно принципу Гюйгенса, для нахождения фронта волны в момент $t + \Delta t$ необходимо построить огибающую этих волн; ею будет поверхность конуса с вершиной в точке A_1, причем угол θ (рис. III.89) определится из соотношения

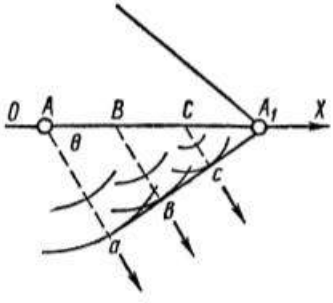


Рис. III.89

$$\cos\theta = Aa/AA_1 = c/v.$$

Излучение электрона оказывается направленным вперед по движению и имеющим максимум интенсивности на поверхности этого ко-

нуса. В этом отношении оно напоминает «носовую волну» от тела (корабля, самолета), движущегося в воде или в воздухе со скоростью, большей скорости распространения механической волны в этой среде, т. е. со сверхзвуковой скоростью.

При большой скорости электрона излучение Вавилова—Черенкова может оказаться в оптическом диапазоне в области видимого света. Заметим, что сверхсветовая скорость движения частиц в среде ($v > c$) всегда м е н ь ш е скорости света в вакууме c_0, которая, согласно теории относительности, является предельно возможной скоростью перемещения тел, имеющих инертную массу. Отношение c_0/c есть показатель преломления данной среды по отношению к вакууму, и оно всегда больше единицы, т. е. излучение происходит при $v > c_0/n$, но при условии $v < c_0$.

Излучение Вавилова—Черенкова получило применение в измерительной технике; регистрируя импульсы излучения от каждой пролетающей частицы и измерив угол θ, можно сосчитать число частиц и определить их скорость.

Часть IV
ОПТИКА И ФИЗИКА АТОМА

Глава 1
ОСНОВЫ ВОЛНОВОЙ ОПТИКИ

§ 1. ВОЛНОВАЯ (ЭЛЕКТРОМАГНИТНАЯ) И КОРПУСКУЛЯРНАЯ (ФОТОННАЯ) ТЕОРИИ СВЕТА

Для объяснения световых явлений в физике используются две теории света — волновая и корпускулярная: 1) по волновой *(электромагнитной)* теории, световое излучение представляет собой *электромагнитные волны*, длина которых лежит в пределах от 0,38 до 0,77 мкм; более короткие и более длинные волны не вызывают в зрительном органе человека светового ощущения и воспринимаются специальной аппаратурой; 2) по корпускулярной *(фотонной)* теории, световое излучение представляет собой *поток особых частиц — фотонов*; фотоны обладают энергией, массой и импульсом.

При помощи волновой теории объясняются законы распространения света (отражение, преломление, интерференция, дифракция света и т. п.); при помощи фотонной теории объясняются законы взаимодействия между светом и веществом (поглощение и рассеяние света электронами, излучение и поглощение света атомами и т. п.). Ограничиться только одной из этих теорий для объяснения всех явлений излучения, распространения и поглощения света не удается. Таким образом, в световых явлениях наблюдается своеобразный дуализм волновых и корпускулярных свойств. Заметим, что частота колебаний ν, которая присуща световой волне по волновой теории, определяет энергию $\varepsilon_ф$, массу $m_ф$ и импульс $p_ф$ фотонов, из которых состоит этот свет по корпускулярной теории:

$$\varepsilon_ф = h\nu; \quad m_ф = h\nu/c^2; \quad p_ф = h\nu/c, \tag{1.1}$$

где $h = 6{,}626 \cdot 10^{-27}$ эрг·с $= 6{,}626 \cdot 10^{-34}$ Дж·с есть постоянная Планка; c — скорость света.

Электромагнитная волна (см. ч. III, § 29) есть распространение взаимосвязанных электрических и магнитных полей; в каждой точке волны происходит периодическое изменение напряженностей **E** и **H** электрического и магнитного полей. Оба эти вектора взаимно перпендикулярны друг другу и, кроме того, перпендикулярны направлению распространения волны.

Допустим, что в точке A (рис. IV.1) вектор **E** изменяется со временем по формуле $E = E_0 \sin \omega t$; тогда в точке A_1, отстоящей от A на расстоянии x, его значение будет равно

$$E_1 = E_0 \sin \omega \left(t - \frac{x}{c} \right) = \\ = E_0 \sin 2\pi \left(\frac{t}{T} - \frac{x}{\lambda} \right), \quad (1.2)$$

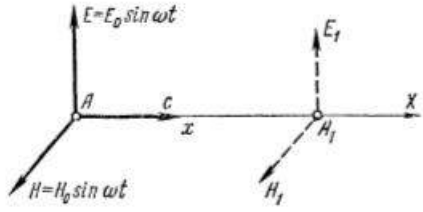


Рис. IV.1

где ω — угловая частота ($\omega = 2\pi\nu = 2\pi/T$); T — период; $\lambda = c/\nu = cT$ — длина волны; c — скорость распространения волны.

По такой же формуле рассчитывается напряженность магнитного поля H_1. Очевидно, формула (1.2) применима, если только вдоль направления распространения волны амплитудное значение напряженности поля E_0 остается постоянным.

СПЕКТРАЛЬНЫЙ СОСТАВ ИЗЛУЧЕНИЯ

Реальное излучение содержит в себе не одну определенную частоту колебаний, а некоторый набор различных частот, называемый *спектром* или спектральным составом данного излучения. Излучение называется *монохроматическим*, если оно содержит очень узкий интервал частот $\Delta\nu$ (или длин волн $\Delta\lambda$). В видимой области монохроматическое излучение вызывает световое ощущение определенного *цвета*; например, излучение, охватывающее интервал длин волн от 0,55 до 0,56 мкм, воспринимается как зеленый цвет. Чем уже интервал частот $\Delta\nu$ данного излучения, тем более монохроматическим оно является. Формула (1.2) относится к идеально монохроматическому излучению, содержащему одну частоту колебаний.

Раскаленные твердые и жидкие тела испускают *непрерывный* (или сплошной) спектр электромагнитных волн очень широкого интервала частот. Светящиеся разреженные газы испускают *линейчатый* спектр, состоящий из отдельных монохроматических излучений, называемых *спектральными линиями*; каждая спектральная линия характеризуется определенной частотой колебаний (или длиной волны), расположенной в середине охватываемого ею узкого интервала частот. Если источниками излучения являются не отдельные (изолированные, свободные) атомы, а молекулы газа, то спектр состоит из полос *(полосатый спектр)*; каждая полоса охватывает более широкий непрерывный интервал длин волн, чем спектральная линия.

Линейчатый (атомный) спектр каждого вещества является характерным для него; благодаря этому возможен *спектральный анализ*, т. е. определение химического состава вещества по длинам волн спектральных линий испускаемого им излучения.

Допустим, что электромагнитная волна распространяется вдоль некоторой прямой, которую будем называть *лучом*. Можно интересоваться изменением вектора **E** в определенной точке луча с течением

времени; возможно, что в этой точке изменяется не только величина вектора **E**, как это следует из формулы (1.2), но и ориентировка вектора **E** в пространстве. Далее можно зафиксировать величину и направление вектора **E** в различных точках луча, но в определенный момент времени. Если окажется, что в различных точках вдоль луча все векторы **E** лежат в одной плоскости, то излучение называется *плоскополяризованным* или *линейно-поляризованным;* такое излучение дает источник, который в процессе излучения сохраняет плоскость колебаний. Если же плоскость колебаний источника волны со временем изменяется, то вектор **E** в волне не лежит в определенной плоскости и излучение не будет плоскополяризованным. В частности, можно получить волну, в которой вектор **E** равномерно вращается вокруг луча. Если же вектор **E** изменяет свою ориентировку вокруг луча совершенно беспорядочно, то излучение называется *естественным*. Такое излучение получается от светящихся твердых, жидких и газообразных тел, у которых плоскости колебаний элементарных источников излучения — атомов и молекул — ориентированы в пространстве беспорядочно.

Таким образом, простейшим излучением является монохроматическая плоскополяризованная волна. Плоскость, в которой лежат вектор **E** и вектор направления распространения волны, называется *плоскостью колебаний волны*; плоскость, перпендикулярная плоскости колебаний (т. е. плоскость, в которой лежит вектор **H**), называется *плоскостью поляризации*.

Скорость распространения электромагнитных волн в вакууме есть одна из важнейших констант физики и равна

$$c = 2{,}9979 \cdot 10^8 \text{ м/с} \approx 3 \cdot 10^8 \text{ м/с}.$$

В других средах она меньше и определяется по формуле (см. ч. III, § 29)

$$c = \frac{1}{\sqrt{\varepsilon_0 \varepsilon \mu_0 \mu}} \approx \frac{3 \cdot 10^8}{\sqrt{\varepsilon\mu}} \text{ м/с},$$

где ε и μ — соответственно диэлектрическая и магнитная проницаемости среды.

При переходе излучения из одной среды в другую частота колебаний ν в волне сохраняется, но длина волны λ изменяется; обычно, если это не оговорено, λ обозначает длину волны в вакууме.

Выше указывалось, что видимое излучение (которое мы называем *светом*) охватывает длины волн от 400 до 770 нм (нм = 10^{-9}, м); при специальной тренировке глаз может воспринимать свет длиной волны от 320 до 900 нм. Более широкий интервал длин волн от 1 см до 10^{-9}, см, охватывающий также ультрафиолетовую и инфракрасную области, называется *оптическим излучением*.

ХАРАКТЕРИСТИКИ ИЗЛУЧЕНИЯ

Допустим, что источник непрерывно излучает электромагнитные волны определенного спектрального состава. Мысленно проведем в пространстве некоторую плоскую поверхность S, перпендикуляр-

ную лучам (тогда нормаль к каждой точке S будет совпадать с направлением распространения излучения). Энергия излучения W, проходящая через S в единицу времени, называется *потоком излучения*, обозначается через Φ_e и выражается в единицах мощности (Дж/с или Вт):

$$\Phi_e = \frac{W}{t}, \quad \text{или} \quad \Phi_e = \frac{dW}{dt}.$$

Поток излучения, проходящий через единичную площадку в единицу времени, т. е. отношение

$$\varphi = \frac{\Phi_e}{S}, \quad \text{или} \quad \varphi = \frac{d\Phi_e}{dS},$$

называется *поверхностной плотностью потока излучения* (или интенсивностью излучения) и выражается в ваттах на квадратный метр (Вт/м^2). Согласно формуле (6.13), приведенной в ч. I, § 30, поверхностная плотность потока излучения (интенсивность волны) равна $I = wc$, где w — энергия, содержащаяся в единице объема потока излучения, а c — скорость распространения волны в данной среде. Аналогичная формула может быть получена для оптического излучения:

$$\varphi = wc$$

(по электромагнитной теории, $w = \frac{\varepsilon_0 \varepsilon E^2}{2} + \frac{\mu_0 \mu H^2}{2}$, по фотонной теории, $w = N_1 h\nu_1 + N_2 h\nu_2 + \ldots = \Sigma N_i h\nu_i$, где N_i — число фотонов в единице объема, имеющих энергии $h\nu_i$).

Непрерывный спектр, из которого состоит данный поток излучения, можно разделить на узкие (монохроматические) участки, каждый из которых содержит некоторый интервал частот $\Delta\nu$ или длин волн $\Delta\lambda$. Весь поток излучения (т. е. мощность данного излучения) можно представить как сумму:

$$\Phi_e = \sum \Delta\Phi_{e\lambda},$$

где $\Delta\Phi_{e\lambda}$ — поток излучения (мощность), приходящийся на участок спектра от λ до $\lambda + \Delta\lambda$. Отношение $\Delta\Phi_{e\lambda}/\Delta\lambda$ или $d\Phi_{e\lambda}/d\lambda$ есть та мощность излучения, которая приходится на единичный интервал длин волн. Это отношение, как правило, различно в различных местах спектра данного излучения и, кроме того, зависит от полной мощности излучения Φ_e. Более удобной характеристикой излучения является относительная величина

$$\psi_\lambda = \frac{1}{\Phi_e} \frac{d\Phi_{e\lambda}}{d\lambda}, \tag{1.3}$$

которая является *функцией распределения* мощности данного излучения по его спектру. Если эта функция известна, то для заданного потока излучения можно рассчитать мощность, приходящуюся на часть спектра, ограниченную длинами волн λ_1 и λ_2:

$$\Delta\Phi_e(\lambda_1, \lambda_2) = \Phi_e \int_{\lambda_1}^{\lambda_2} \psi_\lambda \, d\lambda. \tag{1.4}$$

На рис. IV.2 показано различие между непрерывным (*а*), полосатым (*б*) и линейчатым (*в*) спектрами. Заштрихованная на графике непрерывного спектра площадка показывает ту долю полной мощности излучения, которая приходится на малый участок спектра от λ до $\lambda + d\lambda$.

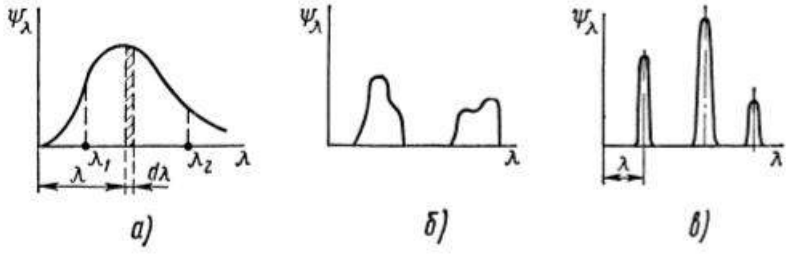


Рис. IV.2

Кроме распределения мощности излучения по длинам волн можно интересоваться распределением мощности по различным ориентациям плоскостей колебаний вектора **E**. Допустим, что излучение распространяется в направлении, перпендикулярном плоскости чертежа (рис. IV.3), так, что для читателя это направление проецируется в точку *O*. Далее предположим, что имеется такой приемник излучения, который воспринимает энергию электромагнитной волны только при определенной ориентации вектора **E** этой волны относительно прибора. Если при повороте приемника падающее на него излучение оказывает одно и то же воздействие, то можно утверждать, что в этом излучении энергия распределена по различным направлениям плоскостей колебаний равномерно (рис. IV.3, *а*). Однако существуют излучения, у которых распределение энергии по плоскостям колебаний неравномерное. На рис. IV.3, *б* условно показано распределение, имеющее максимум колебаний в плоскости $a - a$ и минимум — в плоскости $b - b$. Такое излучение называется *частично поляризованным*. Чем больше разница в энергиях, приходящихся на плоскости $a - a$ и $b - b$, тем более поляризованным является излучение. В пределе, если указанный выше приемник излучения обнаруживает энергию излучения только в строго определенной плоскости (рис. IV.3, *в*), т. е. если вектор **E** колеблется в одной плоскости, излучение будет идеально поляризованным *(линейно поляризованным)*; в реальных условиях может быть получено поляризованное излучение, у которого отклонение вектора **E** от определенной плоскости колебаний очень мало.

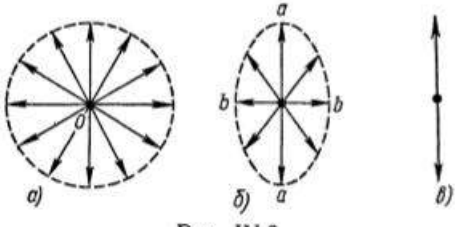


Рис. IV.3

Отметим несколько положений, важных при описании электромагнитных волн.

1) гармоническая электромагнитная волна определенной частоты ν или длины волны λ, описываемая формулой (1.2), энергии не содержит и не переносит. Согласно формуле (1.3), для узкого участка спектра

$$\Delta\Phi_{e\lambda} = \Phi_e \psi_\lambda \Delta\lambda;$$

следовательно, для определенной длины волны λ имеем $\Delta\lambda = 0$ и $\Delta\Phi_{e\lambda} = 0$. Энергия содержится только в некотором, не равном нулю интервале частот $\Delta\nu$ или длин волн $\Delta\lambda$. Поэтому наличие такого интервала у спектральных линий (см. рис. IV.2, *в*) следует рассматривать как обязательное условие того, что эта линия содержит и переносит энергию (заметим, что мощность монохроматического излучения зависит не только от интервала длин волн $\Delta\lambda$, но и от значений ψ_λ в пределах этого интервала; для волны, имеющей определенную длину волны λ, $\Delta\lambda = 0$ и конечное значение мощности возможно только при ψ_λ, равной бесконечности);

2) плоскополяризованная волна, в которой вектор напряженности электрического поля лежит точно в одной плоскости, энергии не содержит и не переносит. Если распределение энергии излучения по плоскостям колебаний выражается непрерывной функцией (которую мы обозначим ψ_α), то мощность той части излучения, у которой плоскости колебаний вектора **E** составляют с координатной плоскостью (проходящей через направление распространения волны) углы α, $\alpha + \Delta\alpha$, будет равна

$$\Delta\Phi_{e\alpha} = \Phi_e \psi_\alpha \Delta\alpha.$$

Следовательно, при $\Delta\alpha = 0$ имеем $\Delta\Phi_{e\alpha} = 0$. Мощность, которая содержится в частично поляризованной волне ($\Delta\alpha \neq 0$), зависит не только от интервала углов, в которых лежат плоскости колебаний, но и от значений ψ_α в пределах этого интервала. Для строго поляризованной волны $\Delta\alpha = 0$ и конечное значение энергии возможно только при $\psi_\alpha = \infty$.

Таким образом, гармоническая электромагнитная волна, описываемая формулой (1.2), имеющая строго определенную частоту колебаний, является абстракцией, такой же, как, например, «точечное тело» (материальная точка, обладающая конечной массой), «точечный заряд» и т. п. Абстракцией является также волна, у которой вектор **E** лежит строго в одной плоскости. Дело не в том, что мы не можем обнаружить или получить такие волны, а в том, что они в природе не существуют. Однако понятие гармонической плоскополяризованной волны значительно облегчает описание и изучение волновых процессов, так же как в теоретической механике — понятие материальной точки или в учении об электричестве — понятие точечного заряда и т. п.

Источники излучения подразделяются на: 1) *первичные* (самосветящиеся), преобразующие свою внутреннюю энергию в энергию излучения, и 2) *вторичные*, отражающие или пропускающие через себя падающее на них излучение.

Выделим на поверхности излучающего тела элементарную площадку dS и обозначим через $d\Phi_{e0}$ весь лучистый поток, испускаемый этой площадкой по всем направлениям (в одну сторону), а через $d\Phi_e$ — лучистый поток, исходящий из этой площадки в данном направлении $O - O_1$ внутрь элементарного телесного угла $d\omega$ (рис. IV.4). Тогда излучательную способность этой площадки можно характеризовать следующими величинами:

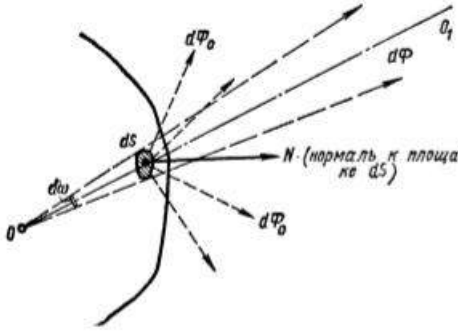


Рис. IV.4

$$R_e = \frac{d\Phi_{e0}}{dS}; \quad I_e = \frac{d\Phi_e}{d\omega}.$$

Величина R_e равна лучистой энергии, испускаемой единичной площадкой, взятой на поверхности тела, в единицу времени (выражается в Вт/м²). Эту величину называют *энергетической светимостью* данной поверхности. Величина I_e *(энергетическая сила света)* равна потоку излучения, который приходится на единицу телесного угла (выражается в ваттах на стерадиан: Вт/ср). Для точечного источника излучения

$$I_e = \frac{\Phi_e}{4\pi}; \quad \Phi_e = 4\pi I_e,$$

где Φ_e — весь поток излучения, исходящий из этого источника по всем направлениям; 4π — телесный угол вокруг точки.

Допустим теперь, что на поверхность какого-нибудь тела падает лучистый поток, причем на элементарную площадку dS приходится мощность излучения $d\Phi_e$. Отношение

$$E_e = \frac{d\Phi_e}{dS}$$

называется *энергетической освещенностью* этой площадки (выражается в Вт/м²). Энергетическая освещенность есть лучистая энергия, ежесекундно падающая на единицу облучаемой поверхности; она может быть одинаковой или различной вдоль этой поверхности.

Излучение, падающее на поверхность тела, оказывает на него давление. Максвелл указал на существование *светового давления*, исходя из электромагнитной теории света. Вектор **E** волны приводит в упорядоченное движение элементарные заряды в веществе тела, а магнитное поле **H** действует на эти заряды с силой Лоренца. Эта сила оказывается направленной в сторону распространения излучения

(рис. IV.5); равнодействующая всех этих сил воспринимается как давление, оказываемое излучением на тела. Допустим, что на единицу поверхности тела перпендикулярно ей падает поток излучения с поверхностной плотностью $\varphi = wc$. Максвелл показал, что если излучение полностью поглощается телом, то оказываемое им давление равно $p = \varphi/c = w$, т. е. равно энергии, содержащейся в единице объема излучения. Если из потока излучения с поверхностной плотностью φ некоторая часть $\rho\varphi$ отражается (ρ — коэффициент отражения), то давление света оказывается равным

$$p = (1+\rho)\frac{\varphi}{c} = (1+\rho)\,w. \qquad (1.5)$$

Для абсолютно отражающей поверхности ($\rho = 1$) давление излучения вдвое больше, чем для абсолютно поглощающей поверхности ($\rho = 0$). Давление солнечного излучения (в ясный день) на поглощающую поверхность, ориентированную перпендикулярно к лучам, приблизительно равно $4 \cdot 10^{-6}$ Па. Измерение светового давления впервые было выполнено П. Н. Лебедевым (1900).

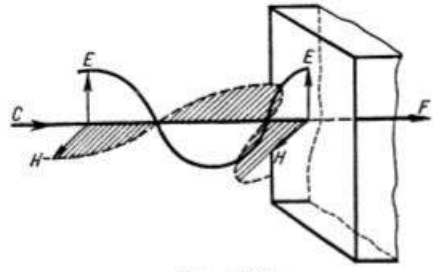


Рис. IV.5

Существование давления, оказываемого на тела излучением, следует также из фотонной теории. Допустим, что в единицу времени на единицу поверхности перпендикулярно ей падает N фотонов, каждый из которых имеет импульс, равный

$$m_\Phi c = h\nu/c.$$

Из них N_1 фотонов поглощается телом, а N_2 — отражается, так что $N_2/N = \rho$ — коэффициент отражения. Так как изменение количества движения на поверхности тела при поглощении фотона равно $h\nu/c$, а при отражении будет вдвое больше и равно $2h\nu/c$, то на поверхности тела происходит ежесекундное изменение количества движения, равное

$$N_1\frac{h\nu}{c} + N_2 \cdot 2\frac{h\nu}{c} = N\frac{h\nu}{c}(1+\rho). \qquad (1.6)$$

Эту величину следует приравнять силе, действующей на рассматриваемую единичную площадку. Так как $Nh\nu = I$ — энергия, ежесекундно падающая на эту поверхность, то мы получаем формулу (1.5). Результат не изменится, если на тело падает излучение, состоящее из фотонов с различными частотами.

Энергия электромагнитного излучения воспринимается и измеряется по ее воздействию на различные приемники излучения, в которых происходит превращение этой энергии в другие виды. Среди них особое значение имеет глаз. Каждый из приемников излучения имеет различную чувствительность (реакцию) на различные участки спектра. Допустим, что монохроматический поток $d\Phi_{e\lambda}$ (содержащий излуче-

ние в пределах λ, $\lambda + d\lambda$) вызвал в приемнике «показание», равное dX_λ. «Показанием приемника излучения» может быть: изменение температуры или электрического сопротивления датчика, сила тока в цепи фотоэлемента, электродвижущая сила в цепи термоэлемента, количество вещества, выделившегося при химической реакции, и т. д.

Отношение

$$\Psi_\lambda = \frac{dX_\lambda}{d\Phi_{e\lambda}} \tag{1.7}$$

называется *спектральной чувствительностью* данного приемника излучения к длине волны λ. Если для одной длины волны λ_0 эта чувствительность максимальная и равна ψ^0_λ, то величину $\psi_\lambda/\psi^0_\lambda$ называют *относительной* спектральной чувствительностью. Для излучения, охватывающего широкий интервал длин волн от λ_1 до λ_2, введено понятие *интегральной* чувствительности

$$\Psi = X/\Phi_e = \left(\int_{\lambda_1}^{\lambda_2} dX_\lambda\right) \Big/ \left(\int_{\lambda_1}^{\lambda_2} d\Phi_{e\lambda}\right).$$

Глаз является приемником излучения, обладающим различной чувствительностью в пределах видимого спектра. Спектральную чувствительность глаза как функцию от длины волны обозначают через K_λ и называют *спектральной световой эффективностью*. Реакцию глаза на воспринимаемое им излучение называют *световым ощущением*; произведение лучистого потока $d\Phi_{e\lambda}$, содержащего интервал длин волн от λ до $\lambda + d\lambda$, на спектральную чувствительность глаза для этого участка спектра K_λ

$$K_\lambda\, d\Phi_{e\lambda} = d\Phi_\lambda$$

называют *монохроматическим световым потоком*, соответствующим участку спектра λ, $\lambda + d\lambda$. Световой поток выражается в люменах (лм), поэтому K_λ выражается в лм/Вт.

Наибольшей чувствительностью глаз обладает для $\lambda_0 = 0{,}555$ мкм, причем $K_м = 625$ лм/Вт. Это означает, что монохроматический лучистый поток, содержащий излучение с длиной волны около 0,555 мкм, мощностью в 1 Вт, воспринимается как световой поток желто-зеленого цвета, равный 625 лм. Обратная величина $M = 1/K_м = 0{,}0016$ Вт/лм называется *механическим эквивалентом света* (для $\lambda_0 = 0{,}555$ мкм).

Относительная спектральная световая эффективность (чувствительность) глаза $V_\lambda = K_\lambda/K_м$ (иногда называемая *функцией видности*) имеет, по усредненным данным, следующие значения, мкм:

λ	0,4	0,5	0,555	0,590	0,650	0,700	0,750
V_λ	0,0004	0,323	1	0,757	0,107	0,0041	0,00006

График этой функции приведен на рис. IV.6.

По спектральной плотности излучения в данном участке спектра $\varphi_\lambda = d\Phi_{e\lambda}/d\lambda$ можно найти соответствующую спектральную плотность в световом потоке умножением на K_λ.

Заметим, что белый свет, испускаемый, например, раскаленными твердыми телами, содержит все длины волн видимого спектра; однако

не всякое излучение, содержащее все видимые волны, будет восприниматься глазом как белый свет. Для этого необходимо еще определенное распределение энергии по спектру.

Перечислим фотометрические величины, характеризующие источник света:

1) *световой поток* Φ есть величина

$$\Phi = \int_{400}^{700} K_\lambda \, d\Phi_{e\lambda},$$

где $d\Phi_{e\lambda}$ — часть потока излучения, приходящаяся на участок спектра от λ до $\lambda + d\lambda$. Отношение $\Phi/\Phi_{e\lambda}$ показывает, сколько люменов светового потока соответствует потоку излучения в 1 Вт. Для монохроматического излучения длины волны λ 1 Вт лучистого потока равен $625 \cdot V_\lambda$ люменов светового потока;

2) *световая энергия*, излучаемая источником за время t, выражается в люмен-секундах (лм·с):

$$Q = \int_0^t \Phi(t)\, dt = \Phi_{ср} t,$$

где $\Phi_{ср}$ — среднее значение светового потока за время t;

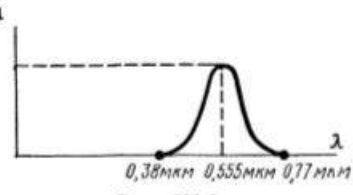


Рис. IV.6

3) *светимость* поверхности источника — отношение излучаемого светового потока к площади светящейся (собственным или отраженным светом) поверхности — выражается в лм/м²:

$$R_{ср} = \frac{\Delta\Phi}{\Delta S}; \quad R = \frac{d\Phi}{dS}.$$

Светимость киноэкранов должна быть 50—100 лм/м², бумаги при письме или чтении — 35—50 лм/м²; светимость снежного покрова в безлунную ночь — 0,002 лм/м²;

4) *сила света* источника в данном направлении есть отношение светового потока $d\Phi$, испускаемого им в этом направлении внутрь телесного угла $d\Omega$, к величине этого угла:

$$I_{ср} = \frac{\Delta\Phi}{\Delta\Omega}; \quad I = \frac{d\Phi}{d\Omega}. \qquad (1.8)$$

Сила света измеряется в *канделах* (ранее эта единица силы света называлась «свечой»; определение канделы дано ниже). Источник света в 1 кд излучает по всем направлениям световой поток, равный $4\pi I$ люменов;

5) *яркость* светящейся поверхности в данном направлении есть отношение светового потока $\Delta\Phi$, излучаемого площадкой ΔS в этом направлении внутрь телесного угла $\Delta\Omega$, к величине этого угла и к проекции площадки ΔS на плоскость, перпендикулярную указанному направлению (рис. IV.7):

$$B = \frac{\Delta\Phi}{\Delta\Omega \, \Delta S \cos i}.$$

Яркость светящейся поверхности измеряется в *нитах* (нт). Так как $\Delta\Phi/\Delta\Omega = \Delta I$ есть сила света площадки ΔS в данном направлении, то для определения яркости необходимо эту силу света ΔI разделить на указанную проекцию площадки $\Delta S \cos i$, где i — угол между данным направлением и нормалью к площадке:

$$B = \frac{\Delta I_e}{\Delta S \cos i}.$$

Нит — яркость площадки, которая в перпендикулярном ей направлении имеет отношение силы света к площади dI/dS_0, равное 1 кд/м². Если это отношение равно 1 кд/см², то яркость равна одному *стильбу* (сб), следовательно, 1 нт = 10^{-4} сб.

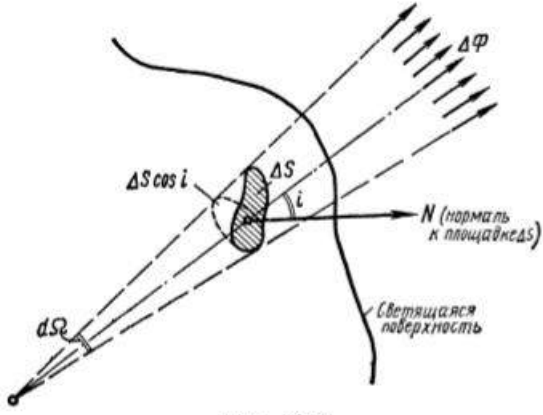


Рис. IV.7

Яркость листа белой бумаги при чтении и письме должна быть не менее 0,001 сб; люминесцентной лампы — 0,7 сб; нити накаливания электрической лампочки — около 500 сб; поверхности Солнца — 120 000 сб; белой поверхности, освещенной полной Луной, — около $6 \cdot 10^{-6}$ сб; ночного (безлунного) неба — 10^{-8} сб. Наименьшая яркость светящихся поверхностей, воспринимаемая глазом в полной темноте, — 10^{-9}, сб — называется пороговой яркостью для человеческого глаза.

В Международной системе единиц дано следующее определение единицы силы света: к а н д е л а — сила света, испускаемого с площади 1/600000 м² сечения полного излучателя в перпендикулярном этому сечению направлении при температуре излучателя, равной температуре затвердевания платины при давлении 101 325 Па (определение полного излучателя, т. е. так называемого абсолютно черного тела, будет дано в § 15).

Яркость светящихся поверхностей различна в различных направлениях; если она одинакова по всем направлениям, то говорят, что источник света подчиняется з а к о н у Л а м б е р т а. С некоторым приближением такими источниками являются: поверхность Солнца; колпак из матового стекла, освещенный изнутри; поверхность, покрытая окисью магния, и т. д. Для таких источников легко рассчиты-

ваются полный световой поток $\Delta\Phi$, излучаемый площадкой ΔS по всем направлениям в одну сторону $\Delta\Phi = \pi B \Delta S$, и светимость $R = \pi B$.

При падении светового потока на поверхность тела определяют:

1) *освещенность* поверхности — отношение светового потока, равномерно распределенного по этой поверхности, к ее площади (выражается в люксах, 1 лк = 1 лм/м^2):

$$E_{ср} = \frac{\Delta\Phi}{\Delta S}; \quad E = \frac{d\Phi}{dS}.$$

Освещенность земной поверхности при полной Луне (в зените) равна 0,2 лк, в безлунную ночь — 0,002 лк;

2) *световая экспозиция* — произведение освещенности на время освещения [измеряется в люкс-секундах (лк·с)]:

$$H = E_{ср} t; \quad H = \int_0^t E(t)\, dt.$$

Поверхности тел рассеивают падающий на них световой поток. Если каждый элемент поверхности рассеивает свет по всем направлениям (в одну сторону, т. е. в телесный угол 2π для каждой точки) с одинаковым коэффициентом отражения ρ, то поверхность называется *диффузно рассеивающей*; для нее соблюдаются следующие соотношения между светимостью, освещенностью и яркостью:

$$R = \rho E; \quad B = \rho \frac{E}{\pi}. \tag{1.9}$$

Коэффициент диффузного отражения для белых экранов, покрытых окисью магния, $\rho = 0{,}98$, белой клеевой краской — 0,8. Если $\rho = 1$ для всех направлений и длин волн, то поверхность является идеальным рассеивателем; для нее $R = E$, $B = E/\pi$; яркость рассеивающих поверхностей выражают в *апостильбах* (1асб = $1/\pi$ нт).

§ 3. РАСПРОСТРАНЕНИЕ, ОТРАЖЕНИЕ, ПРЕЛОМЛЕНИЕ И ПОГЛОЩЕНИЕ СВЕТА. ДИСПЕРСИЯ. РАССЕЯНИЕ СВЕТА

Допустим, что в некоторой среде точечный источник света в момент $t_0 = 0$ начал излучать монохроматическую световую волну. Спустя время t эта волна, распространяясь по всем направлениям, достигнет некоторой поверхности, называемой *фронтом световой волны*. Форма этой поверхности (в каждый данный момент времени) определяется оптическими свойствами среды. В пространстве между излучателем и фронтом волны можно выделить также *волновые поверхности*, в различных точках которых вектор **E** имеет одинаковую фазу. Среда называется «оптически однородной и изотропной», если в этой среде скорость распространения оптического излучения одинакова во всех местах (однородность) и по всем направлениям (изотропность). В такой среде по истечении времени t волна от точечного источника достигнет поверхности, имеющей вид сферы радиуса $r = ct$ (рис. IV.8, *а*). Если же среда анизотропная, например скорость распространения

волны постепенно изменяется при переходе от одного направления к соседнему, достигая максимума в направлении $Z - Z$, минимума — в направлениях, лежащих в плоскости XOY, то форма фронта волны, а также волновой поверхности будет эллипсоидом вращения с осью $Z - Z$ (рис. IV.8, *б*).

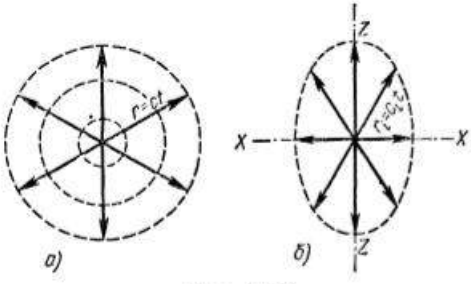


Рис. IV.8

Линии, вдоль которых распространяется световая энергия, называются *световыми лучами*. В изотропной среде световые лучи направлены по нормалям к волновой поверхности; в анизотропных средах лучи и нормали к волновым поверхностям не совпадают.

Если известны форма и расположение фронта световой волны в некоторый момент времени t, то фронт волны в последующие моменты времени $t + \Delta t$ можно найти по принципу Гюйгенса:

каждую точку заданного фронта волны в момент t следует рассматривать как самостоятельный точечный источник света, начинающий излучать световую волну в этот момент t. Построив волновые поверхности элементарных источников по истечении времени Δt, следует провести их огибающую, которая и будет искомым фронтом волны в момент $t + \Delta t$.

На рис. IV.9 показаны применения принципа Гюйгенса для плоского (*а*) и сферического (*б*) фронта волны в однородной и изотропной

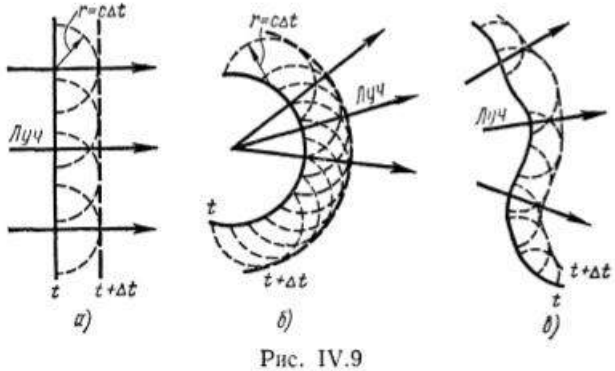


Рис. IV.9

среде, а также для волны произвольной формы в неоднородной среде (*в*).

В дальнейшем ограничимся рассмотрением световых волн, фронт которых является плоскостью; такие волны для краткости называют *плоскими волнами*.

Допустим, что плоская волна падает на границу двух различных сред, в которых скорости света равны c_1 и c_2. На границе свет частично отражается, частично проходит во вторую среду. Для нахождения законов отражения и преломления световых лучей применим принцип Гюйгенса; найдем огибающую элементарных волн для точечных источников света, которые расположены на поверхности раздела двух сред (рис. IV.10, *а*). Когда луч *1* дойдет до точки A, из этой точки начнет распространяться сферическая волна в первой среде со скоростью c_1,

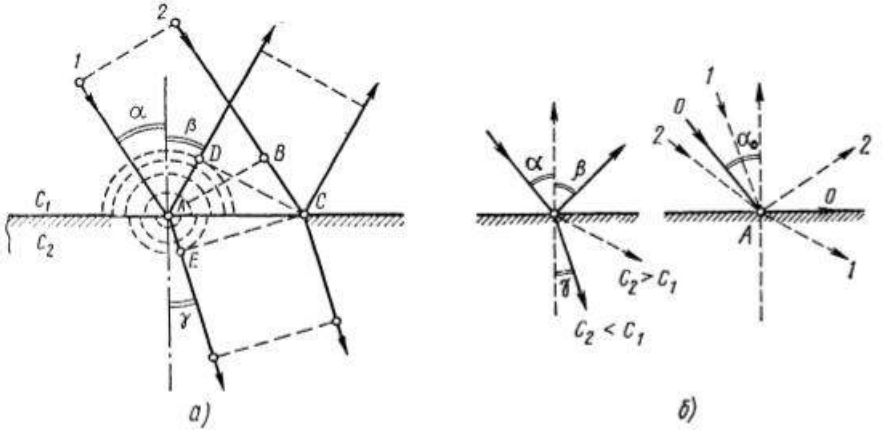


Рис. IV.10

во второй — со скоростью c_2; спустя время BC/c_1, когда до границы раздела дойдет луч *2*, радиусы этих сфер сделаются равными соответственно:

$$r_1 = c_1 \frac{BC}{c_1} = BC; \quad r_2 = c_2 \frac{BC}{c_1} = BC \frac{c_2}{c_1}.$$

Опуская построение волновых поверхностей для промежуточных лучей, можно получить фронты отраженной и преломленной волны, если из точки C провести касательные плоскости к сферическим волнам, исходящим из точки A в первую и вторую среды. Так как $BC = AD$, $AE/AD = c_2/c_1$, то получаются следующие результаты:

1) угол отражения β лучей от границы двух сред равен углу падения α; этот закон соблюдается для всех длин волн и для любых сред;

2) отношение синуса угла падения α к синусу угла преломления γ равно отношению скорости света в первой среде (откуда идет свет) к скорости света во второй среде (куда идет свет):

$$\frac{\sin \alpha}{\sin \gamma} = \frac{c_1}{c_2}. \qquad (1.10)$$

Это отношение не зависит от значений углов α и γ и определяется только оптическими свойствами сред; его называют *относительным показателем преломления* второй среды по отношению к первой.

Если в качестве первой среды всегда брать вакуум, в котором скорость света равна c_0, то отношение

$$c_0/c = n \tag{1.11}$$

называют *абсолютным показателем преломления* данной среды. По абсолютным показателям преломления двух данных сред $n_1 = c_0/c$ и $n_2 = c_0/c_2$ можно найти их относительный показатель преломления

$$n_{12} = c_1/c_2 = n_2/n_1. \tag{1.12}$$

Абсолютный показатель преломления является одной из важнейших оптических характеристик вещества.

При переходе света из более преломляющей среды в менее преломляющую ($n_1 > n_2$) угол преломления γ будет больше угла падения α. С увеличением α при некотором $\alpha = \alpha_0$ можно получить $\gamma = \pi/2$, т. е. преломленный луч будет скользить по поверхности раздела этих сред (рис. IV.10, *б*). Угол, удовлетворяющий условию $\sin \alpha_0 = n_2/n_1$, при котором $\gamma = \pi/2$, $\sin \gamma = 1$, называется *предельным углом* падения луча. При $\alpha < \alpha_0$ падающий луч частично отражается, частично проходит во вторую среду, испытав преломление; при $\alpha > \alpha_0$ луч не преломляется, а полностью отражается обратно в первую среду. Это явление *полного внутреннего отражения* используется, например, в стеклянных отражающих призмах, с успехом заменяющих металлические зеркала (поверхность которых со временем тускнеет).

Рассмотрим *отражение* света от границы двух сред. Обозначим световые потоки: Φ — падающий, $\Phi_{отр}$ — отраженный и $\Phi_п$ — прошедший во вторую среду. Коэффициент отражения $\rho = \Phi_{отр}/\Phi$ зависит от следующих факторов:

1) от показателей преломления обеих сред; при нормальном падении лучей

$$\rho = \left(\frac{n-1}{n+1}\right)^2.$$

На границе между воздухом и стеклом $n = 1{,}5$ и $\rho \approx 0{,}04$ ($\approx 4\%$);

2) от длины волны падающего излучения. Например, для серебряного зеркала, мкм:

λ	0,251	0,305	0,316	0,326	0,338	0,420	0,450	0,500	0,700	1,0
ρ	0,34	0,091	0,012	0,146	0,555	0,866	0,905	0,913	0,96	0,975

3) от спектрального состава излучения; если коэффициент отражения различен для различных λ, то распределение энергии по спектру при отражении изменится; этим объясняется, почему некоторые тела при освещении их белым светом выглядят окрашенными;

4) от угла падения лучей;

5) от угла между вектором **E** падающего излучения и плоскостью падения (в которой лежат падающий луч и перпендикуляр к отражающей поверхности).

Та часть светового потока $\Phi - \Phi_{отр} = \Phi_п$, которая пройдет во вторую среду через границу раздела, будет поглощаться в этой среде. Если толщина среды недостаточна для полного поглощения, то часть

светового потока поглотится ($\Phi_{погл}$), а другая часть ($\Phi_{пр}$) выйдет из этой среды. Отношения

$$\alpha = \frac{\Phi_{погл}}{\Phi_п}; \quad \tau = \frac{\Phi_{пр}}{\Phi_п}$$

зависят от оптических свойств среды и толщины поглощающего слоя.

Выведем формулу для расчета поглощаемого светового потока в зависимости от толщины слоя среды. Сначала проведем этот расчет для полного потока (мощности) излучения. Допустим, что в элементарном слое толщиной dx поглощается некоторая мощность излучения $d\Phi_e$. Можно полагать, что $d\Phi_e$ прямо пропорциональна Φ_e и толщине слоя dx:

$$-d\Phi_e = a\Phi_e\, dx.$$

Знак минус означает потерю мощности, а коэффициент a учитывает поглощательные свойства среды. Из этого соотношения получаем:

$$\frac{d\Phi_e}{\Phi_e} = -a\, dx; \quad \ln\frac{\Phi_e}{\Phi_{e0}} = -ax,$$

где x — толщина слоя, Φ_{e0} — значение мощности при $x = 0$. Таким образом,

$$\Phi_e = \Phi_{e0}e^{-ax}. \tag{1.13}$$

Эта формула (*закон Бугера*) показывает, что мощность излучения в поглощающей среде уменьшается по экспоненциальному закону.

У многих веществ коэффициент поглощения a зависит от длины волны. Некоторые вещества обнаруживают *селективное* поглощение, т. е. имеют очень большой коэффициент поглощения для какого-то (обычно очень узкого) участка спектра. Например, серебро очень сильно поглощает ультрафиолетовое излучение с $\lambda = 0{,}316$ мкм.

Соотношение (1.13) применимо также и для светового потока

$$\Phi = \Phi_0 e^{-ax}, \tag{1.14}$$

где Φ_0 — световой поток на входной границе поглощающего слоя ($x = 0$). Ввиду зависимости коэффициента поглощения от длины волны средние значения a для всего потока излучения (Φ_e) и для видимой его части (Φ) могут быть различными.

Если на пластинку (слой поглощающей среды) с толщиной d падает световой поток Φ, то

$$\Phi = \Phi_{отр} + \Phi_{погл} + \Phi_{пр}; \quad \Phi_{погл} = (\Phi - \Phi_{отр})\, e^{-ad}.$$

Таким образом,

$$\rho + \alpha + \tau = 1.$$

СВЕТ И ЦВЕТ

Зависимость коэффициентов отражения (ρ) и поглощения (α) от длины волны является физической причиной *окрашенности* тел, не излучающих собственного света. Если, например, какое-либо

тело при его освещении белым светом имеет красный цвет, то это означает, что тело имеет относительно большие коэффициенты отражения для красного участка спектра и поглощения для зелено-фиолетовой части спектра. Заметим также, что цвет тела зависит, кроме того, от спектрального состава падающего на него света. Если указанное выше тело осветить, например, синим светом, то оно будет казаться почти черным, так как синие лучи будут им почти полностью поглощаться и отраженный свет будет отсутствовать. Окраска тел, пропускающих свет, будет определяться зависимостью коэффициента поглощения от длины волны. Если вещество очень сильно поглощает излучение всех длин волн, кроме, например, зеленых, то можно подобрать достаточно большую толщину слоя этого вещества с таким расчетом, чтобы при пропускании через него белого света в прошедшем излучении почти отсутствовали все остальные участки спектра, кроме зеленого. Такие вещества (обычно различные краски) используются для изготовления *светофильтров*. Обычно подбирают смесь красок и вводят их в состав стекла, слоя желатина или прозрачной пластмассы. При малой толщине поглощающего слоя прошедшее излучение не будет в достаточной степени монохроматическим. Однако с увеличением толщины слоя уменьшается также и интенсивность выделяемого монохроматического излучения.

ДИСПЕРСИЯ СВЕТА

Световые волны различных частот колебаний распространяются в вакууме с одинаковыми скоростями ($c \approx 3 \cdot 10^8$ м/с), а в средах — с различными скоростями. Например, в обыкновенном стекле красный свет распространяется с большей скоростью, чем фиолетовый. Вследствие этого показатель преломления оказывается различным для света различных частот колебаний.

Дисперсией вещества по отношению к световым волнам (дисперсией света) называется зависимость скорости распространения этих волн от частоты колебаний (или длины волны); вследствие этого абсолютный показатель преломления данного вещества также зависит от частоты или длины волны проходящего через него света. Дисперсия света в веществе определяется видом функции $n = n(\nu)$ или $n = n(\lambda)$. В различных участках спектра дисперсия характеризуется также тем изменением показателя преломления, который приходится на единичный интервал длин волн (величина $dn/d\lambda$ есть скорость изменения показателя преломления в данном месте спектра).

В справочных таблицах показатели преломления различных веществ обычно дают для желтой линии натрия $\lambda_D = 0{,}5893$ мкм, синей $\lambda_F = 0{,}4861$ мкм и красной $\lambda_C = 0{,}6563$ мкм линий водорода. Соответствующие им показатели преломления обозначаются n_D и n_F и n_C. Разность $n_F - n_C$ называется *средней дисперсией*.

Величина $\frac{n_F - n_C}{n_D - 1}$ называется *относительной дисперсией*, а обратная ей величина — *коэффициентом дисперсии среды*; вещества, у которых $n_F - n_C$ мала, а коэффициент дисперсии велик, имеют более

плавное изменение показателя преломления от длины волны и называются «веществами с малой дисперсией».

Зависимость $n = n(\lambda)$ в оптической области спектра имеет сложный характер; для тех участков спектра, которые слабо поглощаются данным веществом (т. е. для которых это вещество прозрачно), зависимость показателя преломления от длины волны может быть с удовлетворительной точностью представлена в виде формулы Коши

$$n = n_0 + \frac{a}{\lambda^2} + \frac{b}{\lambda^4} \quad \text{или} \quad n \approx n_0 + \frac{a}{\lambda^2}, \tag{1.15}$$

где n_0, a и b — некоторые постоянные для данного вещества величины. При $\lambda \to \infty$ $n \to n_0$. Дисперсию вещества для этих участков спектра называют *нормальной*; здесь

$$\frac{dn}{d\lambda} \approx -\frac{2a}{\lambda^3}.$$

Для тех участков спектра, которые сильно поглощаются веществом, показатель преломления с увеличением длины волны изменяется иначе: сначала он резко уменьшается, приобретая значения, меньшие n_0, затем быстро увеличивается и, достигнув максимума, вновь резко уменьшается. В этом случае дисперсию вещества называют *аномальной*. На рис. IV.11 изображен ход зависимости n от λ («кривая дисперсии»), в которой выделяются области нормальной (*I* и *III*) и аномальной (*II*) дисперсий.

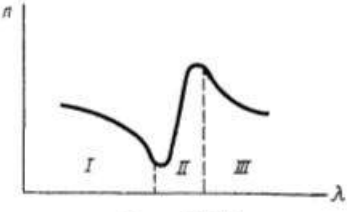


Рис. IV.11

В большом интервале длин волн у каждого вещества обнаруживается несколько таких областей аномальной дисперсии. Согласно теории дисперсии, которая здесь не рассматривается, аномальная дисперсия должна наблюдаться при резонансе между колебаниями вектора **E** проходящей волны и собственными колебаниями электрических зарядов в атомах и молекулах вещества. Поэтому по измеренным частотам областей аномальной дисперсии можно определять частоты собственных колебаний в атомах и молекулах вещества. Кроме того, при резонансе должна наблюдаться также интенсивная передача энергии от проходящей волны к атомам и молекулам вещества, т. е. должно иметь место интенсивное поглощение веществом энергии проходящего излучения.

Измерения показали, что в области аномальной дисперсии кривая зависимости коэффициента поглощения от длины волны имеет резкий максимум.

РАССЕЯНИЕ СВЕТА

Рассмотрим рассеяние света различными веществами. При прохождении световой волны через вещество электрические заряды в его атомах и молекулах под действием переменного светового вектора **E** совершают вынужденные колебания с той же частотой. При этом ча-

стицы среды сами становятся вторичными излучателями электромагнитных волн, которые распространяются по различным направлениям. Таким образом, часть энергии волны, проходящей через вещество, поглощается и вновь излучается его частицами и вследствие этого *рассеивается* по всевозможным направлениям. Рассеяние света может произойти также и при отражениях и преломлениях света на границах мельчайших частиц (пылинок, капелек, пузырьков), содержащихся в данной среде.

Прозрачные среды (оптические стекла, очень чистые прозрачные жидкости и газы) почти не рассеивают света; это объясняется тем, что вторичные волны, излучаемые частицами среды, вследствие интерференции взаимно гасятся по всем направлениям, кроме направления распространения проходящего света. При этом, как показал Л. И. Мандельштам, важна однородность среды, так как для полного гашения необходима не только когерентность, но и равенство интенсивностей интерферирующих волн.

При наличии неоднородностей интенсивности вторичных волн будут иметь в различных местах и направлениях различные значения и поэтому полного гашения их не получится. Особенно сильно рассеивается свет в так называемых «мутных средах» (молочное стекло, туманы, дым, молоко, суспензии и эмульсии и т. д.).

Наблюдения и расчеты показали, что:

1) интенсивность рассеянного света пропорциональна четвертой степени частоты или обратно пропорциональна четвертой степени длины волны:

$$I = \text{const}_1 \cdot \nu^4 = \frac{\text{const}_2}{\lambda^4}$$

(закон Релея). Вследствие этого при прохождении белого света через рассеивающую среду рассеянный свет имеет голубоватый, а прошедший — красноватый оттенок. Практически рассеяние света, по закону Релея, имеет место при $\lambda \gg a$, где a — параметр, характеризующий линейные размеры рассеивающих частиц среды;

2) интенсивность рассеянного света различна в различных направлениях и может быть вычислена по формуле

$$I_\alpha = \frac{1}{2} I_{\text{макс}} (1 + \cos^2 \alpha),$$

где I_α — интенсивность рассеянного света в направлении, составляющем угол α с направлением проходящего света; $I_{\text{макс}}$ — максимальная интенсивность вторичного (рассеянного) излучения (это имеет место в направлении проходящего света);

3) свет, рассеянный под углом $\alpha = \pi/2$ к направлению проходящего излучения, плоскополяризован.

Рассеяние света в однородных средах возможно ввиду того, что в объеме этой среды при беспорядочном (тепловом) движении молекул возможны случайные отклонения плотности среды от среднего по всему объему значения; в некоторых местах происходит временное скопление молекул и увеличение плотности, в других — уменьшение плотности. Эти флуктуации плотности среды означают появление оп-

тической неоднородности, так как показатель преломления зависит от плотности вещества. Рассеяние света на этих неоднородностях называется *молекулярным рассеянием*. Интенсивность его небольшая; например, воздух в нормальных условиях рассеивает приблизительно $3 \cdot 10^{-7}$ поступающей в его объем энергии, вода — $5 \cdot 10^{-5}$ и т. д. Молекулярным рассеянием объясняется голубой цвет неба. Флуктуации плотности происходят особенно интенсивно в критическом состоянии вещества; они вызывают заметное помутнение вещества при прохождении его через это состояние.

§ 4. ИНТЕРФЕРЕНЦИЯ СВЕТА; КОГЕРЕНТНОСТЬ СВЕТОВЫХ ЛУЧЕЙ. ИНТЕРФЕРОМЕТРЫ

Допустим, в точку A одновременно поступает монохроматическое плоскополяризованное излучение от двух источников света: S_1 и S_2. Если векторы напряженности электрических полей этих волн $\mathbf{E}_1$ и $\mathbf{E}_2$ имеют в точке A одинаковое направление (рис. IV.12), то суммарная напряженность поля в этой точке A будет равна $\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2$. Единица объема, взятая в окрестности точки A, ежесекундно получит от первой световой волны энергию электрического поля, равную $\varepsilon_0 E_1^2/2$ (рассуждения ведутся для вакуума), а от второй волны — энергию $\varepsilon_0 E_2^2/2$. Однако энергия в единице объема, подсчитанная по суммарной напряженности E,

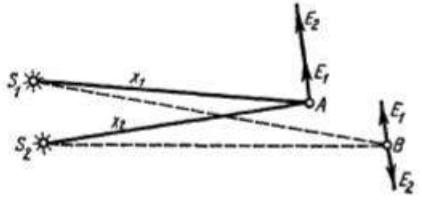


Рис. IV.12

$$\frac{\varepsilon_0 E^2}{2} = \frac{\varepsilon_0 (E_1 + E_2)^2}{2} = \frac{\varepsilon_0 E_1^2}{2} + \frac{\varepsilon_0 E_2^2}{2} + \varepsilon_0 E_1 E_2$$

оказывается больше суммы энергий, подсчитанных по напряженностям E_1 и E_2 каждой волны в отдельности. Казалось бы, закон сохранения энергии здесь нарушается, так как энергия в некотором объеме вокруг точки A оказывается больше, чем сумма энергий, посылаемых туда каждым источником света S_1 и S_2 в отдельности. Но мы не учли, что в соседних точках среды (например, в точке B), куда световые волны от S_1 и S_2 приходят, имея противоположно направленные векторы $\mathbf{E}_1$ и $\mathbf{E}_2$, энергия в единице объема, подсчитанная по суммарной напряженности $\mathbf{E} = \mathbf{E}_1 - \mathbf{E}_2$, будет меньше, чем сумма энергий, $\varepsilon_0 E_1^2/2 + \varepsilon_0 E_2^2/2$, посылаемых в этот объем источниками света S_1 и S_2. Следовательно, в результате сложения электрических (а также и магнитных) полей световых волн происходит *перераспределение энергии* этих волн в пространстве. В некоторых местах две световые волны, складываясь, дают «темноту»; в этих местах энергии $\varepsilon_0 E_1^2/2$ и $\varepsilon_0 E_2^2/2$ (которые поступали бы от каждой световой волны при отсутствии другой) при совместном существовании этих волн оказываются «передвинутыми» в соседние области.

Интерференцией света называется сложение нескольких световых волн, в результате которого образуются чередующиеся светлые и тем-

ные области, т. е. происходит перераспределение энергии этих волн в пространстве (вдоль фронта волны). Рассмотрим простейший случай интерференции монохроматических световых волн от двух одинаковых источников, причем предположим, что эти источники стали излучать одновременно, поэтому фазы векторов $\mathbf{E}_1$ и $\mathbf{E}_2$, определяющие их значения, согласно уравнению волны (1.2), зависят только от расстояний x_1 и x_2 (рис. IV.12). Тогда

$$E_1 = E_0 \sin \omega \left(t - \frac{x_1}{c}\right), \quad E_2 = E_0 \sin \omega \left(t - \frac{x_2}{c}\right);$$

$$E = E_1 + E_2 = 2E_0 \cos \omega \frac{x_2 - x_1}{2c} \sin \omega \left(t - \frac{x_1 + x_2}{2c}\right). \qquad (1.16)$$

В точках, где аргумент косинуса равен нечетному числу $\pi/2$, $\cos \omega \frac{x_2 - x_1}{2c} = 0$, суммарная напряженность электрического поля E в любой момент времени равна нулю и световые волны взаимно «гасятся». Подставляя $\omega = 2\pi/T = 2\pi c/\lambda$, можно найти расположение тех точек, в которых происходит взаимное «гашение» двух монохроматических световых волн (с одинаковой амплитудой E_0):

$$2\pi \frac{c}{\lambda} \frac{x_2 - x_1}{2c} = (2k+1) \frac{\pi}{2}; \quad x_2 - x_1 = \frac{\lambda}{2}(2k+1), \qquad (1.17)$$

где $k = 0, 1, 2, \ldots$ — целое число. В точках, отстоящих от источников света S_1 и S_2 на расстояниях, удовлетворяющих условию (1.17), света не будет. Расположение точек, в которых амплитуда суммарной напряженности E имеет максимум, определяется условием $\cos \omega \frac{x_2 - x_1}{2c} = \pm 1$, $x_2 - x_1 = k\lambda$. Таким образом, точки с максимумом и минимумом световых колебаний чередуются в пространстве в зависимости от величины $x_2 - x_1$ (очевидно, точки с одинаковым значением вектора $\mathbf{E}$ лежат на гиперболоиде вращения $x_2 - x_1 = \text{const}$, ось которого проходит через точки S). Весьма важно отметить, что для постоянного взаимного гашения или усиления световых волн в какой-либо точке пространства необходимо, чтобы разность фаз волн, приходящих в данную точку, с течением времени не изменялась.

Световые волны (одинаковой длины волны), которые приходят в данную точку с постоянной (не изменяющейся со временем) разностью фаз, называются *когерентными*. Когерентные волны дают интерференционную картину (распределение интенсивности света в пространстве или на экране), которая с течением времени сохраняется неизменной; ее можно наблюдать визуально, фотографировать, измерять расстояние между местами с максимумами или минимумами света и т. д. Некогерентные лучи не дают стабильной (устойчивой, постоянной) интерференционной картины; непрерывное изменение фаз складываемых колебаний в каждой точке пространства создает быстро меняющуюся картину, которую невозможно использовать для измерительных целей.

Излучение электромагнитных волн элементарными источниками — атомами, молекулами и т. д., получающими энергию при тепловых

столкновениях между собой, — происходит совершенно беспорядочно; фазы испускаемых ими световых волн с течением времени изменяются также беспорядочно. Поэтому приходится обеспечивать когерентность интерферирующих волн искусственным образом. Это достигается разделением каждой волны, испускаемой элементарным источником, на две волны в одном месте и соединением их (для интерференции) в другом месте. Эти волны от места разделения до места встречи проходят различные пути; для обеспечения когерентности свойства среды на пути лучей не должны со временем изменяться. Этот прием осуществляется перечисленными ниже способами:

1) с помощью зеркал заставляют интерферировать две части одной и той же волны (от данного источника света, в частности от каждого атома). Этот способ показан на рис. IV.13, *а*, *б*; результаты

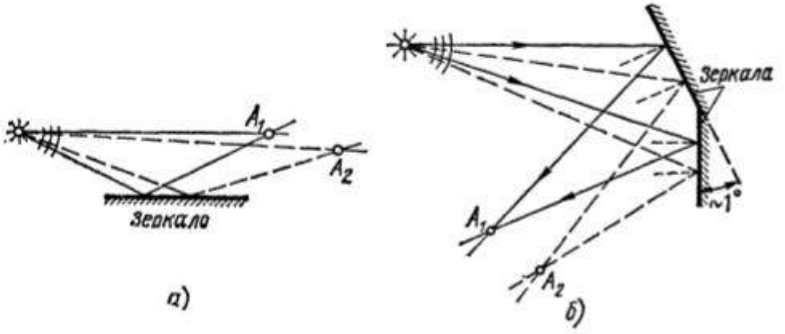


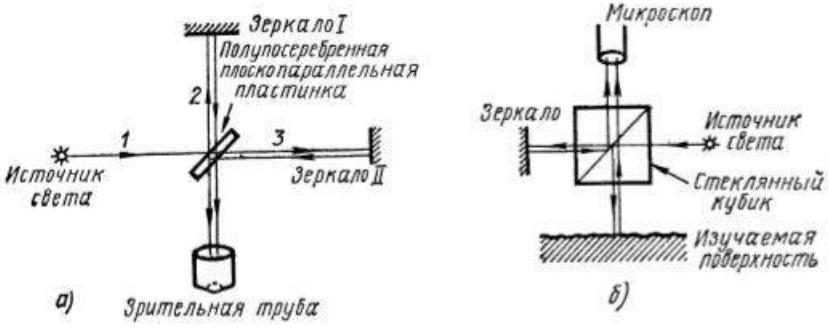
Рис. IV.13

интерференции в точках A_1, A_2, ... определяются разностью фаз между лучами, прошедшими различные пути от источника волны до этих точек (эти лучи показаны на рис. IV.13 сплошными и пунктирными линиями);

2) используется тонкий полупрозрачный слой серебра (нанесенный на стеклянную пластинку), который расщепляет световую волну на две части: отраженную и прошедшую через слой; полученные таким образом когерентные волны при помощи зеркал направляются в прибор, где наблюдается интерференция. Этот способ используется в *интерферометрах Майкельсона* и *Линника*, схемы которых соответственно приведены на рис. IV.14, *а*, *б*.

В интерферометре Майкельсона луч *1* от источника света расщепляется на полупосеребренной поверхности плоскопараллельной пластинки на два луча: *2* и *3*. Первый из них, отразившись от зеркала *I*, частично проходит через пластинку и попадает в зрительную трубу. Луч *3* после отражения от зеркала *II* снова падает на полупосеребренную поверхность и частично отбрасывается в зрительную трубу. Таким образом, в зрительной трубе можно наблюдать интерференцию лучей, разность хода между которыми определяется расположением зеркал относительно пластинки и толщиной этой пластинки.

В интерферометре Линника происходит, по существу, такой же процесс расщепления световой волны источника на две когерентные волны; роль одного из зеркал играет изучаемая поверхность, а полу-



Рйс. IV.14

посеребренной поверхностью является диагональная плоскость кубика (составленного из двух частей);

3) используются тонкие (прозрачные) пластинки или пленки, на поверхности которых интерферируют отраженная и преломленная волны. Этот способ показан на рис. IV.15. Допустим, на прозрачную пластинку, имеющую толщину h и показатель преломления n, падает плоская волна, фронт которой в некоторый момент времени доходит до AD. Пока крайний луч *1* пройдет в пластинке (со скоростью $c = c_0/n$) путь $AB + BC = 2h/\cos\gamma$, затратив на это время $\Delta t = 2h/(c\cos\gamma) = 2hn/(c_0\cos\gamma)$, другой крайний луч *2* пройдет расстояние $DC + CE$ (со скоростью c_0), причем, очевидно, $DC + CE = c_0\Delta t$. Найдем расстояние, на которое второй луч опередит первый. Так как

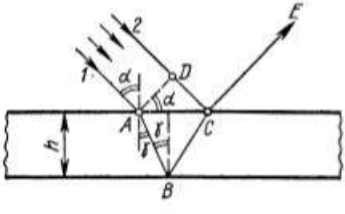


Рис. IV.15

$$DC = AC\sin\alpha; \quad AC = 2h\,\mathrm{tg}\,\gamma,$$

то

$$\Delta = CE = c_0\Delta t - DC = \frac{2hn}{\cos\gamma} - 2h\sin\alpha\,\mathrm{tg}\,\gamma =$$

$$= \frac{2hn}{\cos\gamma} - 2h\frac{\sin\alpha}{\sin\gamma}\frac{\sin^2\gamma}{\cos\gamma} = 2hn\frac{1-\sin^2\gamma}{\cos\gamma} = 2hn\cos\gamma,$$

или, заменяя $\cos\gamma = \sqrt{1-\sin^2\gamma} = \frac{1}{n}\sqrt{n^2-\sin^2\alpha}$, можно получить

$$\Delta = 2hn\cos\gamma = 2h\sqrt{n^2-\sin^2\alpha}.$$

Следует учесть, что при отражении света от границы, по другую сторону которой находится среда с большим показателем преломления, теряется полволны. Если рассматриваемая пластинка имеет

больший показатель преломления, чем среда, откуда идет свет, то потеря $\lambda/2$ происходит у луча *2* в точке C. Окончательно отставание луча *1* от луча *2* или, как говорят, оптическая разность хода этих лучей равна:

$$\Delta = 2hn\cos\gamma - \frac{\lambda}{2} = 2h\sqrt{n^2 - \sin^2\alpha} - \frac{\lambda}{2}. \qquad (1.18)$$

Если $\Delta = (2k - 1)\,\lambda/2$, где $k = 1, 2, 3, \ldots$, то лучи *1* и *2* «гасят» друг друга и в направлении CE отраженного света не наблюдается; если $\Delta = k\lambda$, где k — целое число, то интерферирующие лучи дают в направлении отраженного луча максимум энергии. Число k называется *порядком максимума* или *минимума*.

Однако из формулы (1.18) следует, что при постоянных h, n и α разность фаз лучей *1* и *2* одинакова в любой точке поверхности. Если, например, лучи *1* и *2*, интерферирующие в точке C, друг друга гасят, то при указанных условиях они будут гасить друг друга и в любом месте пластинки (это означает, что световая энергия падающего луча целиком уходит внутрь пластинки и от ее поверхности не отражается). Таким образом, если на однородную ($n = \text{const}$) плоскопараллельную ($h = \text{const}$) пластинку падает световая волна с одинаковым везде углом падения ($\alpha = \text{const}$), то вся поверхность этой пластинки имеет одинаковую светимость (в зависимости от значения разности хода Δ). Но если толщина или показатель преломления различны в различных местах пластинки, то на ее поверхности появляются чередующиеся темные и светлые интерференционные полосы.

Различают:

1) интерференционные *полосы равной толщины*; допустим, что толщина пластинки различна, но показатель преломления и угол падения лучей везде одинаковы. Тогда во всех тех местах пластинки, где толщина h, а следовательно, и разность хода Δ одинаковы, наблюдается один и тот же результат интерференции. Это означает, что вдоль какой-нибудь темной или светлой интерференционной полосы, образующейся на поверхности пластинки, толщина этой пластинки одна и та же.

Интерференционными полосами равной толщины являются, например, полосы, образующиеся на поверхности клина (между линзой и плоской пластинкой они имеют вид колец);

2) интерференционные *полосы равного наклона*; допустим, на плоскопараллельную пластинку ($h = \text{const}$, $n = \text{const}$) падают две световые волны под углами падения α_1 и α_2. Тогда из каждой точки, взятой на поверхности пластинки, будут исходить две отраженные волны (рис. IV.16): одна — от волны *а*, другая — от волны *б*. Разность хода интерферирующих лучей в точке C_1, согласно формуле (1.18), для волны *а* равна Δ_1, для волны *б* — Δ_2; допустим, что эти разности хода таковы, что в точке C_1 две части волны *а* друг друга усиливают, а волны *б* друг друга ослабляют (в частности, гасят). Очевидно, ввиду постоянства толщины пластинки и углов падения α_1 и α_2 из каждой точки C_1, C_2, C_3, ... поверхности пластинки исходит одно и то же излучение, поэтому никаких интерференционных полос на этой поверхности не образуется. Если же воспользоваться линзой, которая собирает все параллельные между собой лучи в одной точке, то можно на экране, расположенном в фокальной плоскости этой линзы, получить светлую точку от волны *а* в одном месте и темную — от волны *б* — в другом месте экрана. Допустим, что на плоскопараллельную пластинку падают световые волны всевозможных направлений; тогда на экране получатся интерференционные полосы, каждая из которых соответствует определенному углу падения α; эти интерференционные полосы называются полосами равного наклона. Их можно наблюдать, например, в интерферометре Майкельсона, если источник света дает расходящийся пучок лучей, т. е. сферическую волну, а плоскости зеркал перпендикулярны.

Из условия максимума (разность хода Δ равна четному числу $\lambda/2$) или минимума (Δ равно нечетному числу $\lambda/2$) следует, что при интерференции немонохроматического излучения максимумы и минимумы для различных длин волн λ получаются в различных местах. Если, например, на клин падает сложный свет, состоящий из некоторого множества монохроматических волн, то на поверхности клина получится окрашенная интерференционная картина; каждая интерференционная полоса соответствует не только определенной толщине, но и определенной длине волны. Возможно, что максимум для одной волны окажется в том месте, где другая волна дает минимум света. Ввиду этого для наблюдения интерференционной картины на пластинках (пленках) от немонохроматического света их толщина не должна быть очень большой. Действительно, допустим, что на пластинку падает световая волна с непрерывным

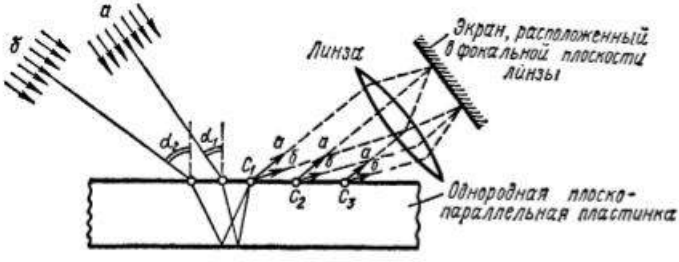


Рис. IV.16

спектром в пределах от λ_1 до λ_2. Очевидно, если на экране или на поверхности пластинки k-й максимум одного края спектра (λ_1) будет совпадать с $(k+1)$-м максимумом другого края спектра (λ_2), то отчетливой интерференционной картины не будет; в этом случае

$$k\lambda_1 = (k+1)\lambda_2; \quad \lambda_1 - \lambda_2 = \frac{\lambda_2}{k}.$$

Поэтому интерференцию можно наблюдать только в том случае, если

$$\lambda_1 - \lambda_2 < \frac{\lambda_2}{k}; \quad k < \frac{\lambda_2}{\Delta\lambda},$$

т. е. при заданной ширине спектра $\Delta\lambda = \lambda_1 - \lambda_2$ на порядок максимума k накладывается определенное ограничение, и, наоборот, если задан порядок максимума k, то этим определится максимально допускаемая ширина интервала $\Delta\lambda$. Порядок же максимума k зависит от толщины пластинки; из выражения (1.18) для k-го максимума имеем условие

$$2hn\cos\gamma - \frac{\lambda}{2} = k\lambda.$$

При больших h порядки максимума k выражаются большими числами, поэтому интерференционную картину можно наблюдать только для очень узкого интервала непрерывного спектра $\Delta\lambda$. Например, если на пластинку падает монохроматический свет от $\lambda_1 = 0{,}55$ до $\lambda_2 = 56$ мкм, то $\lambda_2/\Delta\lambda = 56$ и, для того чтобы k было меньше 56, толщина стеклянной пластинки ($n = 1{,}5$) при нормальном падении лучей ($\cos\gamma = 1$) должна быть менее 8 мкм.

Выше было указано, что если на однородную пластинку постоянной толщины падает волна под определенным углом, то наружная поверхность пластинки (на которую падает свет) будет иметь повсюду одинаковую светимость в зависимости от разности хода Δ лучей, отраженных от обеих границ пластинки. В частности, можно подобрать Δ таким, чтобы наружная поверхность пластинки была темной, т. е. чтобы отражение света от этой поверхности практически отсутствовало. Этим пользуются для «просветления» оптических деталей, т. е. уменьшения коэффициента отражения света от их поверхностей. При помощи специальной обработки на стеклянной поверхности деталей (линз, призм) образуется тонкая (твердая) прозрачная пленка, толщина

и показатель преломления которой подбираются такими, чтобы на наружной границе пленки выполнялось условие гашения $\Delta = \lambda/2, 3\lambda/2, 5\lambda/2$ и т. д. Благодаря этому отражение светового потока от поверхностей деталей будет минимальным и поступающий в оптическую систему свет будет использоваться рационально. «Просветление оптики» особенно важно в системах, содержащих большое число деталей (отражающих поверхностей). Так как условие гашения содержит в себе длину волны, то полное «просветление» можно получить только для очень узкого участка спектра, наиболее важного для данной оптической системы.

Интерференцией света объясняется также различие в коэффициентах зеркального отражения для длинных и коротких волн. Допустим, что поверхность тела имеет шероховатость в виде прямоугольных выступов, высота которых равна h. Та часть светового потока, которая отражается от впадин, будет отставать от другой части, которая отражается от вершин этих выступов, на $\Delta = 2h \cos \alpha$, где α — угол падения лучей. Если высота выступов h велика и угол падения α мал, то в направлении отраженного луча может соблюдаться условие гашения $\Delta = \lambda/2, 3\lambda/2, 5\lambda/2$, тогда отраженный световой поток будет слабым. Очевидно, поверхность может зеркально отразить свет, если высота выступов $h \ll \lambda/2$, и поэтому условие гашения не выполнится для всевозможных углов падения и, в частности, для направления правильного (зеркального) отражения. Заметим, что это условие зависит от длины волны; поверхность, которая зеркальна для длинных волн, будет шероховатой для коротких волн. Кроме того, разность хода лучей Δ зависит также от угла падения α, поэтому при больших углах α можно получить $\Delta \ll \lambda/2$; этим объясняется почти зеркальное отражение света очень шероховатыми поверхностями (бумага, асфальтированное шоссе и др.) при больших углах падения лучей.

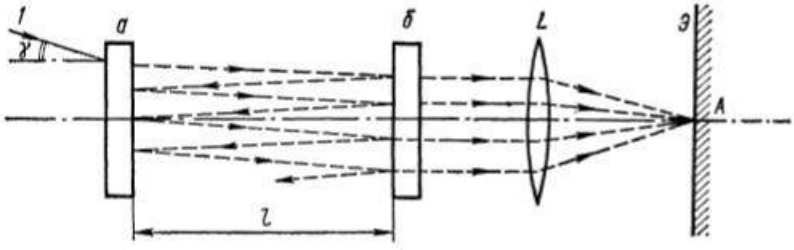


Рис. IV.17

В оптических исследованиях широко применяются *интерферометры Фабри — Перо*, схематическое устройство которых показано на рис. IV.17. Внутренние поверхности пластинок *а* и *б* (стеклянных или кварцевых) тщательно отполированы (до сотых долей длины волны), установлены строго параллельно и покрыты тонким слоем серебра. Допустим, что луч *1* входит в пространство между пластинками под очень малым углом γ к оси прибора. Достигнув пластинки *б*, луч частично отразится к пластинке *а*, частично пройдет через *б* и, пройдя линзу *L*, окажется в точке *A* экрана *Э*. На поверхности пластинки *а* произойдет такое же разделение луча на прошедшую и отраженную составляющие. Вследствие многочисленных отражений между пластинками в точке *A* экрана соберется пучок когерентных между собой частей луча *1*; их интерференция даст в этой точке определенный стабильный результат. Обычно на пластинку *а* извне направляют расходящийся пучок лучей, поэтому на экране получаются чередующиеся интерференционные полосы «равного наклона». Расстояние *l*

между пластинками (доводимое до десятков сантиметров) соответствует толщине h в формуле (1.18), по которой рассчитывается разность хода Δ двух соседних интерферирующих лучей. Поэтому порядок максимума $k = \Delta/\lambda$ в этом интерферометре оказывается очень большим; число интерферирующих в точке A лучей (точнее, отдельных составляющих волны 1, поступившей в прибор) зависит от угла γ и размеров пластинок $а$ и $б$.

Система из двух параллельных пластинок с отражающими плоскими поверхностями получила широкое применение в лазерных устройствах (см. § 18). В отличие от интерферометра источники излучения в лазерах находятся внутри пространства между пластинками; внутренняя поверхность пластинки $а$ посеребрена так, чтобы коэффициент отражения был близок к 1, а пластинки $б$ — к 1/2. При этих условиях из пластинки $б$ выходит почти параллельный пучок лучей, фокусируемый на экран $Э$.

Перечисленные выше приборы были основаны на получении устойчивой интерференционной картины, которая может быть образована только когерентными волнами. Однако понятие когерентности применяется не только для двух волн, интерферирующих в данной точке, но используется также и для характеристики отдельных источников света. Допустим, что излучение с частотой ν (или длиной волны λ), исходящее от данного источника, продолжается τ секунд, после чего наступают паузы. За время τ от источника отойдет цуг волн длиной $l = c\tau$, содержащий N длин волн ($N = l/\lambda$). Для того чтобы от такого источника при помощи перечисленных выше приемов можно было получить устойчивую интерференционную картину, необходимо, чтобы разность хода интерферирующих лучей Δx была значительно меньше l. Величина l называется *длиной когерентности*, а τ — *временем когерентности* данного источника излучения. Заметим также, что ввиду конечности l излучение уже не может полагаться состоящим из одной частоты ν; разлагая это излучение в спектр, мы получим очень узкий интервал частот $\nu \pm \Delta\nu$ (или длин волн $\lambda \pm \Delta\lambda$), причем оказывается, что $\Delta\nu \sim 1/\tau$.

Перечислим важнейшие применения интерференции:

1) измерение длин с очень большой точностью; это позволило дать легко воспроизводимое и достаточно точное определение единицы длины — метра — в зависимости от длины волны оранжевой линии криптона. Интерференционные компараторы позволяют сравнивать размеры до 1 м с точностью до 0,05 мкм; меньшие размеры могут быть измерены с еще большей точностью;

2) изучение и контроль полировки зеркальных поверхностей с точностью до сотых долей длины волны;

3) изучение и контроль однородности веществ (из которых, например, изготовлены оптические детали);

4) определение ряда важнейших величин, характеризующих вещества: коэффициента расширения твердых тел (*дилатометры*), показателя преломления газообразных, жидких и твердых тел (*рефрактометры*) и т. п.;

5) изучение структуры спектров различных веществ и др.

Если на пути световой волны находятся непрозрачные тела или экраны с отверстиями, то за этими телами образуется область тени. Эту область можно очертить геометрически (рис. IV.18, *а*), полагая, что свет распространяется прямолинейно, т. е. световые лучи суть прямые линии. Более детальные наблюдения показывают, что световая волна заходит в область геометрической тени, причем на границе между областями света и тени появляются чередующиеся максимумы и минимумы света, свидетельствующие о некотором перераспределении световой энергии на этой границе. *Огибание световой волной границ непрозрачных тел* с образованием интерференционного перераспределения энергии по различным направлениям называется *дифракцией волны*.

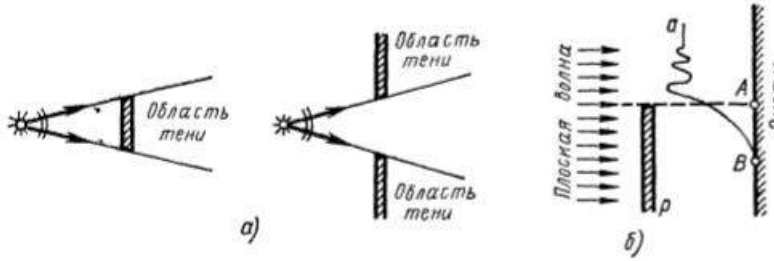


Рис. IV.18

На рис. IV.18, *б* кривая *aB* показывает распределение на экране интенсивности света, испытавшего дифракцию от края пластинки *P*.

Явления дифракции можно объяснить, пользуясь принципом Гюйгенса. Согласно этому принципу, каждую точку фронта световой волны, заданного в некоторый момент времени (например, когда волна дошла до пластинки *P*), можно рассматривать как самостоятельный источник элементарной волны (сферической в однородной и изотропной среде). Огибающая всех этих элементарных волн заходит в область геометрической тени и дает искомое расположение фронта волны в последующие моменты времени. Однако при этом остается открытым вопрос о распределении энергии вдоль фронта волны; если это распределение задано в начале волны, то его необходимо найти и для последующих моментов времени. Эту задачу можно решить; если воспользоваться дополнительным указанием, которое сделал Френель к принципу Гюйгенса, а именно: необходимо найти результат *интерференции элементарных волн* для каждой точки пространства.

Рассмотрим один простой пример; допустим, что через отверстие в непрозрачном экране $\mathcal{Э}_1$ непрерывно проходит плоская монохроматическая световая волна, энергия которой равномерно распределена вдоль фронта волны. Каждую точку фронта волны в сечении $A - B$, согласно принципу Гюйгенса, можно заменить самостоятельным источником элементарной сферической волны. Если построить огибающие этих элементарных волн для различных моментов времени, то оказы-

вается, что фронт волны заходит в область геометрической тени (рис. IV.19, *а*). Необходимо теперь выяснить, какая часть энергии волны $A - B$ заходит в область тени. Для этого, по Френелю, необходимо найти результат интерференции элементарных волн в каждой интересующей нас точке C (в частности, эта точка может лежать на экране, поставленном за отверстием). Обычно за отверстием ставят линзу, которая «сортирует» световые лучи по направлениям; та часть световой волны, которая проходит через отверстие в прямом направлении, соберется линзой в одной точке O (рис. IV.19, *б*). Другие части световой волны, идущие, например, в направлениях AC или AD,

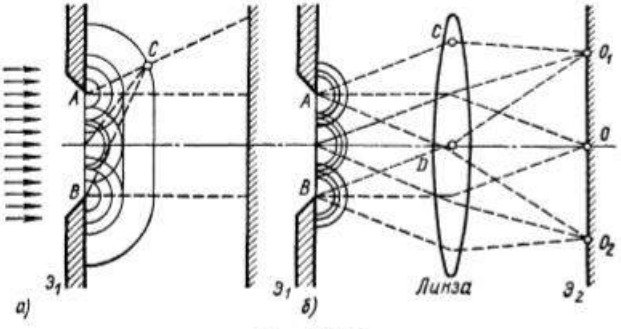


Рис. IV.19

соберутся линзой в других точках O_1, O_2, ... фокальной плоскости, где располагается экран $Э_2$. Таким образом, распределение освещенности на экране, расположенном в фокальной плоскости линзы, показывает, как распределена энергия по различным направлениям у прошедшей через отверстие световой волны.

При определении результатов интерференции (сложения) элементарных волн в каждой интересующей нас точке пространства (например, экрана) необходимо учесть фазы, амплитуды и направления векторов **E** каждой волны. Эта задача очень сложная, и поэтому мы ограничимся рассмотрением простых случаев, где можно воспользоваться элементарными рассуждениями и расчетами.

ДИФРАКЦИЯ ПЛОСКОЙ ВОЛНЫ ОТ ПРЯМОЛИНЕЙНОЙ ТОНКОЙ ЩЕЛИ

При перпендикулярном падении света на плоскость щели все точки фронта волны $A - B$ колеблются в одной фазе. Поэтому лучи, собранные линзой в точке O (рис. IV.20), интерферируя друг с другом, взаимно усиливаются и в этой точке получится максимум освещенности. В другой точке O_1, где соберутся лучи, идущие из разных точек щели под углом φ_1 к основному направлению, результат интерференции будет иной. Следует учесть, что *линза не вносит дополнительной разности хода между лучами*. Это утверждение следует понимать так: параллельные лучи, собранные линзой в какой-нибудь точке, например O или O_1 (см. рис. IV.20, *а*), имеют между собой такую же разность

фаз, какую они имели до линзы в любой плоскости, перпендикулярной этим лучам. Поэтому если в плоскости AB все точки фронта волны колеблются в одной фазе, то исходящие от них лучи, группируясь в точке O, также имеют одинаковые фазы. У лучей, исходящих из щели под углом φ, в точке O_1 будут такие фазы, какие они имеют в перпендикулярной этим лучам плоскости BC.

Обозначим разность хода между крайними лучами пучка, собирающегося в точке O_1, через Δ_1. Очевидно,

$$\Delta_1 = b \sin \varphi_1, \tag{1.19}$$

где b — ширина щели. Допустим, $\Delta_1 = 2\lambda/2$. Тогда рассматриваемый под углом φ_1 пучок лучей (рис. IV.20, *а*) можно разделить на две части, или, как говорят, зоны, причем каждый луч i верхней зоны

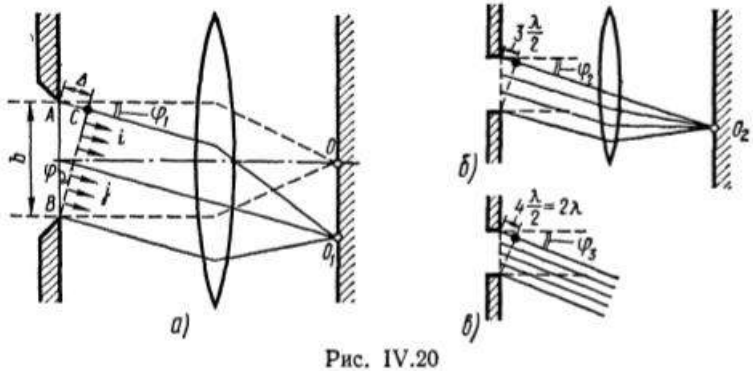


Рис. IV.20

будет на $\lambda/2$ отставать от соответствующего луча j нижней зоны, следовательно, в точке O_1 они «погасят» друг друга. Таким образом, под углом φ_1, удовлетворяющим условию $\Delta_1 = b \sin \varphi_1 = \lambda$, происходит взаимное гашение лучей, прошедших через щель и собранных в точке O_1.

Рассмотрим другое направление — под углом φ_2, для которого $\Delta_2 = b \sin \varphi_2 = 3\lambda/2$. В этом направлении пучок света можно разделить на три зоны (рис. IV.20, *б*), из которых две зоны (первая и вторая или вторая и третья) друг друга погасят, а третья останется непогашенной и даст в соответствующей точке экрана O_2 некоторую освещенность. Очевидно, освещенность в точке O_2 будет значительно меньше, чем освещенность в точке O, так как в точку O_2 приходит непогашенной только одна треть пучка, выходящего из щели в направлении φ_2.

Такое разделение пучка лучей или световой волны на части, которые при интерференции взаимно усиливают или гасят друг друга, называется разделением на *зоны Френеля*.

Продолжая эти рассуждения (например, для $\Delta_3 = b \sin \varphi_3$, рис. IV.20, *в*), можно показать, что в направлениях, в которых $b \sin \varphi$ равно нулю или нечетному числу $\lambda/2$ и, следовательно, пучок света состоит из *нечетного числа* зон Френеля (1, 3, 5 и т. д.), на экране получаются светлые точки (вернее, линии, параллельные щели),

а в направлениях, для которых $b \sin \varphi$ равно четному числу $\lambda/2$ и, следовательно, пучок света делится на *четное число* взаимно гасящих друг друга зон, на экране получаются темные точки (линии). В промежуточных направлениях освещенность экрана, очевидно, должна постепенно изменяться от нуля до соответствующих максимумов (распределение освещенности на экране показано пунктирной линией на рис. IV.21). Так как углы φ очень малы, то $\sin \varphi \approx \varphi$ и условия максимума выполняются под углами

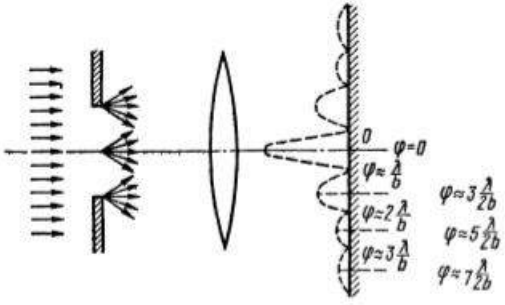


Рис. IV.21

$$\varphi \approx (2k+1)\frac{\lambda}{2b},$$

а условия минимума — под углами

$$\varphi \approx 2k\frac{\lambda}{2b} = k\frac{\lambda}{b}$$

($k = 1, 2, 3, \ldots$). Чем у́же щель, тем дальше друг от друга располагаются максимумы.

Заметим, что *расположение максимумов и минимумов* (в фокальной плоскости линзы) *зависит от длины волны света*. Если падающий свет — сложный, например состоит из двух монохроматических излучений с длинами волн λ_1 и λ_2, то на экране максимумы и минимумы этих волн располагаются в различных местах.

Распределение освещенности на экране (расположенном в фокальной плоскости линзы), полученное вследствие дифракции световой волны, называется *дифракционным спектром*. Рис. IV.21 показывает дифракционный спектр от одной щели, если на щель падает плоская монохроматическая волна.

ДИФРАКЦИЯ ПЛОСКОЙ ВОЛНЫ ОТ НЕСКОЛЬКИХ ЩЕЛЕЙ

Для нахождения дифракционного спектра от двух и более параллельных щелей необходимо учесть не только взаимную интерференцию лучей, вышедших из одной щели, но и интерференцию лучей, прищедших в данную точку экрана из различных щелей.

Согласно рис. IV.19 — IV.21, каждая щель дает на экране ту или иную освещенность по всем направлениям, кроме тех, которые удовлетворяют условиям гашения:

$$b \sin \varphi = \lambda, \quad 2\lambda, \quad 3\lambda, \ldots; \quad \varphi \approx k\frac{\lambda}{b}. \qquad (1.20)$$

В этих направлениях условия гашения (четное число зон) выполняются для каждой щели в отдельности.

Рассмотрим теперь некоторое направление под углом φ и допустим, что первая щель посылает свет, интенсивность которого в соответствующей точке экрана определяется вектором $\mathbf{E}_1$, вторая щель дает

$\mathbf{E}_2$, третья — $\mathbf{E}_3$ и т. д. В точке M, где собираются лучи от всех щелей (рис. IV.22, *а*), освещенность определяется значением суммарного вектора $\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 + \mathbf{E}_3 + \ldots$ Выделим два наиболее простых случая:

а) допустим, что в точке M все векторы $\mathbf{E}_1$, $\mathbf{E}_2$, $\mathbf{E}_3$, ... имеют одинаковое направление. Тогда суммарная напряженность $\mathbf{E}$ будет иметь наибольшее значение. Это возможно, если фазы волн, прищедших из

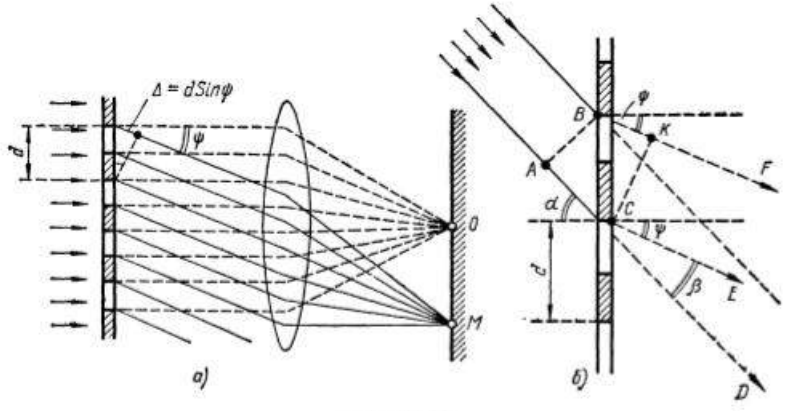


Рис. IV.22

разных щелей, отличаются на 2π или на целое число 2π, т. е. разность хода лучей от соседних щелей ($\Delta = d \sin \psi$) равна λ или целому числу λ. Поэтому условие

$$d \sin \psi = k\lambda \qquad (k = 1, 2, 3, \ldots) \tag{1.21}$$

определяет местонахождение максимумов света (их называют главными максимумами). Здесь d — расстояние между осями двух соседних щелей — называется *периодом решетки*;

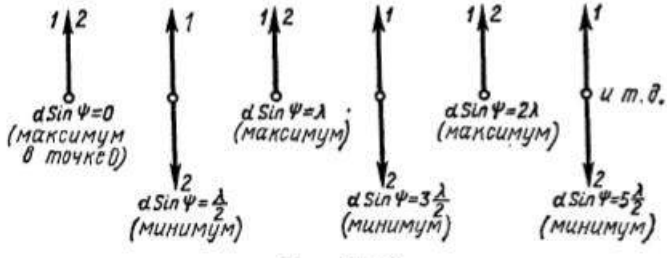


Рис. IV.23

б) допустим, что в данной точке экрана суммарная напряженность $\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 + \mathbf{E}_3 + \ldots = 0$, т. е. волны, пришедшие в эту точку экрана от различных щелей, в результате интерференции гасят друг друга. Это условие может выполняться различным образом в зависимости от числа щелей. Для наглядности воспользуемся векторной

диаграммой сложения колебаний. Допустим, что имеются всего две щели, создающие равные $\mathbf{E}_1$ и $\mathbf{E}_2$. Очевидно, условие $\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 = 0$ выполняется, если $\mathbf{E}_1$ и $\mathbf{E}_2$ противоположны по направлению, т. е. отличаются по фазе на π, 3π, 5π и т. д. Разность хода лучей, исходящих от этих щелей, должна быть равна $d \sin \psi = \lambda/2, 3\lambda/2, 5\lambda/2$ и т. д.

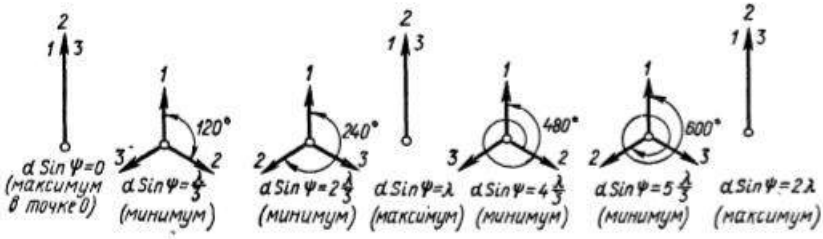


Рис. IV.24

На рис. IV.23 схематически показано образование максимумов и минимумов при различных фазах между интерферирующими лучами *1* и *2*. Заметим, что в случае двух щелей между двумя соседними максимумами имеется один минимум. Для трех щелей направления максимумов по-прежнему определяются условием (1.21), означающим, что все векторы $\mathbf{E}_1$, $\mathbf{E}_2$, $\mathbf{E}_3$ имеют одинаковое направление и складываются; для получения же минимума, в частности для равенства $\mathbf{E}_1 + \mathbf{E}_2 + \mathbf{E}_3 = 0$, необходимо, чтобы фазы этих векторов отличались либо

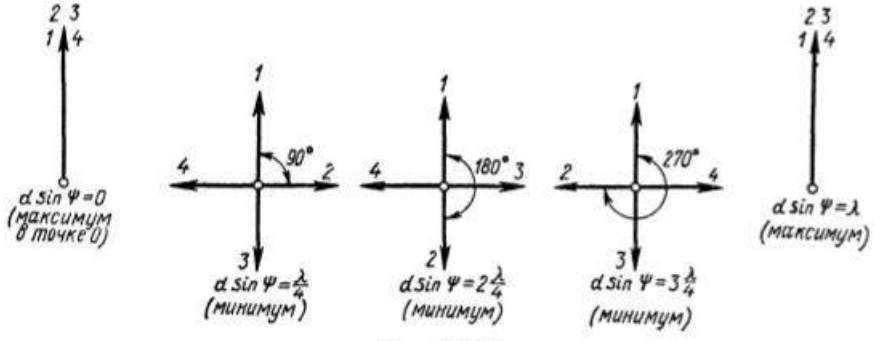


Рис. IV.25

на $2\pi/3$ (120°, или $\lambda/3$), либо на $2 \cdot 2\pi/3$ (240°, или $2\lambda/3$). Приведенная на рис. IV.24 векторная диаграмма иллюстрирует это утверждение. При наличии трех щелей между двумя соседними максимумами появляются два минимума.

Для четырех щелей условия минимума имеют место, если разность фаз между векторами напряженностей от соседних щелей равна $\pi/2$, $2\pi/2$, $3\pi/2$, а разность хода — $\lambda/4$, $2\lambda/4$, $3\lambda/4$ и т. д. Соответствующая векторная диаграмма приведена на рис. IV.25. При наличии четырех щелей между двумя соседними максимумами образуются три минимума.

Продолжая рассуждения для пяти, шести щелей и т. д., можно установить следующее правило: при наличии N щелей между двумя соседними максимумами образуется $N - 1$ минимумов; разность хода лучей $d \sin \psi$ от двух соседних щелей для максимумов должна равняться целому числу λ, а для минимумов — λ/N, $2\lambda/N$, $3\lambda/N$ и т. д. Дифракционный спектр от N щелей имеет вид, показанный на рис. IV.26. Дополнительные максимумы, расположенные между двумя соседними минимумами, создают на экране весьма слабую освещенность (фон).

Основная часть энергии световой волны, прошедшей через дифракционную решетку, перераспределяется между *главными максимумами*, образующимися в направлениях $d \sin \psi = k\lambda$, где $k = 0, 1, 2, 3, \ldots$ называется «порядком» максимума.

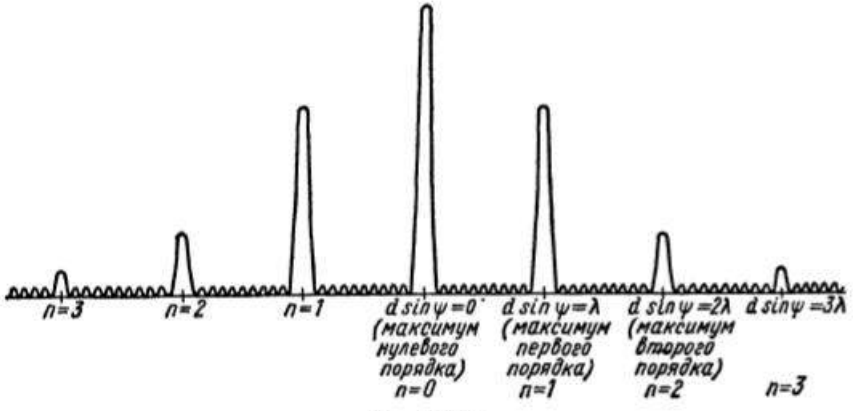


Рис. IV.26

Очевидно, чем больше число щелей N, тем большее количество световой энергии пройдет через решетку, тем больше минимумов образуется между соседними главными максимумами, тем, следовательно, более интенсивными и более острыми будут максимумы.

Если свет, падающий на дифракционную решетку, состоит из двух монохроматических излучений с длинами волн λ_1 и λ_2, то их главные максимумы расположатся в различных местах экрана. Для очень близких друг к другу длин волн (одноцветные излучения) максимумы на экране могут получиться настолько близко друг к другу, что сольются в одну общую светлую полосу (рис. IV.27, *б*). Если же вершина одного максимума совпадает (*в*) или находится дальше (*а*) ближайшего минимума второй волны, то по распределению освещенности на экране можно уверенно установить наличие двух волн (или, как говорят, «разрешить» эти волны).

Выведем *условие разрешимости* двух волн: k-й максимум (т. е. максимум k-го порядка) волны λ_1 получится, согласно формуле (1.21), под углом ψ, удовлетворяющим условию $d \sin \psi = k\lambda_1$. Предельное условие разрешимости требует, чтобы под этим же углом получился

минимум волны λ_2, ближайшей к его k-му максимуму (рис. IV.27, *в*). Согласно сказанному выше, для получения ближайшего минимума к разности хода $k\lambda_2$ следует прибавить дополнительно λ_2/N. Таким образом, условие совпадения углов ψ, под которыми получаются максимум $\lambda_1 (d \sin \psi = k\lambda_1)$ и минимум λ_2 $(d \sin \psi = k\lambda_2 + \lambda_2/N)$, приводит к соотношению

$$k\lambda_1 = k\lambda_2 + \frac{\lambda_2}{N}; \quad \frac{\lambda_2}{\lambda_1 - \lambda_2} = kN. \tag{1.22}$$

Если $\lambda_2/(\lambda_1 - \lambda_2)$ больше, чем произведение числа щелей N, на порядок спектра k, то максимумы λ_1 и λ_2 не будут разрешаться. Очевидно, если два максимума не разрешаются в спектре k-го порядка, то они могут быть разрешены в спектре более высоких порядков. Согласно выражению (1.22), чем больше число N интерферирующих между собой пучков и чем больше разность хода Δ между ними $(k = \Delta/\lambda)$, тем более близкие волны λ_1 и λ_2 могут быть разрешены.

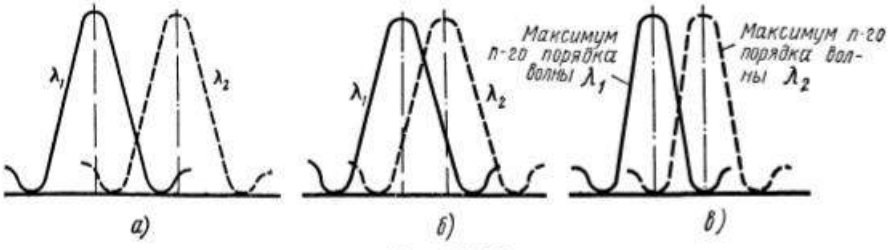


Рис. IV.27

У дифракционной решетки N, т. е. число щелей, велико, но порядок спектра $k = d \sin \psi/\lambda$, который можно использовать для измерительных целей, мал; у интерферометра Майкельсона, наоборот, число интерферирующих пучков N равно двум, но разность хода между ними, зависящая от расстояний до зеркал *I* и *II* (см. рис. IV.14), велика, поэтому порядок k наблюдаемого спектра измеряется очень большими числами.

Угловое расстояние между двумя соседними максимумами двух близких волн зависит от порядка спектра k и периода решетки d:

$$\left.\begin{aligned} d \sin \psi_1 &= k\lambda_1; \quad \psi_1 \approx k\frac{\lambda_1}{d}; \\ d \sin \psi_2 &= k\lambda_2; \quad \psi_2 \approx k\frac{\lambda_2}{d}; \end{aligned}\right| \quad \psi_1 - \psi_2 \approx k\frac{\lambda_1 - \lambda_2}{d}. \tag{1.23}$$

Период решетки можно заменить на число щелей n_1, приходящихся на единицу длины решетки: $d = 1/n_1$.

Выше предполагалось, что лучи, падающие на дифракционную решетку, перпендикулярны ее плоскости. При наклонном падении лучей (см. рис. IV.22, *б*) нулевой максимум будет смещен и получится в направлении ACD. Допустим, что максимум k-го порядка получается в направлении ACE, т. е. разность хода лучей ACE и BF равна $k\lambda$. Тогда $AC - BK = d \sin \alpha - d \sin \psi = k\lambda$. Так как при малых k углы

α и ψ близки друг к другу по величине, то

$\sin\alpha - \sin\psi = 2\cos\frac{\alpha+\psi}{2}\sin\frac{\alpha-\psi}{2} \approx 2\frac{\alpha-\psi}{2}\cos\alpha = \beta\cos\alpha$, следовательно,

$$d\beta\cos\alpha \approx k\lambda;$$

$$\beta \approx k\lambda/d\cos\alpha, \qquad (1.24)$$

где $\beta = \alpha - \psi$ есть угловое отклонение k-го максимума от нулевого. Сравним эту формулу с выражением (1.21), которую запишем в виде $\psi \approx k\lambda/d$; так как $\cos\alpha < 1$, то угловое отклонение при наклонном падении оказывается больше, чем при перпендикулярном падении лучей. Это соответствует уменьшению периода решетки в $\cos\alpha$ раз. Следовательно, при больших углах падения α можно получить дифракционные спектры от коротковолнового (например, рентгеновского) излучения и измерить их длины волн.

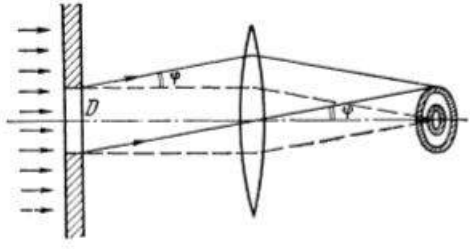


Рис. IV.28

Если плоская световая волна проходит не через щели, а через круглые отверстия малого диаметра D (рис. IV.28), то дифракционный спектр (на плоском экране, расположенном в фокальной плоскости линзы) представляет собой систему чередующихся темных и светлых колец. Первое темное кольцо получается под углом φ_1, удовлетворяющим условию

$$\sin\varphi_1 \approx \varphi_1 = 1{,}22\lambda/D. \qquad (1.25)$$

У второго темного кольца $\varphi_2 = 2{,}232\lambda/D$. На долю центрального светлого круга, называемого *пятном Эйри*, приходится около 85% всей мощности излучения, прошедшей через отверстие и линзу; остальные 15% распределяются между светлыми кольцами, окружающими это пятно. Размеры пятна Эйри зависят от λ, D и фокусного расстояния линзы.

Дифракционные решетки, которые рассматривались выше, состояли из чередующихся «щелей», полностью п р о п у с к а ю щ и х световую волну, и «непрозрачных полосок», которые полностью п о г л о щ а ю т или о т р а ж а ю т падающее на них излучение. Можно сказать, что в таких решетках коэффициент пропускания световой волны имеет только два значения: на протяжении щели он равен единице, а на протяжении непрозрачной полоски — нулю. Поэтому на границе между щелью и полоской коэффициент пропускания скачкообразно изменяется от единицы до нуля.

Однако можно изготовить дифракционные решетки и с другим распределением коэффициента пропускания. Например, если на прозрачную пластинку (или пленку) нанести поглощающий слой с периодически изменяющейся толщиной, то вместо чередования совершенно

прозрачных щелей и совершенно непрозрачных полосок можно получить дифракционную решетку с *плавным* изменением коэффициента пропускания (в направлении, перпендикулярном щелям или полоскам). Особый интерес представляют решетки, у которых коэффициент пропускания изменяется по *синусоидальному* закону. Дифракционный спектр таких решеток состоит не из множества максимумов (как это показано для обычных решеток на рис. IV.26), а только из центрального максимума ($k = 0$) и двух симметрично расположенных максимумов первого порядка ($k = \pm 1$).

Для сферической волны можно изготовить дифракционные решетки, состоящие из множества концентрических кольцевых щелей, разделенных непрозрачными кольцами. Можно, например, на стеклянную пластинку (или на прозрачную пленку) нанести тушью концентрические кольца; при этом центральный круг, охватывающий центр этих колец, может быть либо прозрачным, либо затушеванным. Такие дифракционные решетки называются «зонными пластинками» или решетками. У дифракционных решеток, состоящих из прямолинейных щелей и полосок, для получения отчетливой интерференционной картины было необходимо *постоянство* ширины щели и периода решетки; у зонных пластинок для этой цели должны быть рассчитаны необходимые радиусы и толщины колец. Зонные решетки также могут быть изготовлены с плавным, например синусоидальным, изменением коэффициента пропускания вдоль радиуса.

§ 6. РАСПРОСТРАНЕНИЕ СВЕТА В АНИЗОТРОПНОЙ СРЕДЕ, ПОЛУЧЕНИЕ И ПРИМЕНЕНИЕ ПОЛЯРИЗОВАННОГО СВЕТА

Среда называется *оптически анизотропной*, если ее оптические свойства (скорость распространения света или показатели преломления) различны в различных направлениях. Допустим, что на грань прозрачного кристалла (кварц, исландский шпат и т. д.) падает плоскополяризованный луч, у которого вектор **E** колеблется под некоторым углом к плоскости падения (плоскости чертежа; рис. IV.29). Зная скорость света в первой среде c_0 и измеряя углы падения α и преломления γ, можно найти скорость света во второй среде:

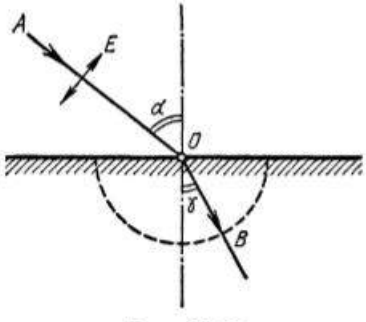


Рис. IV.29

$$\frac{\sin\alpha}{\sin\gamma} = \frac{c_0}{c}; \quad c = c_0 \frac{\sin\gamma}{\sin\alpha}. \quad (1.26)$$

Измерения показывают, что скорость света в кристаллах различна не только в зависимости от направления распространения света, т. е. от углов γ, но и от ориентации вектора **E** относительно плоскости падения. Однако в кристаллах существует одно или несколько направлений, вдоль которых скорость света не зависит от ориентации вектора **E**. Эти направления называются *оптическими осями кристалла*. Следует

иметь в виду, что оптическая ось — это не одна какая-то линия в кристалле наподобие оси симметрии, а определенное направление в кристалле; все прямые, параллельные этому направлению и взятые в любом месте кристалла, являются оптическими осями. Если кристалл имеет одну оптическую ось, то его называют *одноосным*, если две оси — *двухосным* и т. д. Так как вектор **E** перпендикулярен своему лучу, то при распространении света вдоль оптической оси вектор **E** при всех его различных ориентациях в пространстве всегда перпендикулярен также и к оптической оси.

ПОЛЯРИЗОВАННЫЙ СВЕТ

Назовем плоскость, проходящую через данный луч и оптическую ось кристалла, *главной плоскостью*. При распространении света в кристаллах различают:

1) *обыкновенные лучи*, у которых вектор **E** ориентирован перпендикулярно главной плоскости (следовательно, перпендикулярен и оптической оси);



Рис. IV.30

2) *необыкновенные лучи*, у которых вектор **E** лежит в главной плоскости (следовательно, образует с оптической осью некоторые углы).

Обыкновенные лучи распространяются по всем направлениям в кристалле с одной и той же скоростью, которую обозначим c_o. Необыкновенные лучи распространяются в кристалле с различными скоростями в зависимости от угла между вектором **E** и оптической осью; эту скорость обозначим c_e. Очевидно, вдоль оптической оси все лучи являются обыкновенными и в этом направлении существует только одна скорость распространения света.

Допустим, что в некоторый момент времени внутри анизотропной среды точечный источник света S начинает испускать световую волну (рис. IV.30).

В каждом из направлений распространения волны разложим вектор **E** на две составляющие: $\mathbf{E}_o$ — перпендикулярную плоскости, в которой лежат направление луча и оптическая ось, и $\mathbf{E}_e$ — лежащую в этой плоскости. Так как обыкновенная и необыкновенная составляющие

вектора **E** распространяются в данной среде с различными скоростями, то испускаемая источником волна разделится на две волны. Обыкновенная волна по всем направлениям распространяется с одной и той же скоростью c_o и ее фронт волны (или волновая поверхность) имеет форму сферы. Необыкновенная волна распространяется вдоль оптической оси с такой же скоростью, что и обыкновенная, но в других направлениях скорость ее c_e будет зависеть от угла между $\mathbf{E}_e$ и оптической осью. Фронт необыкновенной волны имеет форму эллипсоида (у одноосных кристаллов — эллипсоида вращения); наибольшее расхождение волновых поверхностей обыкновенной и необыкновенной волн будет в направлении, перпендикулярном оптической оси среды. Если все c_e меньше c_o, т. е. эллипсоид необыкновенной волны лежит

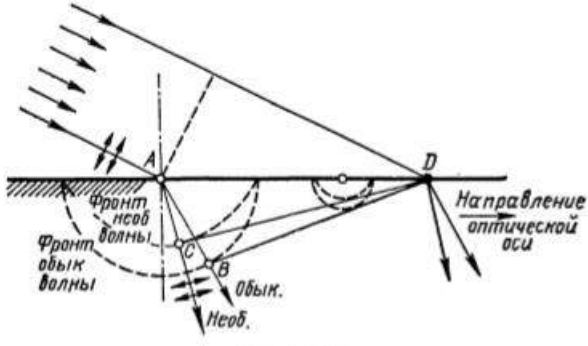


Рис. IV.31

внутри сферы обыкновенной волны, то среду (кристалл) называют *положительной*; если же $c_e > c_o$, то среда называется *отрицательной* (рис. IV.30, *а*, *б*).

Рассмотрим преломление плоской волны на границе анизотропной среды, например положительной. Для простоты предположим, что направление оптической оси параллельно этой границе (рис. IV.31). Применяя принцип Гюйгенса, проведем огибающие элементарных сферических фронтов обыкновенной волны BD и элементарных эллипсоидальных фронтов необыкновенной волны CD. Таким образом, *при преломлении плоской волны на границе анизотропной среды появляются две плоские же волны, распространяющиеся в различных направлениях и с различными скоростями*. Можно сказать, что каждый падающий луч распадается на два луча — обыкновенный и необыкновенный. Обыкновенный луч распространяется в направлении AB, причем $AB \perp BD$; вектор **E** у этого луча колеблется в плоскости, перпендикулярной плоскости чертежа (которая является главной плоскостью). Необыкновенный луч распространяется в направлении AC, причем AC не перпендикулярно CD (C — точка касания огибающей плоскости и фронта эллипсоидальной волны, вышедшей из A); вектор **E** у необыкновенного луча лежит в плоскости чертежа и, следовательно, составляет с оптической осью угол, отличный от прямого. Заметим, что в наших рассуждениях предполагается наличие в падаю-

щей волне колебаний вектора **E** как в плоскости падения, так и в перпендикулярной плоскости. Если, например, падающая волна является плоскополяризованной и имеет вектор **E**, колеблющийся только в плоскости падения (либо только перпендикулярно ей), то при преломлении появится только одна волна — необыкновенная (или во втором случае — обыкновенная). Обе эти волны наблюдаются только в том случае, если падающий свет либо естественный, либо же имеет вектор **E**, колеблющийся под углом к плоскости падения, отличным от нуля или 90°.

При помощи аналогичных построений можно определить ход лучей и фронты преломленных волн на границе отрицательных кристаллов, а также при других ориентациях оптической оси относительно

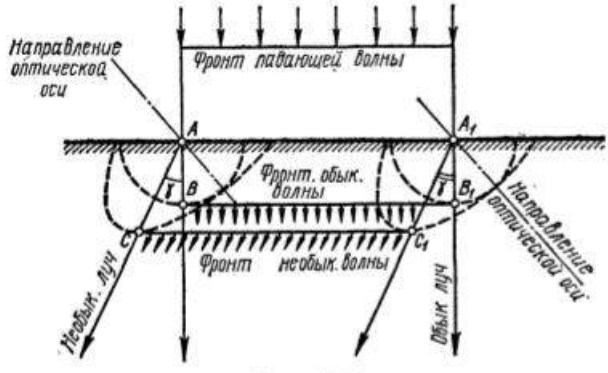


Рис. IV.32

этой границы и при различных углах падения. В качестве примера на рис. IV.32 показано появление обыкновенной и необыкновенной волн при перпендикулярном падении света на границу отрицательной среды; из этого примера видно, что *для необыкновенной волны обычные законы преломления не соблюдаются*. Действительно, по первому закону преломления, $\sin\alpha/\sin\gamma = \text{const}$, поэтому при $\alpha = 0$ должно быть $\gamma = 0$. Это имеет место для обыкновенного луча и не соблюдается для необыкновенного луча. Кроме того, если оптическая ось не лежит в плоскости падения, то необыкновенный луч также выйдет из плоскости падения, следовательно, для него не соблюдается и второй закон преломления (луч падающий, луч преломленный и перпендикуляр в точке падения должны были бы лежать в одной плоскости). Заметим еще одну особенность распространения света в анизотропных средах: направление распространения необыкновенной волны не перпендикулярно ее фронту.

СПОСОБЫ ПОЛУЧЕНИЯ ПОЛЯРИЗОВАННОГО СВЕТА

Перейдем к рассмотрению способов получения поляризованного света. Свет, испускаемый различными источниками, в частности раскаленными твердыми телами или светящимися газами,

обычно естественный. Это объясняется тем, что элементарные источники света — атомы и молекулы — движутся беспорядочно и испускаемые ими световые волны имеют всевозможные направления колебаний вектора **E**. Поэтому необходимо найти способы выделения поляризованного света из естественного. Перечислим важнейшие из них.

1. *Поляризация при отражении и преломлении.* Если естественный свет (рис. IV.33, *а*) падает на отражающую поверхность диэлектрика (стекла, слюды и т. п.) под углом α, удовлетворяющим условию Брюстера:

$$\operatorname{tg}\alpha = n, \tag{1.27}$$

то отраженная волна оказывается плоскополяризованной (n — показатель преломления). При соблюдении условия (1.27) отраженный

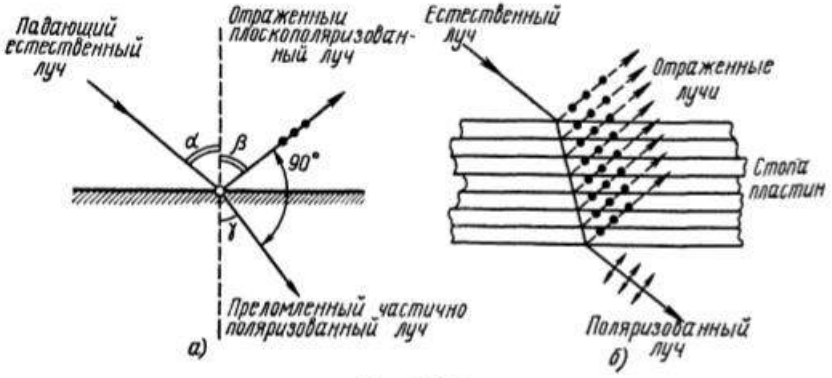


Рис. IV.33

луч оказывается перпендикулярным преломленному лучу ($\sin\alpha/\sin\gamma = n = \operatorname{tg}\alpha$; $\sin\gamma = \cos\alpha$). У отраженной волны вектор **E** перпендикулярен плоскости падения, поэтому в преломленной (прошедшей во вторую среду) волне энергия колебаний в плоскости падения будет больше, чем в перпендикулярной плоскости, и волна частично поляризована.

Недостатком поляризации при отражении является малая доля отражаемого от диэлектриков излучения (например, от стеклянной пластинки отражается 3—5% падающего света). Поэтому пользуются многократным отражением волны от «стопы пластин» (рис. IV.33, *б*); отраженные лучи уносят колебания, перпендикулярные плоскости падения, и проходящий луч, постепенно «очищаясь» от этих колебаний, становится почти плоскополяризованным (с вектором **E**, лежащим в плоскости падения).

2. *Поляризация при двойном лучепреломлении в кристаллах; призма Николя.* При преломлении света на границе оптически анизотропных сред, например кристаллов, естественный луч расщепляется на два луча (обыкновенный и необыкновенный), поляризованные в двух взаимно перпендикулярных плоскостях (рис. IV.34, *а*).

Обыкновенный и необыкновенный лучи имеют в кристалле различные скорости распространения, следовательно, различные показатели преломления n_o и n_e; этим объясняется двойное лучепреломление в точке падения волны на грань призмы: при одном и том же угле падения имеются два угла преломления: γ_o и γ_e ($\sin\alpha/\sin\gamma_o = n_o$; $\sin\alpha/\sin\gamma_e = n_e$). Однако поляризованные лучи выходят из кристалла под очень малым углом друг к другу, что затрудняет их раздельное использование. Чтобы «развести» эти лучи, пользуются различными «поляризующими призмами». Наиболее распространенной является *призма Николя*; из кристалла исландского шпата вырезаются две призмы (рис. IV.34, *б*), которые склеиваются канадским бальзамом. Показатель преломления этого клея ($n = 1{,}550$) лежит между показателями преломления исландского шпата для обыкновенного ($n_o = 1{,}658$) и необыкновенного (n_e) лучей (значение n_e зависит от угла между

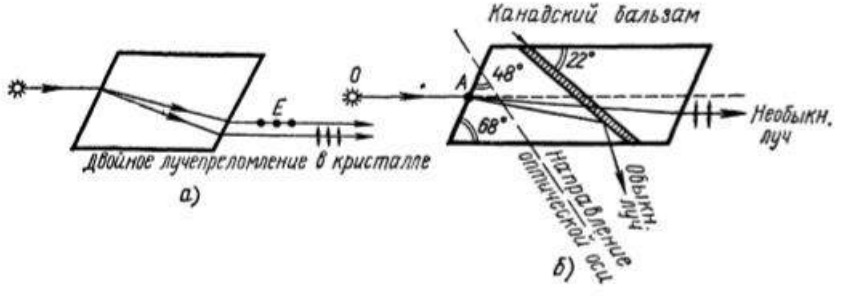


Рис. IV.34

лучом и оптической осью кристалла; его минимальное значение 1,486). Углы в призме подобраны так, чтобы обыкновенный луч на поверхности канадского бальзама испытал полное внутреннее отражение. При помощи этой призмы естественная световая волна разделяется на две плоскополяризованные волны, содержащие почти по 50% падающей энергии (потери в призме невелики). Допустим теперь, что на призму падает плоскополяризованная волна. На рис. IV.35 изображена входная грань призмы; луч падает в точку O перпендикулярно плоскости чертежа. Вектор $\mathbf{E}$ падающей волны следует разложить на две составляющие: $\mathbf{E}_o$ и $\mathbf{E}_e$. Вектор $\mathbf{E}_o$ перпендикулярен оптической оси и соответствующий ей луч — обыкновенный; вектор $\mathbf{E}_e$ лежит в главной плоскости и принадлежит необыкновенному лучу. Очевидно, если $\alpha = 0$, то $\mathbf{E}_e = \mathbf{E}$, $\mathbf{E}_o = 0$ и падающий луч является для призмы необыкновенным; он пройдет без двойного лучепреломления. При $\alpha = 90°$ падающий луч является для призмы обыкновенным и отразится от границы с канадским бальзамом; в этом случае через призму в прямом направлении свет не проходит. Если же α имеет промежуточные значения, то через поляризатор проходит только составляющая $E_e = E\cos\alpha$. Так как энергия электромагнитной волны пропорциональна квадрату электрического вектора, то $W_e/W = E_e^2/E^2 = \cos^2\alpha$ и поэтому прошедший через поляризатор необыкно-

венный луч содержит энергию

$$W_e = W\cos^2\alpha, \tag{1.28}$$

где W — энергия волны, поступающей в поляризатор. Следовательно, если плоскополяризованный свет пропускается через поляризатор, то прошедшая через него энергия пропорциональна квадрату косинуса угла между вектором E поступающего излучения и оптической осью поляризатора (з а к о н М а л ю с а).

Для некоторых целей используются поставленные один за другим два николя. Первый николь является поляризатором (выделяет плоскополяризованный свет из естественного); оптическую ось второго николя располагают под углом α к оптической оси первого николя. Изменяя этот угол от 0 до 90°, можно наблюдать изменение интенсивности света, вышедшего из второго николя («анализатора») по закону (1.28).

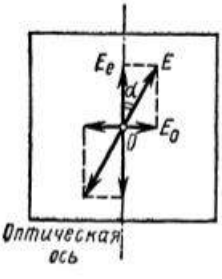


Рис. IV.35

3. *Поляризация при прохождении света через поглощающие анизотропные вещества; полароиды.* Некоторые кристаллические вещества (турмалин, герапатит, т. е. сернокислый иод-хинин, и др.) обладают различным поглощением для лучей с различными ориентациями вектора **E** относительно осей этих кристаллов. Например, турмалиновая пластинка толщиной около 1 мм или чешуйка герапатита толщиной около 0,1 мм почти полностью поглощают обыкновенные лучи (у которых, как указывалось выше, вектор $\mathbf{E}_o$ перпендикулярен оптической оси); необыкновенные же лучи частично поглощаются, частично выходят из пластинки. Если на такую пластинку падает естественный свет, то из пластинки выходит только необыкновенный плоскополяризованный луч. Так же действуют так называемые *полароиды* — целлулоидные пленки, содержащие определенным образом ориентированные мелкие кристаллики герапатита. Следует иметь в виду, что эти вещества обладают селективным (избирательным) поглощением по отношению к различным длинам волн, т. е. коэффициент поглощения их зависит от длины волны. Поэтому если на такие вещества подается не монохроматический, а, например, белый свет, то вышедший из них свет получается окрашенным, причем эта окраска оказывается различной в различных направлениях («дихроизм»).

ИНТЕРФЕРЕНЦИЯ ПОЛЯРИЗОВАННОГО СВЕТА

Интерференция поляризованных лучей имеет некоторые особенности по сравнению с интерференцией естественных лучей. Например, для взаимного гашения двух монохроматических когерентных плоскополяризованных волн кроме равенства амплитуд векторов $\mathbf{E}_1$ и $\mathbf{E}_2$ и наличия разности фаз π, 3π, 5π и т. д. необходимо одинаковое направление колебаний векторов интерферирующих лучей; в противном случае суммарный вектор $\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2$ не будет равен нулю.

Рассмотрим некоторые частные случаи интерференции плоскополяризованных волн. Для получения стабильной интерференционной картины необходимо, чтобы когерентность лучей была обеспечена во всех случаях. Кроме того, предположим, что лучи — монохроматические. Если векторы $\mathbf{E}_1$ и $\mathbf{E}_2$ колеблются в одной и той же плоскости, то в этом случае сохраняются все рассуждения, изложенные в § 5 для естественного света. Рассмотрим теперь более общий случай — интерференцию лучей, у которых векторы $\mathbf{E}$ колеблются в разных плоскостях, в частности во взаимно перпендикулярных плоскостях. Предположим, что интерферирующие волны распространяются в одном и том же направлении; ограничимся двумя частными случаями:

а) во всех точках луча разность фаз векторов $\mathbf{E}_1$ и $\mathbf{E}_2$ равна нулю, т. е. эти векторы одновременно достигают нуля и экстремальных

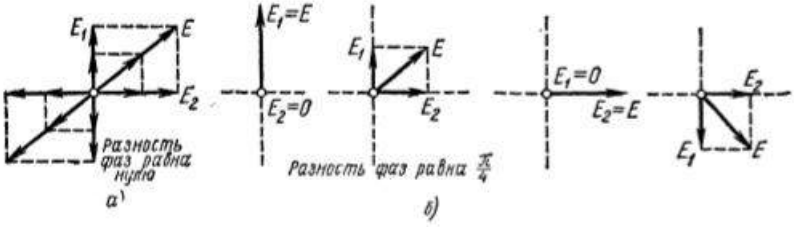


Рис. IV.36

значений (рис. IV.36, *а*). В результате интерференции таких волн получается плоскополяризованная волна, но с иной ориентацией плоскости колебаний суммарного вектора $\mathbf{E}$;

б) фазы векторов $\mathbf{E}_1$ и $\mathbf{E}_2$ интерферирующих волн отличаются на φ, поэтому эти векторы достигают нулевых и экстремальных значений в разное время. На рис. IV.36, *б* показаны ориентации и значения суммарного вектора $\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2$ в разные моменты времени при $\varphi = \pi/2$. В этом случае суммарный вектор $\mathbf{E}$ вращается вдоль луча, сохраняя свое значение при $\mathbf{E}_1 = \mathbf{E}_2$ или меняя его, если $\mathbf{E}_1 \neq \mathbf{E}_2$. Таким образом, если при интерференции двух плоскопараллельных лучей с перпендикулярными плоскостями колебаний разность фаз между векторами напряженности равна нечетному числу $\pi/2$, то результирующий луч поляризуется по кругу при $E_1 = E_2$ или по эллипсу при $E_1 \neq E_2$.

Такую интерференцию можно получить, пропуская плоскополяризованную волну через кристаллическую пластинку соответствующей толщины, вырезанную параллельно оптической оси (рис. IV.37). Действительно, как было указано выше, вектор $\mathbf{E}$ этой волны разлагается в кристалле на составляющие $\mathbf{E}_o$ — обыкновенной и $\mathbf{E}_e$ — необыкновенной волн, которые распространяются в кристалле с различными скоростями c_o и c_e. Если нам необходимо получить разность фаз между $\mathbf{E}_o$ и $\mathbf{E}_e$ по выходе из пластинки, равную $\pi/2$, то нужно подобрать такую толщину d этой пластинки, чтобы один луч вышел раньше (или

позже) другого луча на четверть периода T, т. е.

$$\frac{d}{c_o} - \frac{d}{c_e} = \frac{T}{4}.$$

Умножив это равенство на скорость света в воздухе c и обозначив $c/c_o = n_o$, $c/c_e = n_e$; $Tc = \lambda$, получим формулу, по которой можно рассчитать толщину пластинки d:

$$d(n_o - n_e) = \frac{\lambda}{4}; \quad d = \frac{\lambda}{4(n_o - n_e)}. \tag{1.29}$$

Кристаллическая пластинка, удовлетворяющая этим условиям, называется «пластинкой в четверть волны». Она преобразует плоскопо-

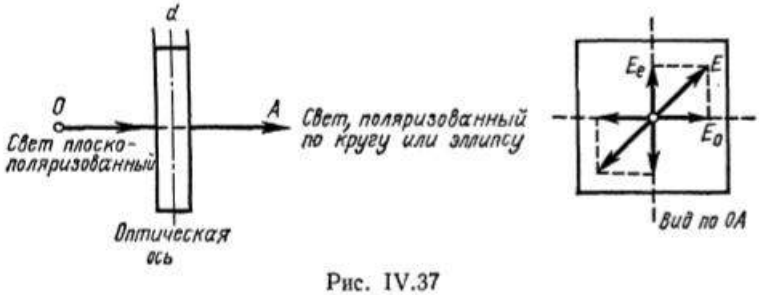


Рис. IV.37

ляризованный свет в свет, поляризованный по кругу (если $\alpha = 45°$ и, следовательно, $E_o = E_e$) или по эллипсу ($\alpha \neq 45°$, $E_o \neq E_e$). Легко доказать, что, обратно, такая пластинка превращает луч, поляризованный по кругу или эллипсу, в плоскополяризованный луч. Очевидно, при помощи такой пластинки можно установить, поляризовано ли данное излучение по кругу (или эллипсу) или же оно является естественным.

ОПТИЧЕСКАЯ АНИЗОТРОПИЯ

Выше упоминались естественные анизотропные среды — кристаллы. Существуют различные способы получения *искусственной анизотропии*, т. е. сообщения оптической анизотропии естественно-изотропным веществам. Например, оптически изотропные вещества (стекло, жидкости) можно превратить в анизотропные, подвергая их односторонней деформации или помещая в электрическое поле. Стеклянный кубик, подвергнутый одностороннему сжатию (или растяжению), становится анизотропным; направление деформации выделяется среди других направлений. Анизотропию жидкостей наблюдают, помещая их в электрическое поле (эффект Керра). Молекулы жидкости, обладающие электрическим моментом (или приобретающие эти моменты под растягивающим действием внешнего поля), ориентируются вдоль поля, поэтому электрические, а следовательно, и оптические свойства жидкости становятся различными вдоль направления поля и перпендику-

лярно ему. Жидкость уподобляется одноосной оптически анизотропной среде.

На рис. IV.38 изображена схема установки для изучения искусственной анизотропии, вызванной электрическим полем; при изменении разности потенциалов $\varphi_1 - \varphi_2$, приложенной к электродам, изменяется степень анизотропии вещества, а следовательно, и интенсивность света, прошедшего через поляризатор и анализатор.

Если убрать внешнее поле, то тепловое движение нарушает упорядоченное расположение электрических осей молекул и жидкость вновь делается изотропной. Это исчезновение анизотропии происходит примерно за 10^{-10} с, поэтому жидкость почти без инерции реагирует на все изменения внешнего поля.

Степень анизотропии, вызванной деформацией (сжатием или растяжением) или электрическим полем, можно оценить по разности

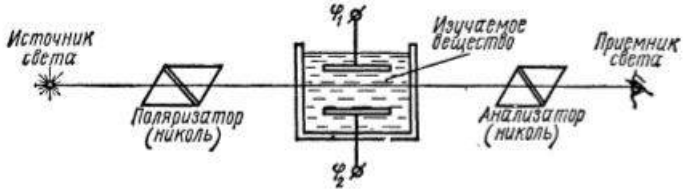


Рис. IV.38

показателей преломления вещества для обыкновенных n_o и необыкновенных лучей n_e (n_e берется для луча, у которого вектор **E** колеблется вдоль направления деформации или внешнего поля, играющего роль оптической оси). Измерения показывают, что для твердых и жидких тел

$$n_o - n_e = K_1\sigma; \quad n_o - n_e = K_2 E^2_{\text{внеш}}, \tag{1.30}$$

где σ — нормальное механическое напряжение; $E_{\text{внеш}}$ — напряженность электрического поля, в котором находится вещество; K_1 и K_2 — постоянные, характеризующие вещество.

Некоторые вещества (*оптически активные*) обладают свойством поворачивать плоскость колебаний вектора **E** вдоль луча. Если световая волна с определенной ориентацией плоскости колебаний (в которой лежат вектор **E** и направление луча) пройдет через такое вещество расстояние l, то плоскость колебаний поворачивается на угол φ, равный

$$\varphi = \alpha l, \tag{1.31}$$

где α — коэффициент, характеризующий вещество. Наблюдения показали, что величина α различна для различных длин волн, поэтому, если проходящий свет — сложный, то различные его монохроматические составляющие повернут свои плоскости колебаний на различные углы (*вращательная дисперсия*). У кварца для $\lambda = 0{,}5893$ мкм коэффициент $\alpha = 21{,}7°$ на миллиметр.

Свойством поворачивать плоскость колебаний, а также вращательной дисперсией обладают не только кристаллические тела (у которых

это свойство наблюдается вдоль оптической оси, когда $c_o = c_e$), но и изотропные среды, например скипидар, раствор сахара в воде и др. Для растворов угол вращения пропорционален концентрации раствора C:

$$\varphi = [\alpha] Cl, \tag{1.32}$$

где $[\alpha]$ — постоянная вращения (различная для различных λ), характеризующая растворенное вещество; этой формулой пользуются для определения концентрации раствора, если измерены φ и l.

Существуют вещества, вращающие плоскость колебаний по часовой стрелке (если смотреть в сторону распространения света); их называют *левовращающими* или отрицательными; вещества, вращающие плоскость колебаний против часовой стрелки, называются *правовращающими* или положительными. Например, существуют две модификации кварца — правовращающая и левовращающая.

Поворот плоскости колебаний волны наблюдается также в веществах, помещенных в магнитное поле (явление Фарадея); поляризованный луч пропускается в направлении поля. Угол поворота плоскости колебаний φ пропорционален пройденному в веществе расстоянию l и напряженности поля H:

$$\varphi = \rho l H,$$

где ρ — величина, характеризующая вещество и зависящая от длины волны света; для различных веществ ρ составляет несколько тысячных долей угловой минуты на ампер. Поворот плоскости колебаний волны происходит также при отражении плоскополяризованного луча от поверхности намагниченного ферромагнетика.

Перечислим важнейшие применения поляризованного света:

1) исследование механических напряжений в сложных деталях машин. Модели (прозрачные) этих деталей помещают между двумя поляризаторами и подвергают деформации; наблюдаемое распределение оптической анизотропии по объему детали показывает распределение механической деформации;

2) изучение различных физических свойств вещества (в частности, полимеров); состава и строения молекул; структуры кристаллических решеток; исследование минералов; количественный анализ веществ и т. д. Для этих целей изготовлен ряд специальных поляризационных приборов, в частности поляризационный микроскоп, сахариметры (для определения концентрации раствора сахара), фотометры и др.;

3) изучение быстропротекающих процессов (скорости распространения света, скорости реакций; звукозапись и воспроизводство звука; фотографирование с очень малой экспозицией порядка 10^{-8} с и т. п.). Для этой цели применяется *ячейка Керра* (рис. IV.38) — сосуд с двумя электродами, между которыми находится вещество, изменяющее степень своей анизотропии под действием приложенного электрического поля. Эта ячейка при перпендикулярном расположении оптических осей поляризатора и анализатора представляет собой быстродействующий оптический затвор; интенсивность проходящего света почти без отставания следует за изменениями приложенного к пластинкам электрического напряжения.

§ 7. ЛИНЗЫ; ИХ ОПТИЧЕСКАЯ СИЛА. АБЕРРАЦИИ. ДИАФРАГМЫ В ОПТИЧЕСКИХ ПРИБОРАХ

Линзами называются прозрачные для данного излучения тела, ограниченные двумя поверхностями различной формы (сферической, цилиндрической и т. д.). Образование сферических линз показано

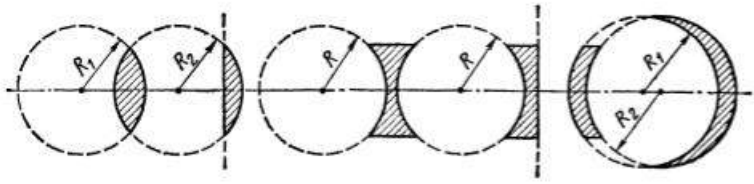


Рис. IV.39

на рис. IV.39. Одна из ограничивающих линзу поверхностей может быть сферой бесконечно большого радиуса, т. е. плоскостью.

Ось, проходящая через центры образующих линзу поверхностей, называется *оптической осью*; у плосковыпуклой и плосковогнутой линз оптическая ось проводится через центр сферы перпендикулярно плоскости.

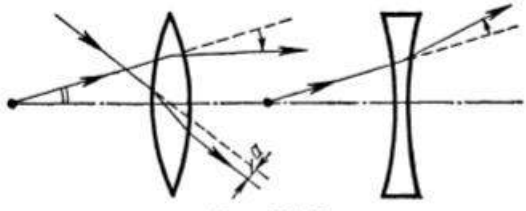


Рис. IV.40

Линза называется *тонкой*, если ее толщина значительно меньше радиусов кривизны образующих поверхностей. В тонкой линзе можно пренебречь смещением a лучей, проходящих через центральную часть (рис. IV.40). Линза является *собирающей*, если она преломляет проходящие через нее лучи в сторону оптической оси, и *рассеивающей*, если она отклоняет лучи от оптической оси.

ФОРМУЛА ЛИНЗЫ

Рассмотрим преломление лучей сначала на одной сферической поверхности линзы. Обозначим точки пересечения оптической оси с рассматриваемой поверхностью через O, с падающим лучом — через O_1 и с преломленным лучом (или его продолжением) — через O_2; точка O_3 есть центр сферической поверхности (рис. IV.41); обозначим расстояния $OO_1 = d$; $OO_2 = f$; $OO_3 = R$ (R — радиус кривизны поверхности). В зависимости от угла падения лучей на сферическую поверхность возможны различные расположения точек O_1, O_2 и O_3 относительно точки O. На рис. IV.41 показан ход лучей, падающих на выпуклую поверхность под разными углами падения α при условии $n_1 < n_2$, где n_1 — показатель преломления среды, откуда идет падающий луч, а n_2 — показатель преломления среды, куда идет преломленный луч. Предположим, что падающий луч — *параксиальный*, т. е.

составляет с оптической осью очень малый угол φ; тогда углы α, γ, ψ и θ также малые и можно считать:

$$O_1A \approx O_1O = d; \quad O_2A \approx O_2O = f;$$
$$\varphi \approx h/d; \quad \psi \approx h/f; \quad \theta \approx h/R. \tag{1.33}$$

На основании закона преломления при малых углах α и γ

$$\sin\alpha/\sin\gamma = n_2/n_1; \quad \sin\alpha \approx \alpha; \quad \sin\gamma \approx \gamma; \quad n_1\alpha \approx n_2\gamma. \tag{1.34}$$

Из рис. IV.41, *а* следует:

$$\alpha = \theta + \varphi = \frac{h}{R} + \frac{h}{d}; \quad \gamma = \theta - \psi = \frac{h}{R} - \frac{h}{f}.$$

Подставив эти выражения в формулу (1.34), получим после сокращения на h формулу преломляющей сферической поверхности:

$$\frac{n_2 - n_1}{R} = \frac{n_1}{d} + \frac{n_2}{f}. \tag{1.35}$$

Зная расстояние d от «предмета» O_1 до преломляющей поверхности, можно по этой формуле рассчитать расстояние f от поверхности до

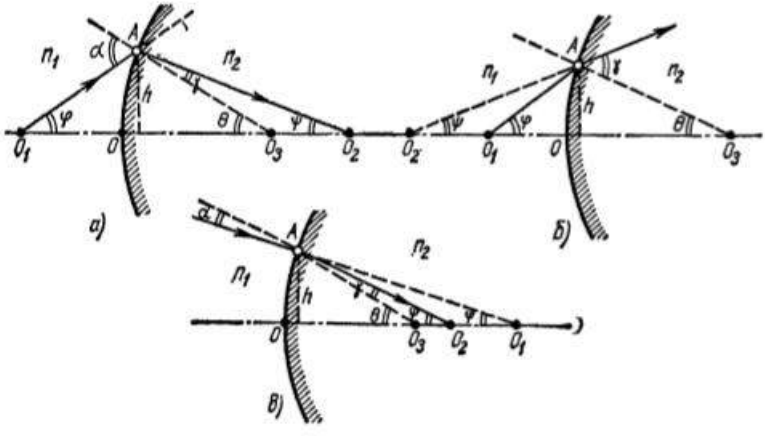


Рис. IV.41

«изображения» O_2. Заметим, что при выводе формулы (1.35) величина h сократилась; это означает, что все параксиальные лучи, вышедшие из точки O_1, какой бы угол φ они ни составляли с оптической осью, соберутся в точке O_2.

Проведя аналогичные рассуждения для других углов падения (рис. IV.41, *б*, *в*), получим соответственно:

$$\frac{n_2 - n_1}{R} = \frac{n_1}{d} + \frac{n_2}{-f}; \quad \frac{n_2 - n_1}{-R} = \frac{n_1}{d} + \frac{n_2}{-f}. \tag{1.36}$$

Отсюда получаем правило знаков (полагая расстояние d всегда положительным): если точка O_2 или O_3 лежит на той же стороне преломляющей поверхности, на которой находится точка O_1, то расстояния f

и R следует брать со знаком минус; если же точка O_2 или O_3 находится по другую сторону поверхности по отношению к точке O_1, то расстояния f и R следует брать со знаком плюс. Такое же правило знаков получится, если рассматривать преломление лучей через вогнутую сферическую поверхность. Для этой цели можно воспользоваться теми же чертежами, приведенными на рис. IV.41, если только изменить направление лучей на обратное и переменить обозначения у показателей преломления.

Линзы имеют две преломляющие поверхности, радиусы кривизны которых R_1 и R_2 могут быть одинаковыми или различными. Рассмотрим двояковыпуклую линзу; для луча, проходящего через такую линзу, первая (входная) поверхность является выпуклой, а вторая (выходная) — вогнутой. Формулу для расчета f_2 по данным d_1, n_1, n_2, R_1 и R_2 можно получить, если воспользоваться формулами (1.35) для входной и (1.36) для выходной поверхности (с обратным ходом лучей, так как луч переходит из среды n_2 в среду n_1):

$$\frac{n_2-n_1}{R_1}=\frac{n_1}{d_1}+\frac{n_2}{f_1};\quad \frac{n_1-n_2}{R_2}=\frac{n_2}{d}+\frac{n_1}{-f_2}. \tag{1.37}$$

Так как «изображение» O_2 от первой поверхности является «предметом» O_1 для второй поверхности, то $f_1=d_2$. Тогда из формулы (1.37) получаем, заменив d_1 на d, f_2 на f,

$$(n_2-n_1)\left(\frac{1}{R_1}+\frac{1}{R_2}\right)=n_1\left(\frac{1}{d}+\frac{1}{f}\right).$$

Из этого соотношения видно, что $1/d+1/f$ — постоянная величина, т. е. d и f взаимосвязаны. Обозначим $1/d+1/f=1/F$, где F — фокусное расстояние линзы ($1/F=D$ называется *оптической силой* линзы и измеряется в *диоптриях*). Следовательно,

$$\frac{1}{F}=\left(\frac{n_2}{n_1}-1\right)\left(\frac{1}{R_1}+\frac{1}{R_2}\right). \tag{1.38}$$

Если же расчет провести для двояковогнутой линзы, то получим

$$\frac{1}{-F}=\frac{1}{d}+\frac{1}{-f};\quad \frac{1}{F}=\left(\frac{n_2}{n_1}-1\right)\left(\frac{1}{-R_1}+\frac{1}{-R_2}\right). \tag{1.39}$$

Сравнивая результаты, можно прийти к выводу, что для расчета оптической силы линзы любой формы следует пользоваться одной формулой (1.38) с соблюдением правила знаков: радиусы кривизны выпуклых поверхностей подставлять со знаком плюс, вогнутых — со знаком минус. Отрицательная оптическая сила ($D<0$), т. е. отрицательное фокусное расстояние ($F<0$), означает, что расстояние f имеет знак минус, т. е. «изображение» находится на той же стороне, где расположен «предмет». В этом случае «изображение» является *мнимым*. Линзы с положительной оптической силой являются собирающими и дают действительные изображения, пока $d>F$; при $d<F$ расстояние f приобретает знак минус и изображение получается мнимым. Линзы с отрицательной оптической силой — рассеивающие и дают всегда мнимое изображение; для них $1/f=-1/F-1/d$ и ни при каких числовых значениях d нельзя получить положительное расстояние f.

Формула (1.38) выведена при условии, что по обе стороны линзы находится одна и та же среда. Если же показатели преломления сред, граничащих с поверхностями линзы n_1 и n_1', различны (например, у хрусталика глаза), то фокусные расстояния справа F и слева F' от линзы не равны, причем

$$\frac{F}{F'} = \frac{n_1'}{n_1}; \quad \frac{F}{d} + \frac{F'}{f} = 1, \tag{1.40}$$

где F — фокусное расстояние с той стороны, где находится предмет.

Заметим, что, согласно формуле (1.38), оптическая сила линзы определяется не только ее формой, но и соотношением между показателями преломления вещества линзы и окружающей среды. Например, двояковыпуклая линза в среде с большим показателем преломления ($n_1 > n_2$) имеет отрицательную оптическую силу, т. е. является рассеивающей

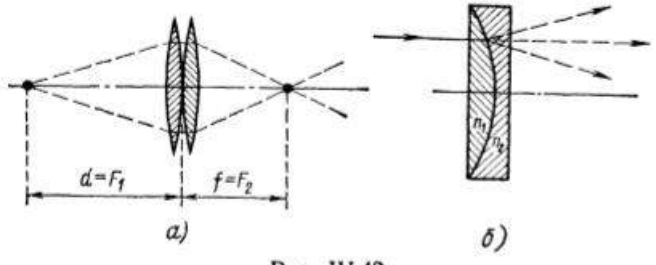


Рис. IV.42

линзой. Наоборот, двояковогнутая линза в такой же среде имеет положительную оптическую силу, т. е. является собирающей линзой.

Рассмотрим систему из двух линз (рис. IV.42, *а*); допустим, что точечный предмет находится в фокусе первой линзы. Луч, вышедший из первой линзы, будет параллельным оптической оси и, следовательно, пройдет через фокус второй линзы. Рассматривая эту систему как одну тонкую линзу, можем написать $1/d + 1/f = 1/F$. Так как $d = F_1$, $f = F_2$, то

$$\frac{1}{F_{\text{сист}}} = \frac{1}{F_1} + \frac{1}{F_2}. \tag{1.41}$$

Этот результат верен и для более сложной системы тонких линз (если только сама система может рассматриваться как «тонкая»): оптическая сила системы тонких линз ($D_{\text{сист}} = 1/F_{\text{сист}}$) равна сумме оптических сил ($D_i = 1/F_i$) составных частей:

$$D_{\text{сист}} = D_1 + D_2 + D_3 + \ldots \tag{1.42}$$

(у рассеивающих линз оптическая сила имеет отрицательный знак). Например, плоскопараллельная пластинка, составленная из двух тонких линз (рис. IV.42, *б*), может быть собирающей (если $n_1 > n_2$) или рассеивающей (если $n_1 < n_2$) линзой. Для двух тонких линз, находящихся на расстоянии a друг от друга (рис. IV.43), оптическая сила является функцией от a и фокусных расстояний линз F_1 и F_2.

Заметим, что при $a = F_1 + F_2$ пучок параллельных лучей, входящих в систему, выходит из нее также параллельным пучком.

Увеличение, которое дает тонкая линза, зависит от расположения предмета и от оптической силы линзы или ее фокусного расстояния F. Если предмет находится за фокусом (рис. IV.44, *а*), то изображение получается действительным и увеличенным в $\beta = f/d$ раз. Пользуясь соотношением $1/F = 1/d + 1/f$, можно увеличение линзы представить в виде

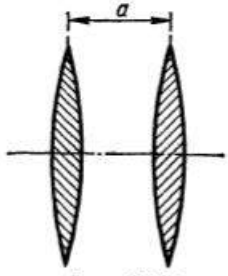


Рис. IV.43

$$\beta = \frac{f}{d} = \frac{f-F}{F} = \frac{F}{d-F}. \qquad (1.43)$$

Величина β называется *линейным увеличением*; из этой формулы видно, что увеличение линзы тем больше, чем ближе расположен предмет к фокусу линзы.

Если же предмет находится между фокусом и линзой (рис. IV.44, *б*), то изображение получается мнимым и увеличенным также в $\beta = f/d$ раз. Из треугольника O_2AF_2 следует, что это увеличение $\beta = AO_2/BO$ равно O_2F_2/OF_2, т. е.

$$\beta = \frac{f+F}{F} = \frac{\Delta}{F}. \qquad (1.44)$$

Очевидно, увеличение при мнимом изображении тем больше, чем ближе предмет к фокусу (F_1). Положение глаза за линзой определяется

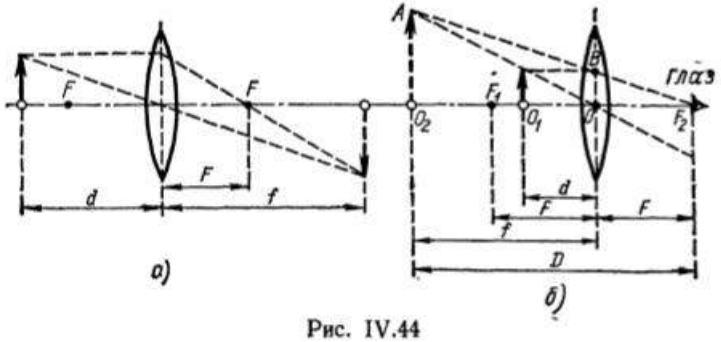


Рис. IV.44

условием, чтобы изображение AO_2 находилось на расстоянии наилучшего зрения, равного для нормального глаза 25 см.

Простейший *микроскоп* состоит из объектива, дающего действительное и увеличенное изображение, и окуляра, дающего мнимое изображение, дополнительно увеличенное. Если изображение необходимо сфотографировать, то от второй линзы получают действительное изображение. Для получения наибольшего увеличения предмет помещается вблизи «наружного» фокуса объектива (немного дальше этого фокуса), а изображение, даваемое объективом, располагается вблизи «внутреннего» фокуса окуляра (ближе к окуляру). Общее *увеличение микроскопа* равно произведению увеличений объектива $\beta_{об}$ (см. фор-

мулу (1.43)) и окуляра $\beta_{ок}$ (см. формулу (1.44)):

$$\beta = \beta_{об}\beta_{ок} = \frac{(f_{об} - F_{об})}{F_{об}} \cdot \frac{(f_{об} + F_{ок})}{F_{ок}}. \qquad (1.45)$$

Так как изображение от объектива помещается около фокуса окуляра, то разность $(f_{об} - F_{об})$ равна расстоянию между внутренними фокусами объектива и окуляра.

ТОЛСТЫЕ ЛИНЗЫ. АБЕРРАЦИИ

Для построения изображений в тонких линзах достаточно знать расположение их фокусов. Для толстых линз и для сложных систем линз этого недостаточно. Рассмотрим *центрированные оптические системы*, т. е. такие системы линз, у которых центры кривизны всех преломляющих поверхностей лежат на одной прямой — оптической оси. Построение изображений в таких приборах облегчается заданием: двух *главных фокусов* (рис. IV.45) — переднего F_1 и заднего F_2; двух

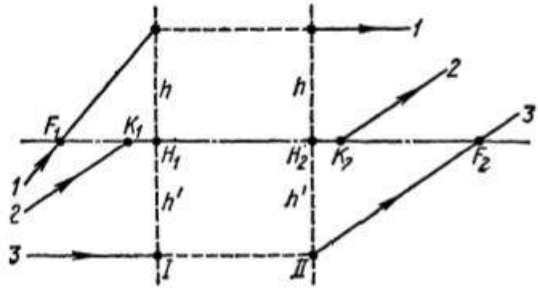


Рис. IV.45

узловых точек — K_1 и K_2 и *двух главных плоскостей I* и *II*, перпендикулярных оптической оси системы. Точки пересечения H_1 и H_2 плоскостей *I* и *II* с оптической осью называются *главными точками*. Луч *1*, идущий от предмета через фокус, например F_1, и пересекающий главную плоскость *I* на расстоянии h от оптической оси, выходит из главной плоскости *II* на таком же удалении h от оптической оси и параллельно ей. Луч *2*, идущий от предмета к одной узловой точке, выходит из другой узловой точки параллельно своему направлению до системы. Луч *3*, направленный от предмета параллельно оптической оси, выходит из второй главной плоскости на таком же удалении h' от оптической оси и проходит через второй фокус. Пользуясь этими лучами, можно построить изображение любой точки предмета, а затем определить и увеличение. Если по обе стороны системы находится одна и та же среда, то узловые точки сливаются с главными точками. На рис. IV.46 показано расположение главных плоскостей у некоторых толстых линз; у тонких линз эти плоскости сливаются в одну, проходящую через центр линзы, а у толстых они могут быть расположены даже вне объема линзы.

Обозначим через d расстояние от предмета до первой главной плоскости, а f — расстояние изображения от второй главной плоскости.

Расстояния от главных плоскостей до фокусов H_1F_1 и H_2F_2 обозначим через F_1 и F_2. Тогда оказывается, что

$$\frac{F_1}{d} + \frac{F_2}{f} = 1. \tag{1.46}$$

По этой формуле можно рассчитать f, если задано d. Если предмет и изображение лежат в одной и той же среде, т. е. по обе стороны

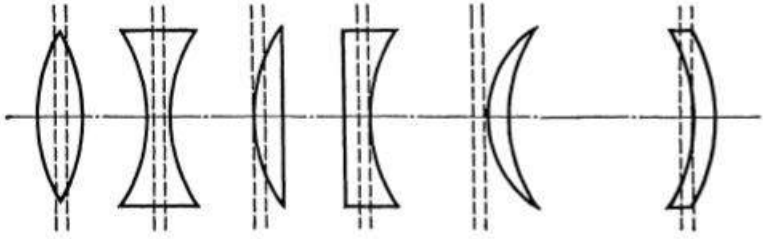


Рис. IV.46

системы находится одна среда (например, воздух), то $F_1 = F_2 = F$ и формула (1.46) принимает вид

$$\frac{1}{d} + \frac{1}{f} = \frac{1}{F}. \tag{1.47}$$

Линзы могут дать отчетливое изображение предмета только в том случае, если лучи, выходящие из одной точки предмета по разным направлениям, после преломления в линзе собираются (пересекаются)



Рис. IV.47

в одной точке. Изображение называют *стигматическим*, если каждой точке предмета соответствует одна точка изображения. Стигматичность соблюдается, если, во-первых, все лучи, дающие изображение, параксиальные, т. е. составляют небольшие углы с оптической осью, во-вторых, если все лучи преломляются в линзе одинаково, т. е. нет дисперсии.

Перечислим возможные недостатки (*аберрации*) линз:

1) *сферическая аберрация*. Если изображение получается при помощи не только параксиальных лучей ($п$ на рис. IV.47, *а*), но и краевых лучей ($к$), то стигматичность нарушается: лучи, вышедшие из одной точки предмета, не пересекаются в одной точке. Различают:

а) продольную аберрацию, когда изображение точки растянуто вдоль оптической оси, и б) поперечную, когда изображение деформировано в перпендикулярном направлении. Собирающие линзы дают изображение параксиального пучка лучей дальше, чем краевых, а рассеивающие линзы, наоборот, отбрасывают дальше краевые лучи. Благодаря этому удается, комбинируя выпуклые и вогнутые линзы, ослабить сферическую аберрацию. Относительно малой сферической аберрацией обладают линзы, имеющие различные радиусы кривизны (например, плосковыпуклая), причем они должны быть обращены к падающим лучам наиболее выпуклой стороной;

2) *ко́ма*. Лучи, идущие в линзу широким пучком из точек предмета, не лежащих на оптической оси линзы, после преломления собираются не в одной точке, а дают расплывчатое изображение, похожее на короткую запятую. Ко́му устраняют теми же приемами, что и сферическую аберрацию, либо же применяют диафрагмы;

3) *хроматическая аберрация*. Так как вещество линзы (и окружающей среды, если это не воздух) обладает дисперсией (т. е. показатель преломления зависит от длины волны света, дающего изображение), то лучи различных цветов преломляются под различными углами и, следовательно, пересекаются в различных точках. Если свет, исходящий от предмета, состоит, например, из двух цветов, то линза даст два цветных изображения, находящихся в различных местах. Отдельные участки линз действуют как призмы, поэтому у собирающих, например стеклянных, линз фиолетовые лучи, преломляясь, отклоняются к оптической оси сильнее, чем красные, а у рассеивающих линз, наоборот, фиолетовые лучи после преломления отклоняются от оптической оси сильнее, чем красные.

Вещества, из которых делают линзы (различные сорта стекла, кварц, флюорит, каменная соль), обладают различной дисперсией. Этим пользуются для того, чтобы ослабить хроматическую аберрацию хотя бы для двух или трех цветов. Если составить сложные (склеенные) линзы, состоящие из выпуклых и вогнутых линз с различной дисперсией, то удается получить изображения предмета, образованные, например, синими и желтыми лучами, в одной плоскости и одинаковым увеличением (причем другие лучи в этой плоскости резкого изображения не дают). Такие сложные линзы (системы линз, объективы) называются *ахроматами*.

Если изображение воспринимается приемником излучения, особенно чувствительным к определенным цветам (например, глаз, фотографическая эмульсия и др.), то для них изготовляют ахроматы, собирающие вместе эти цвета в одном изображении.

Заметим, что увеличение, которое дает линза, зависит от фокусного расстояния; последнее, согласно формуле (1.38), зависит от показателя преломления линзы для данного цвета. Таким образом, лучи различных цветов дают изображения различных размеров;

4) *астигматизм* — это нарушение стигматичности, при котором каждая точка предмета дает изображение в виде короткой черточки, причем если экран расположен в плоскости *1*, то черточка́ перпендикулярна чертежу (рис. IV.47, *б*); если же экран расположен в плос-

кости *II*, то черточка лежит в этой же плоскости перпендикулярно оптической оси. Астигматизмом обладают линзы, имеющие несимметричные относительно оптической оси преломляющие поверхности, например цилиндрические линзы.

Для устранения астигматизма изготовляют сложные линзы (анастигматы) с соответствующим подбором кривизны преломляющих поверхностей и показателей преломления веществ;

5) *дисторсия* — это недостаток линз, при наличии которого прямая линия, взятая на предмете, изображается в виде кривой линии. Такое искажение объясняется различным увеличением различных частей предмета. Чаще всего увеличение монотонно растет или убывает по мере удаления от оптической оси. Дисторсия особенно вредна при аэрофотосъемках, производимых для составления карт. Ее устраняют, комбинируя линзы, дающие искривления в противоположных направлениях.

В высококачественных объективах эти недостатки должны быть по возможности сведены к минимуму. Кроме того, дополнительно они должны удовлетворять другим требованиям в зависимости от их применения. Например, при фотографировании необходимо, чтобы изображение «объемного» предмета было получено на плоскости, где располагается фотографическая пластинка или пленка. Кроме того, чтобы изображение было точной (неискаженной) копией предмета, необходимо одинаковое увеличение всех его частей, т. е. *увеличение должно быть одинаковым во всех местах изображения по всем направлениям*. В частности, если объектив есть ахромат, соединяющий вместе изображения в двух цветах, то увеличения, даваемые лучами каждого цвета, должны быть одинаковыми, т. е. изображения от лучей двух цветов должны получаться в одном месте и иметь одинаковые размеры.

Для получения отчетливых изображений в линзах и оптических системах применяют диафрагмы, отрезающие лучи, идущие в прибор под большими углами к оптической оси, или ограничивающие ширину пучков лучей, поступающих в прибор от каждой точки предмета. Половина угла φ (рис. IV.48, *а*), образованного крайними лучами (*1*, *2*), попадающими в изображение от одной точки предмета, называется *апертурным углом*. Лучи, находящиеся вне этого угла, могут пройти через входную линзу, но они отрезаются диафрагмой и в соответствующую точку изображения не попадут. *Диафрагма* D_1, *ограничивающая апертурный угол, называется действующей* или *апертурной*. Роль этой диафрагмы может играть оправа входной линзы (или оправа какой-нибудь из внутренних линз).

Угол ψ (рис. IV.48, *б*), ограниченный лучами (*1*, *2*), исходящими от крайних точек предмета, еще изображаемых в системе, выделяет «поле зрения». Точки предмета M_3, лучи от которых находятся вне этого угла, могут пройти через входную линзу, но они отрезаются диафрагмой D_2, называемой *диафрагмой поля зрения*. Роль этой диафрагмы может играть оправа какой-нибудь линзы; возможно, что апертурная диафрагма одновременно является и диафрагмой поля зрения.

Диафрагмы, а также оправы линз являются отверстиями, которые отрезают определенные участки фронта сферических волн, исходящих

от каждой точки предмета. Дифракция этих волн за отверстиями в диафрагмах вносит дополнительное искажение изображения («размазывание» светового потока, поступившего в прибор). Кроме того, эта дифракция не позволяет установить, исходит ли свет, поступивший в прибор, от одной или от двух очень близких друг к другу точек.

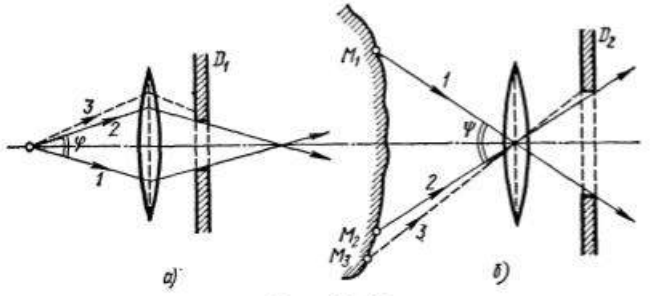


Рис. IV.48

Допустим, что от двух очень близких точек *1* и *2* в прибор поступили две волны. На плоскости, где получаются изображения этих точек, их дифракционные спектры накладываются друг на друга. Согласно условию разрешимости (см. § 5), можно установить наличие двух спектров (и, следовательно, «разрешить» точки *1* и *2*), если главные максимумы этих спектров расположены достаточно далеко друг от друга или, в пределе, если главный максимум одной волны совпадает с первым минимумом второй волны. Для круглого отверстия (см. § 5, формула (1.25)) значение угла $\delta\varphi$, удовлетворяющего этому условию (рис. IV.49), равно $\delta\varphi = 1{,}22\,\lambda/D$, где λ — длина волны, D — диаметр отверстия. Тогда наименьшее расстояние между «разрешимыми» точками будет равно $\delta y = l\delta\varphi$, где l — расстояние от этих точек до отверстия. Вместо диаметра диафрагмы можно ввести апертурный угол, который часто обозначают u ($\sin u = D/2l$), и тогда

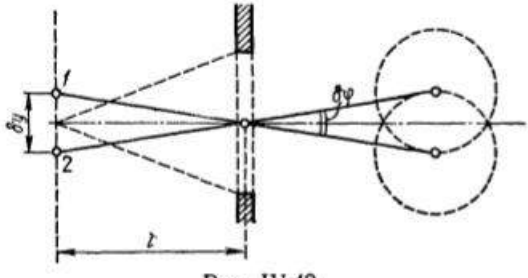


Рис. IV.49

$$\delta y = 0{,}61\,\frac{\lambda}{\sin u}. \qquad (1.48)$$

В микроскопах рассматриваемые объекты иногда помещаются в среду с показателем преломления n, что увеличивает разрешающую способность микроскопа приблизительно в n раз:

$$\delta y = 0{,}61\,\frac{\lambda}{n \sin u}.$$

Оптические системы (бинокли, зрительные трубы, микроскопы, телескопы и т. д.) составляются в основном из линз, призм и зеркал. В этих системах потери светового потока при прохождении через детали и отражения от их поверхностей, а также различные аберрации должны быть по возможности минимальными. Это предъявляет высокие требования к качеству вещества, из которого делаются детали, к точности расчета и изготовления их поверхностей и т. д.

Для больших систем, например телескопов, возникают дополнительные требования: диаметр стеклянной линзы — объектива (в рефракторах) или зеркала (в рефлекторах, рис. IV.50, *а*) должен быть

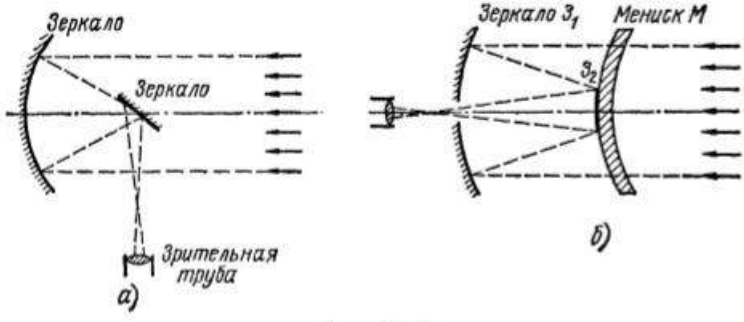


Рис. IV.50

по возможности большим. Однако технологические трудности изготовления ограничивают эти размеры: линзы, диаметр которых достигает 100 см, должны быть сделаны из совершенно однородного стекла; зеркала, диаметр которых 3—5 м, должны иметь не сферическую, а параболическую форму; изготовление таких зеркал связано с большими трудностями. Д. Максутов предложил телескоп, обладающий весьма высокими качествами (рис. IV.50, *б*): свет проходит через стеклянный мениск M, ограниченный сферическими поверхностями, и затем отражается от двух зеркал: $З_1$ и $З_2$. Соответствующим подбором размеров мениска и радиусов зеркал удается значительно уменьшить все аберрации этой системы по сравнению с другими. Кроме того, размеры, вес и стоимость изготовления этого телескопа значительно меньше, чем у равных по оптическим параметрам телескопов других систем. Менисковые системы Д. Максутова применяются также в геодезической, киносъемочной и другой оптической аппаратуре.

§ 8. ПОНЯТИЕ О ГОЛОГРАФИИ

Голография (впервые предложенная Ю. Габором в 1947 г.) есть особый метод регистрации на фотопластинке (или пленке) волнового фронта излучения, идущего от предмета, и затем восстановления этого фронта для получения мнимого или действительного изображения предмета. Голография основана на фотографировании результата интерференции двух частей волны (плоской или сферической), испус-

каемой *когерентным* источником, например лазером: одна часть поступает на фотопластинку непосредственно и называется «опорной волной», а вторая часть направляется для освещения изучаемого объекта и после отражения от него поступает на ту же пластинку; эту часть волны называют *предметной*. На фотопластинке после проявления оказывается зафиксированной сложная интерференционная картина, называемая *голограммой*. В каждой точке голограммы результат интерференции опорной и предметной волн зависит не только от амплитуд их электрических векторов, но и от разности фаз между ними. Ввиду этого сведения о предмете, записанные на голограмме, оказываются более полными, чем при обычном фотографировании.

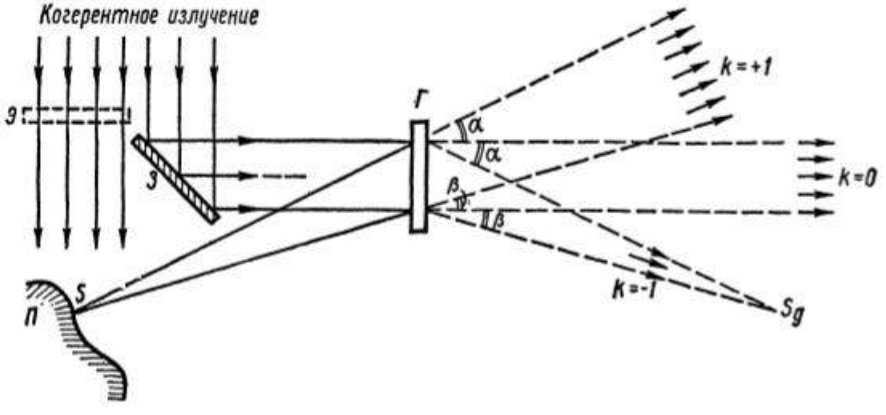


Рис. IV.51

На рис. IV.51 показана схема одной из простейших установок для получения сначала голограммы, а затем и изображения предмета. Когерентная (и монохроматическая) волна от лазера частично отражается от зеркала *З* и поступает на фотопластинку *Г*, частично же направляется на предмет *П*. От каждой точки *S* предмета на фотопластинку поступает сферическая волна, которая на всей освещаемой части пластинки интерферирует с плоской опорной волной. Можно показать, что в результате этой интерференции на фотопластинке получаются чередующиеся темные и светлые концентрические кольца с синусоидальным распределением коэффициента пропускания (см. § 5). Эта интерференционная картина фиксируется на фотопластинке и является голограммой данной точки *S* предмета. Каждая точка предмета в зависимости от разности фаз между волной, исходящей от этой точки предмета, и опорной волной образует на фотопластинке свою голограмму. Таким образом, на голограмме предмета окажутся записанными голограммы всех его точек.

Для того чтобы получить изображение предмета по его голограмме, направляют на голограмму то же опорное излучение, при помощи которого она была снята. В схеме, указанной на рис. IV.51, можно, например, просто удалить предмет или воспользоваться экраном *Э*

и закрыть ту часть излучения, которая раньше направлялась на предмет при фотографировании голограммы. Изображения предмета получаются в результате дифракции опорной волны на периодической структуре голограммы. Расчеты показывают, что плоская опорная волна, прошедшая через голограмму, образованную одной точкой предмета, в результате дифракции разделяется на три волны: одну плоскую, распространяющуюся в том же направлении ($k = 0$), и две сферические волны — расходящуюся ($k = +1$) и сходящуюся ($k = -1$). На рис. IV.51 показано, что каждый луч опорной волны после дифракции разделяется на прямой луч ($k = 0$) и два симметрично расположенных луча ($k = \pm 1$); эти лучи показаны пунктиром. Лучи с $k = -1$ сходятся в точке $S_д$ и образуют действительное изображение точки S предмета; лучи с $k = +1$ расходятся и могут создать мнимое изображение (которое получается в том месте, где раньше находилась точка S предмета). Таким образом будут получены действительные и мнимые изображения каждой точки предмета, а следовательно, и соответствующее полное изображение всего предмета. Разумеется, голограмма предмета представляет собой весьма сложное распределение почернения на фотопластинке. Однако можно полагать, что сравнительно простая интерференционная картина от каждой отдельной точки предмета, условно выражаясь, «действует» независимо от других, благодаря чему получается очень отчетливое изображение всего предмета. Заметим также, что на голограмме будет записана не только интерференция предметной и опорной волн, но и результаты интерференции отдельных частей самой предметной волны между собой; однако интерференционная картина, вызванная одной только предметной волной, оказывается значительно более слабой, чем основная картина, образованная совместно предметной и опорной волнами.

Для удобства наблюдения действительного изображения предмета (и фотографирования его) голограмму освещают не с левой, как показано на рисунке, а с правой стороны (под тем же углом). Тогда действительное изображение получается в том месте, где находился предмет при снятии голограммы.

Перечислим некоторые достоинства голографии по сравнению с обычным фотографированием. Объектив фотоаппарата (допустим, идеальный) преломляет лучи, вышедшие из одной точки предмета, таким образом, что они пересекаются в одной точке, которую мы называем *изображением*. По существу, объектив дает объемное изображение предмета; однако изображения некоторых точек предмета оказываются либо ближе, либо дальше плоского светочувствительного слоя. Почернение пластинки в различных местах определяется только энергией, которая поступает на эти участки за время экспозиции, поэтому на фотографии невозможно установить, какие части предмета находились ближе или дальше от фотоаппарата. В голограмме же интерференционная картина зависит от разности фаз между опорной волной и волной, идущей от данной точки предмета. Эта разность фаз различна для различных точек предмета, поэтому от голограммы можно получить почти не искаженное объемное изображение предмета.

Далее, при обычном фотографировании различные части предмета фиксируются в различных местах фотопластинки, поэтому сведения о предмете можно получить, изучая всю фотографию; вследствие этого размеры фотографии не могут быть очень малыми. В голограмме же сведения как о каждой точке предмета, так и о предмете в целом записаны на всей площади фотопластинки, поэтому от каждой, даже очень малой части пластинки, на которой записана голограмма, можно получить изображение всего предмета.

Важным достижением голографии оказалась возможность записи интерференционной картины (имеющей вид стоячих волн) в толстом слое фотоэмульсии, т. е. получение *трехмерных голограмм* (метод Ю. Денисюка, 1962 г.); толщина слоя должна быть значительно больше расстояний между соседними интерференционными поверхностями. Такие голограммы представляют собой трехмерные дифракционные структуры, которые при освещении дают объемные изображения предмета. При этом голограмма должна быть освещена излучением той же самой длины волны, при помощи которого производилась запись. На одной фотопластинке могут быть одновременно записаны голограммы нескольких тел, если воспользоваться опорными излучениями с различными длинами волн (и под различными углами падения). Далее, можно получить голограммы, дающие объемное цветное изображение предмета, если записать на одной пластинке голограммы этого предмета в трех основных цветах — красном, зеленом, синем (изображение предмета получается при освещении белым светом). Голография дает возможность получить также и увеличенные изображения предмета, если голограмма снимается на коротких волнах, а изображение получают при освещении длинноволновым излучением.

Известны многие применения голографии: для записи и хранения большого информационного материала на малых по размерам участках голограммы; для получения изображения от предметов, рассеивающих звуковые волны; для изучения физических процессов на поверхности тел и т. д.

Глава 2

ФИЗИКА АТОМОВ И МОЛЕКУЛ. ОСНОВЫ КВАНТОВОЙ ТЕОРИИ

§ 9. СТРОЕНИЕ АТОМА. ОПЫТЫ РЕЗЕРФОРДА, ФРАНКА И ГЕРЦА. ЭЛЕМЕНТАРНАЯ ТЕОРИЯ АТОМА ВОДОРОДА

Возникшее в конце XIX в. представление о том, что в атоме кроме электронов имеются еще и положительные заряды, требовало дальнейшего уточнения. Необходимо было выяснить, «размазан» ли этот положительный заряд по всему предполагаемому объему атома или же сосредоточен в очень малом объеме в виде некоторой положительно заряженной частицы. Этот вопрос можно было решить экспериментально, например изучая рассеяние атомом пучка каких-нибудь заряженных частиц, в частности альфа-частиц, испускаемых радиоактив-

ными веществами. Так как масса альфа-частиц почти в 7350 раз больше массы электрона, то участием электронов, содержащихся в атоме, в процессе рассеяния альфа-частиц можно пренебречь; характер рассеяния будет определяться главным образом размерами того объема, в котором сосредоточены инертная масса и положительный заряд атома.

ОПЫТЫ РЕЗЕРФОРДА

Допустим, что масса и положительный заряд атома, согласно предположению Дж. Томсона, равномерно распределены по сфере радиуса r. На рис. IV.52, *а* показаны траектории альфа-частиц, проходящих мимо такой сферы. Полагая, что альфа-частица и положительный заряд рассеивающего атома взаимодействуют по закону Кулона, можно рассчитать угол φ отклонения альфа-частиц в зависимости от «прицельного расстояния» b. Для заряда, равномерно распределенного по объему сферы, напряженность электрического поля равна нулю

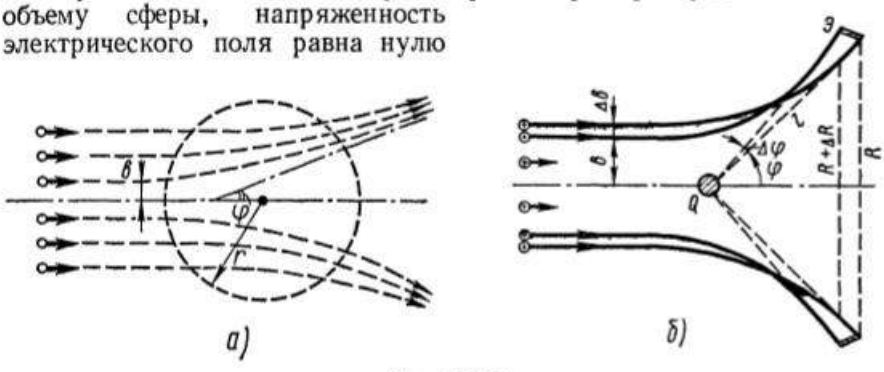


Рис. IV.52

в центре и максимальна на поверхности (см. ч. III, § 3), поэтому наибольший угол отклонения будет при $b = r$. Расчет показывает, что при таких предположениях атомы золота ($r \approx 10^{-10}$ м) будут отклонять альфа-частицы, летящие со скоростью $v \approx 2 \cdot 10^7$ м/с, на угол, не превышающий 0,02°. Пользуясь теорией случайных событий, можно определить средний угол отклонения альфа-частиц при прохождении их через тонкую фольгу из золота, содержащую около 10 000 атомных слоев. Этот угол оказывается равным приблизительно 1°, что в основном наблюдалось на опыте. Однако эксперименты Резерфорда показали, что некоторая (хотя и очень незначительная) часть альфа-частиц отклонялась на большие углы (1/8000 часть — на угол в 90° и т. п.). Вероятность того, что очень большие отклонения могли получиться при многократном рассеянии на отдельных атомах, оказывается настолько маленькой, что такое объяснение приходится отвергнуть. Таким образом, представление о положительном заряде атома, «размазанном» по его объему, оказалось в противоречии с данными наблюдений.

Э. Резерфорд, анализируя результаты этих опытов, высказал предположение, что положительные заряды как альфа-частиц, так и рассеи-

вающего атома сосредоточены в весьма малых объемах. Если для расчета угла отклонения φ их полагать почти точечными телами, то углы φ могут получиться любыми (от 0 до 180°). Допустим, что параллельный пучок альфа-частиц (рис. IV.52, *б*) рассеивается точечным зарядом Q. Выделим частицы, которые имеют «прицельное расстояние» в пределах b, $b + \Delta b$; их число будет пропорционально площади $\pi b^2 \Delta b$. В результате кулоновского взаимодействия с зарядом Q эти частицы будут отклонены на углы, лежащие в пределах φ, $\varphi + \Delta\varphi$, и попадут на кольцевую площадку экрана $Э$, ограниченную радиусами R, $R + \Delta R$. Так как угол отклонения φ зависит от b, то в окончательном расчете можно исключить b и получить зависимость между относительным числом $\Delta N/N$ альфа-частиц, падающих на единицу площади экрана, и углом φ в виде (формула Резерфорда)

$$\frac{\Delta N}{N} = nd\left[\frac{Ze^2}{4\pi\varepsilon_0 lmv^2 \sin(\varphi/2)}\right]^2, \qquad (2.1)$$

где d — толщина фольги, n — число атомов в единице объема фольги, Z — число элементарных положительных зарядов в атоме, т. е. порядковый номер золота в периодической системе Менделеева, m и v — масса и скорость альфа-частицы, размер l указан на рисунке. Эта формула получила экспериментальное подтверждение на измерительной установке, схема которой показана на рис. IV.53. Альфа-частицы, испускаемые радиоактивным препаратом P, проходили через отверстие в свинцовой пластинке C и параллельным пучком падали на тонкую фольгу $Ф$. Рассеянные фольгой альфа-частицы регистрировались на флюоресцирующем экране $Э$, который мог перемещаться в направлениях, указанных стрелкой. Каждая частица, упавшая на экран, вызывала вспышку (сцинтилляцию); измерялось число таких вспышек, приходящихся на единицу площадки экрана.

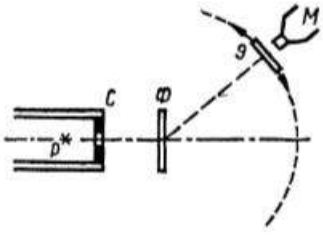


Рис. IV.53

Экспериментальное подтверждение формулы Резерфорда (2.1) послужило основанием для разработки планетарной модели атома, согласно которой масса и положительный заряд атома сосредоточены в весьма малом объеме, названном *ядром* атома; вокруг ядра по замкнутым орбитам вращаются электроны. Для оценки размера ядра можно воспользоваться тем, что резерфордовское представление допускает отклонение альфа-частицы также и на 180°, соответствующее движению этой частицы по направлению к ядру с прицельным расстоянием $b = 0$. Наименьшее расстояние r_0, до которого могут сблизиться альфа-частицы и ядро атома золота ($Z = 79$), можно найти из равенства кинетической и потенциальной энергий. Связав систему отсчета с центром масс взаимодействующих частиц и пренебрегая кинетической энергией ядра атома золота, можно написать

$$\frac{mv^2}{2} \approx \frac{Ze \cdot 2e}{4\pi\varepsilon_0 r_0}.$$

Энергия альфа-частицы была равна $1{,}2 \cdot 10^{-12}$ Дж, поэтому $r_0 \approx 3 \cdot 10^{-14}$м. Очевидно, что размеры ядра атома золота должны быть меньше рассчитанного r_0.

Формула (2.1) может быть использована также и для определения числа Z рассеивающего атома по известным значениям n, e, энергии альфа-частицы и измеренным $\Delta N/N$. Для платины, серебра и меди Чадвик получил значения Z, очень близкие к порядковому номеру этих элементов.

ОПЫТ ФРАНКА И ГЕРЦА

Другим экспериментом, давшим важные для теории атомов результаты, был опыт Франка и Герца: исследовалось поглощение энергии атомами при их столкновениях с электронами. Схема опыта представлена на рис. IV.54. В трубке T с парами ртути находятся подогреваемый катод K (источник электронов) и сетка C. При помощи батареи $Б$, многоомного сопротивления R и движка $Д$ в трубке создаётся электрическое поле, разгоняющее электроны по направлению к сетке C (движок $Д$ позволяет изменять разность потенциалов V между катодом и сеткой). Между сеткой и вторым электродом A приложена небольшая разность потенциалов ($V_{торм} \approx 0{,}5$ В), тормозящая движение электронов, вследствие чего электроны, прошедшие через сетку с малыми скоростями, не могут попасть на электрод A. На опыте измерялись напряжение V между катодом и сеткой и ток I в участке CA, регистрируемый чувствительным гальванометром $Г$.

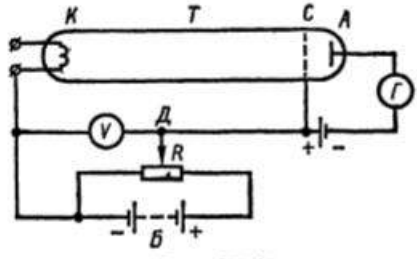


Рис. IV.54

Допустим, что столкновения электронов с атомами ртути являются упругими, при которых сохраняются импульсы и кинетические энергии соударяющихся тел (см. ч. 1, § 13). При каждом отдельном столкновении произойдет некоторое изменение скорости электрона по величине; однако при большом числе столкновений с различными ориентациями скоростей и при отсутствии электрического поля изменение энергии электрона в среднем будет равно нулю. Если же к трубке приложено электрическое поле, то оно будет непрерывно увеличивать скорость электронов независимо от сложности их траектории. Вблизи сетки энергия электронов будет равна eV (если пренебрегать начальной энергией при выходе из катода), и так как $V > V_{торм}$, то электроны смогут преодолеть тормозящее поле на участке CA и попасть на электрод A. Поэтому следует ожидать, что по мере увеличения разгоняющего потенциала V ток I в цепи гальванометра должен монотонно увеличиваться. Однако измерения показали, что ток при определенных значениях V, кратных $\approx 4{,}9$ В, резко убывает (рис. IV.55). Для объяснения этого результата пришлось сделать два предположения:

1) столкновения между электронами и атомами ртути могут быть неупругими, при которых кинетическая энергия электрона расходуется на изменение внутреннего состояния атома ртути (на возбуждение или ионизацию);

2) при этих столкновениях атом ртути поглощает энергию только определенными «порциями», равными 4,9 эВ. Если энергия электрона меньше этой величины, то поглощения энергии не происходит и столкновение оказывается упругим. Если же электрон имеет энергию, превышающую 4,9 эВ, то столкновение может быть неупругим; при этом энергия, равная 4,9 эВ, поглощается атомом ртути, а избыток остается у электрона в кинетической форме.

Резкое падение тока I при $V = 4{,}9$ В показывает, что первое неупругое столкновение произошло вблизи сетки C, приближаясь к которой электрон успел приобрести необходимую для этого энергию; оставшаяся после столкновения энергия будет недостаточна для преодоления тормозящего потенциала $V_{торм}$ и электроны не достигнут электрода A. При постепенном увеличении V место неупругого столкновения будет отодвигаться от сетки; на оставшемся до сетки участке пути электроны будут вновь ускоряться и ток I с увеличением V будет опять монотонно возрастать. Второе падение тока при $V = 9{,}8$ В объясняется тем, что на участке $K - C$ произошли два неупругих столкновения: одно — посередине, второе — вблизи сетки. Последующие падения тока I будут происходить при таких значениях V, при которых на участке $K - C$ возможны три, четыре неупругих столкновения и более (причем последнее из таких столкновений — возле сетки).

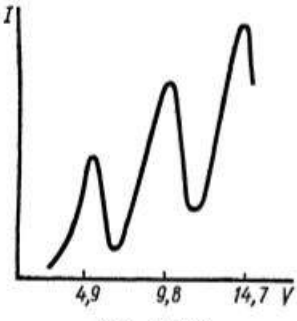


Рис. IV.55

Эти опыты были повторены с другими газами; для паров натрия резкие изменения тока I наблюдались при напряжениях, кратных 2,1 В, для цинка — при 4,1 В, для гелия — при 19,75 В и т. д.

На основе приведенного выше объяснения опыта Франка и Герца можно сделать еще одно предположение: если энергия 4,9 эВ затрачивается на перевод атома ртути в возбужденное состояние, то можно ожидать, что при возвращении в нормальное состояние эти атомы будут испускать фотоны с энергиями $\varepsilon_ф = h\nu = 4{,}9$ эВ. Это предположение оказалось верным: обнаруженное от трубки ультрафиолетовое излучение имело длину волны $\lambda = 0{,}2537$ мкм; расчет же дает

$$\lambda = \frac{c}{\nu} = \frac{hc}{h\nu} = \frac{hc}{\varepsilon_ф} = \frac{3 \cdot 10^8 \cdot 6{,}62 \cdot 10^{-34}}{4{,}9 \cdot 1{,}6 \cdot 10^{-19}} \approx 0{,}2533 \cdot 10^{-6} \text{ м}.$$

СПЕКТРЫ ИЗЛУЧЕНИЯ АТОМОВ

Опыты по рассеянию альфа-частиц и по поглощению энергии электронов при столкновениях с атомами были попытками получить некоторые сведения о структуре и свойствах атомов путем внешних воздей-

ствий. Однако ценная информация может быть получена также и непосредственно от атомов при исследовании испускаемого ими излучения. Для того чтобы характер излучения данного атома определялся только его структурой и свойствами, он должен быть по возможности предохранен от воздействия других атомов (которые, вообще говоря, могут внести нежелательные искажения). Такие условия существуют в разреженных газах: атомы вещества в среднем находятся на относительно больших расстояниях друг от друга и бóльшую часть времени пребывают в свободном состоянии. Поэтому вещество, приведенное в состояние разреженного газа, испускает излучение, спектр которого состоит из определенного набора частот, характерного только для данного вещества, тогда как твердые и жидкие вещества излучают непрерывный спектр, в котором определяющим фактором оказывается сильное взаимодействие между атомами и молекулами, а их индивидуальные свойства отходят на второй план.

Спектр излучения свободных атомов данного вещества характеризуется:

1) значениями частот $\nu_1, \nu_2, \ldots$ (или длин волн $\lambda_1, \lambda_2, \ldots$) излучаемых ими «спектральных линий». Так как каждая спектральная линия (см. § 1) охватывает некоторый весьма узкий интервал частот $\Delta\nu$ (или длин волн $\Delta\lambda$), то обычно указывается средняя частота в пределах каждого интервала;

2) распределением полной энергии излучения между спектральными линиями.

Наиболее простой спектр испускается водородом. Удалось установить, что частоты спектральных линий водорода могут быть охвачены одной общей формулой

$$\nu = R_0\left(\frac{1}{k^2} - \frac{1}{n^2}\right), \tag{2.2}$$

где R_0 — некоторая постоянная, равная $3{,}290 \cdot 10^{15}$ с$^{-1}$, k и n — целые числа: $k = 1, 2, 3, \ldots$; $n = k + 1, k + 2, k + 3 \ldots$

Вместо частоты колебаний обычно используют *волновое число* $1/\lambda = \nu/c$, показывающее число длин волн, укладывающихся на единице длины — сантиметре или метре. Тогда формула (2.2) перепишется в виде

$$\frac{1}{\lambda} = R\left(\frac{1}{k^2} - \frac{1}{n^2}\right), \tag{2.3}$$

где величина $R = R_0/c$ называется *постоянной Ридберга*. Его численное значение определено с очень большой точностью; для атома водорода

$$R_{\text{H}} = 109677{,}581 \text{ см}^{-1}. \tag{2.4}$$

Из полного набора частот (2.2) можно выделить отдельные серии, лежащие в различных областях спектра. Условием $k = 1$, $n = 2, 3, 4\ldots$ выделяется *серия Лаймана*, лежащая в ультрафиолетовой области. Линии, лежащие в видимой области, — *серия Бальмера* — выделяются значениями $k = 2$, $n = 3, 4, 5, \ldots$ В инфракрасной области расположена *серия Пашена*, для которой $k = 3$, $n = 4, 5, 6, \ldots$ и т. д.

Таким образом, для разработки теории атома водорода можно было воспользоваться следующими представлениями:

1) атом водорода состоит из протона и электрона, вращающихся вокруг общего центра масс. Так как масса протона в 1836 раз больше массы электрона, то в первом приближении можно полагать, что центр массы атома совпадает с центром массы протона, а вращение электрона происходит по круговой орбите;

2) центростремительной силой, удерживающей электрон на орбите является кулоновское притяжение к протону:

$$\frac{mv^2}{r} = \frac{e^2}{4\pi\varepsilon_0 r^2} \tag{2.5}$$

(ε_0 — электрическая постоянная; см. ч. III, § 2);

3) полная энергия атома состоит из кинетической энергии электрона и потенциальной энергии взаимодействия электрона с протоном (потенциальная энергия является отрицательной; см. ч. I, § 12):

$$E = \frac{mv^2}{2} - \frac{e^2}{4\pi\varepsilon_0 r}. \tag{2.6}$$

При сопоставлении этой формулы с (2.5) выясняется, что полная энергия атома водорода имеет отрицательный знак и равна

$$E = -\frac{e^2}{8\pi\varepsilon_0 r} = -\frac{mv^2}{2}. \tag{2.7}$$

Однако эти три положения недостаточны для объяснения наблюдаемых свойств атома; потребовались дополнительные предположения. Дело в том, что, согласно законам электродинамики, вращающийся электрон должен излучать электромагнитные волны (см. ч. III, § 29) и поэтому продолжительное пребывание его на орбите невозможно. Постепенно теряя свою энергию, электрон должен упасть на протон. Кроме того, согласно формуле (2.5), по мере приближения к протону линейная и угловая скорости электрона будут увеличиваться. Вследствие этого частота излучения, определяемая частотой обращения электрона вокруг протона, должна была бы непрерывно возрастать. Следовательно, спектр излучения атома водорода окажется непрерывным, но не линейчатым.

Для того чтобы объяснить наблюдаемую устойчивость атома водорода и его линейчатый спектр, Н. Бор предложил следующие постулаты:

1) в атоме водорода существуют особые орбиты, вращаясь по которым электрон не излучает;

2) эти устойчивые («стационарные») орбиты выделяются следующим условием: момент импульса электрона по орбите равен целому числу $h/2\pi$ (h — постоянная Планка):

$$mvr = n\frac{h}{2\pi}, \tag{2.8}$$

где n — номер орбиты. Радиус первой орбиты ($n = 1$) есть наименьшее расстояние, до которого могут сблизиться электрон и протон. Если электрон находится на первой орбите, то энергия атома имеет минимально возможное значение E_1; это состояние атома называется *нормальным* или *основным*. Если же электрон вращается по другим, более удаленным от протона орбитам, то $E > E_1$ и состояние атома называется *возбужденным*. Таким образом, энергия атома водорода может иметь только определенные значения: E_1, E_2...;

3) излучение энергии атомом происходит при переходе электрона из одной орбиты (n) на другую ($k < n$), причем частота испускаемой электромагнитной волны зависит от разности энергии атома в начальном и конечном состояниях:

$$\nu = \frac{E_n - E_k}{h}. \tag{2.9}$$

Таким образом, атом водорода может излучать энергию только определенными монохроматическими «порциями», равными $h\nu$; первоначально их называли «квантами излучения», но в дальнейшем величина $h\nu$ стала рассматриваться как энергия элементарной частицы, названной *фотоном* (заслуга создания фотонной теории света принадлежит М. Планку и А. Эйнштейну).

Из этих представлений следует, что и поглощение энергии атомом должно происходить определенными дозами; для того чтобы перебросить электрон с ближних орбит на дальние, атому должна быть сообщена энергия, равная $(E_n - E_k)$. Эта энергия может быть получена атомом как при столкновениях с какими-нибудь частицами, так и при поглощении соответствующего «кванта излучения» $h\nu = E_n - E_k$. Таким образом легко объясняются результаты описанного выше опыта Франка и Герца ($\Delta E = 4{,}9$; 9,8 эВ).

Уравнения (2.5), (2.7) и (2.8) позволяют рассчитывать радиусы стационарных орбит, скорости электронов на этих орбитах и соответствующие значения энергии атома:

$$r_n = n^2 \frac{\varepsilon_0 h^2}{\pi m e^2}; \quad v_n = \frac{1}{n} \frac{e^2}{2\varepsilon_0 h}; \quad E_n = -\frac{1}{n^2} \frac{me^4}{8\varepsilon_0^2 h^2}. \tag{2.10}$$

Для первых трех орбит получаем:

$r_1 = 0{,}53 \cdot 10^{-10}$ м; $v_1 = 2{,}18 \cdot 10^6$ м/с; $E_1 = -13{,}55$ эВ;

$r_2 = 2{,}12 \cdot 10^{-10}$ м; $v_2 = 1{,}09 \cdot 10^6$ м/с; $E_2 = -3{,}37$ эВ;

$r_3 = 4{,}77 \cdot 10^{-10}$ м; $v_3 = 0{,}55 \cdot 10^6$ м/с; $E_3 = -1{,}50$ эВ.

Подставив значения энергии из (2.10) в формулу (2.9) для двух значений орбит n и k, получим полный набор частот или длин волн, которые должен (при соблюдении перечисленных выше предположений) испускать атом водорода:

$$\nu = \frac{me^4}{8\varepsilon_0^2 h^3}\left(\frac{1}{k^2} - \frac{1}{n^2}\right); \quad \frac{1}{\lambda} = \frac{me^4}{8\varepsilon_0^2 ch^3}\left(\frac{1}{k^2} - \frac{1}{n^2}\right). \tag{2.11}$$

Сравнивая с сериальной формулой (2.2), соответствующей реальному излучению, получаем для постоянной Ридберга теоретическое значение

$$R_\infty = \frac{me^4}{8\varepsilon_0^2 ch^3} = 109737{,}317 \text{ см}^{-1}.$$

Измеренные и теоретически полученные значения этой величины оказались очень близкими; это обстоятельство показывает, что использованные выше предположения, в том числе постулаты Бора, в основном правильны. Небольшое различие удалось устранить при учете совместного вращения электрона и протона вокруг общего центра масс O (рис. IV. 56). По определению, центр масс системы двух тел удовлетворяет условию (M — масса протона)

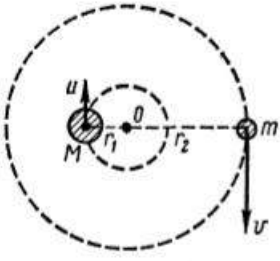


Рис. IV.56

$$\frac{r_1}{r_2} = \frac{m}{M}.$$

Кроме того, скорости электрона (v) и протона (u) связаны соотношением

$$Mu = mv; \quad \frac{u}{v} = \frac{m}{M}.$$

Равенство кулоновской и центростремительной сил запишется в виде, несколько отличном от (2.5):

$$\frac{mv^2}{r_2} = \frac{e^2}{4\pi\varepsilon_0 (r_1 + r_2)^2}.$$

По предположению Бора полный момент импульса атома $mvr_2 + Mur_1 = nh/(2\pi)$, тогда

$$r = r_1 + r_2 = n^2 \frac{\varepsilon_0 h^2}{\pi m e^2}\left(1 + \frac{m}{M}\right).$$

Полная энергия атома будет равна

$$E = \frac{mv^2}{2} + \frac{Mu^2}{2} - \frac{e^2}{4\pi\varepsilon_0 r} = -\frac{mv^2}{2}\left(1 + \frac{m}{M}\right) = -\frac{e^2}{8\pi\varepsilon_0 r}.$$

Подставив значение r_2, получим уточненное выражение для энергии атома водорода в состоянии, когда электрон находится на n-й орбите:

$$E = -\frac{1}{n^2}\frac{me^4}{8\varepsilon_0^2 h^2}\frac{1}{1 + m/M}. \qquad (2.12)$$

Это выражение отличается от (2.10) множителем $M/(M + m)$. С учетом этого множителя теоретически вычисленная постоянная Ридберга приобретает значение, почти равное экспериментально установленному (2.4).

Заметим, что при переходе из дальних орбит на ближние работа кулоновской силы (т. е. изменение потенциальной энергии атома) оказывается численно равной $2h\nu$; половина этой работы расходуется на увеличение кинетической энергии электрона (так как, согласно

формуле (2.10), скорость электрона на ближних орбитах больше, чем на дальних), а вторая половина расходуется на образование энергии испускаемого при этом фотона.

Однако внимательный анализ приведенного выше вывода сериальной формулы спектра водорода показывает, что при этом кроме перечисленных предположений используются также и некоторые другие, например:

1) утверждая, что источником энергии фотона и электрона является только потенциальная энергия взаимодействия протона и электрона, мы тем самым предполагаем, что *внутри атома никаких других процессов, которые могли бы выделять или поглощать энергию, нет*;

2). согласно теории относительности (см. Заключение, § 2), каждая частица с инертной массой m обладает энергией mc^2, причем масса физической системы всегда меньше суммы масс ее составных частей. При увеличении энергии системы на ΔW (например, при получении энергии извне) масса системы увеличивается на

$$\Delta m = \Delta W/c^2,$$

а при уменьшении энергии (например, при излучении) масса системы, согласно этому же соотношению, уменьшается. Ввиду этого масса атома водорода при излучении фотона с энергией $h\nu$ должна уменьшиться на $h\nu/c^2$. При выводе сериальной формулы (2.11) источником энергии фотона полагается потенциальная энергия взаимодействия, но не изменение массы атома, т. е. предполагается, что *при переходе электрона из дальних орбит на ближние масса электрона* (*а также и масса протона*) *остается неизменной*;

3) полагая, как это обычно делается, что энергия содержится не только у частиц вещества (mc^2), но и в окружающем их поле, полную энергию атома водорода следовало бы представить не в виде (2.6), а в виде

$$E = Mc^2 + mc^2 + U,$$

где M и m — массы протона и электрона, U — энергия связанных с ними полей. Так как электрон имеет скорость и поэтому вокруг него существует кроме электрического еще и магнитное поле, то энергия U должна быть рассчитана по формуле

$$U = \int \left(\frac{\varepsilon_0 E^2}{2} + \frac{\mu_0 H^2}{2} \right) dV,$$

где dV — элемент объема, в пределах которого напряженности электрического и магнитного полей равны E и H. При интегрировании должны быть исключены объемы, занимаемые самими частицами; если полагать, что электрон и протон имеют сферические формы, то результат интегрирования (см. ч. III, § 7) будет зависеть от их радиусов. Поэтому дополнительно к перечисленным выше предположениям необходимо *условие постоянства размеров протона и электрона*; кроме того, полагая, что потенциальная энергия в атоме обусловлена только электрическим взаимодействием, *игнорируется энергия, содержащаяся в магнитном поле.*

Сериальная формула (2.3) или (2.11) охватывает лишь длины волн в спектре излучения атома водорода. Вопрос об интенсивности отдельных спектральных линий не может быть решен на основании рассуждений, приведших к этой формуле, и должен обсуждаться отдельно. Допустим, что имеются N одинаково возбужденных атомов водорода, у которых электрон находится на n-й орбите (будем полагать, что n очень велико). Переход в нормальное состояние у различных атомов может происходить различным образом. Возможны переходы с n-й орбиты сразу на первую, но возможны также «ступенчатые переходы» любого характера, например с n-й орбиты на вторую, третью и т. д., с последующим переходом со второй на первую, с третьей на вторую или сразу на первую и т. д. Так как все орбиты — стационарные, то, находясь на них, электрон не излучает и поэтому в каждом из возбужденных состояний атом может находиться любое время. Ввиду этого переход электрона с дальних орбит на ближние следует рассматривать как случайный процесс. Среднее время пребывания электрона на дальних орбитах может быть обусловлено либо нестабильностью внутреннего состояния в атоме, либо внешними воздействиями случайного характера (упругие столкновения с другими атомами, флуктуирующие внешние электрические поля и т. п.). Однако случайный характер электронных переходов вовсе не означает их одинаковую вероятность, поэтому вопрос об относительной интенсивности спектральных линий связан с относительной вероятностью соответствующих переходов, рассчитанных, разумеется, для одинаковых внешних условий, в которых находятся излучающие атомы.

§ 10. КОРПУСКУЛЯРНО-ВОЛНОВЫЕ СВОЙСТВА ЧАСТИЦ. ФОТОЭЛЕКТРИЧЕСКИЙ ЭФФЕКТ. ФОРМУЛА ДЕ БРОЙЛЯ. ДИФРАКЦИЯ ЭЛЕКТРОНОВ

Существование волновых и корпускулярных свойств у физических объектов было сначала обнаружено в световых явлениях. В гл. 1 были рассмотрены оптические явления, которые объяснялись на основе волновой (электромагнитной) теории; это были в основном явления распространения света и взаимодействия световой волны с макроскопическими телами: линзами, призмами, дифракционными решетками, телами, отражающими или поглощающими энергию световой волны, и т. д. Однако в некоторых явлениях, течение которых определяется взаимодействием света с микрофизическими объектами — отдельными заряженными частицами, атомами или молекулами, обнаруживается расхождение между предсказаниями волновой теории и результатами наблюдений и измерений. Например:

1) согласно волновой теории, фронт электромагнитной волны, излучаемой отдельным атомом, на больших расстояниях от него должен иметь сферическую форму, а энергия этой волны (в вакууме или в однородной среде) должна быть равномерно распределена по направлениям; ввиду этого плотность энергии в волне должна убывать обратно пропорционально квадрату расстояния.

Ниже будет описан один из экспериментов — опыт А. Ф. Иоффе и Н. И. Добронравова, не подтверждающий этого положения волновой

теории; его удовлетворительное объяснение было получено на основе предположения, что атомы излучают не непрерывную световую волну, а отдельные «световые частицы» (фотоны);

2) по волновой теории, если световая (электромагнитная) волна проходит через вещество, то заряженные частицы, содержащиеся в этом веществе (электроны, ионы), могут поглощать энергию волны в *любых количествах* (любыми дозами) независимо от длины этой волны. Частота колебаний в волне имеет значение только в явлениях *резонансного* поглощения, когда собственная частота колебаний заряженных частиц в атомах, молекулах или в узлах кристаллической решетки вещества совпадает с частотой действующей на них световой волны.

Однако этому представлению противоречат законы фотоэлектрического эффекта, рассмотренные ниже. Оказывается, свободные электроны, содержащиеся в металлах, поглощают энергию падающего на них излучения не любыми дозами, а только в количествах, равных $h\nu$, где ν — частота колебаний этого излучения.

Таких примеров можно привести много. Они показывают, что волновая теория необходима только для объяснения распространения света (отражения, преломления, дифракции). Но если попытаться применить эту теорию к явлениям взаимодействия излучения с веществом, то обнаруживаются явные расхождения с экспериментальными результатами. Ниже рассмотрим некоторые из этих экспериментов, наглядно показывающих необходимость использования корпускулярных представлений о свете как о потоке особых частиц — фотонов.

ФОТОЭЛЕКТРИЧЕСКИЙ ЭФФЕКТ

Допустим, что на поверхность металла падает монохроматическое излучение с частотой ν, причем в поверхностном слое металла амплитуда напряженности электрического поля волны равна E. В результате поглощения энергии волны из поверхности металла вылетают электроны с различными скоростями. Измерения показали, что:

1) число электронов N, вылетающих с единицы поверхности металла в единицу времени, пропорционально интенсивности φ падающего излучения (или поверхностной плотности потока излучения, см. ч. IV, § 1), следовательно, пропорционально объемной плотности энергии в падающей волне $w = \varepsilon_0 E^2/2$;

2) максимальная скорость, которая обнаруживается у вылетающих электронов, не зависит от интенсивности φ (или плотности энергии w) падающей волны, но *зависит от частоты колебаний* этой волны и от металла, подвергающегося облучению;

3) фотоэффект наблюдается только при частотах падающей волны, превышающих некоторое *минимальное* значение ν_0, которое, однако, различно для различных металлов;

4) фотоэффект возникает и исчезает почти *одновременно* с началом и прекращением облучения; расхождение во времени не превышает 10^{-9} с.

Перечисленные законы «внешнего» фотоэффекта не могут быть объяснены волновой теорией света. Только первый закон согласуется с этой

теорией: чем больше интенсивность падающего излучения, тем большее число электронов получат энергию, необходимую для выхода из металла (эта энергия должна превышать «работу выхода» A, указанную в ч. III, § 12). Разумеется, не вся энергия излучения, ежесекундно поглощаемая металлом (W), передается вылетающим электронам. Допустим, что каждый из N вылетающих электронов получает от падающей волны одну и ту же энергию w; тогда отношение

$$\eta = \frac{Nw}{W}$$

будет показывать «коэффициент полезного действия» фотоэффекта. Если этот коэффициент остается постоянным в некоторых пределах изменения энергии W, то число вылетающих электронов N будет прямо пропорционально этой энергии.

Однако объяснение остальных законов фотоэффекта волновой теорией оказывается невозможным. Согласно этой теории, свободные электроны, имеющиеся в объеме металла, раскачиваясь под действием электрического поля волны E, могут поглощать ее энергию независимо от амплитуды E_0 и частоты колебаний в этой волне. Процесс накапливания необходимой энергии электронами потребует некоторого времени, которое, очевидно, будет зависеть от амплитуды напряженности E. С уменьшением интенсивности падающего излучения φ уменьшается значение напряженности поля E и, следовательно, время, необходимое для достижения электронами энергии A, будет увеличиваться; при очень слабых интенсивностях это время может оказаться очень большим, что явно противоречит упомянутому выше отсутствию расхождения во времени между началом облучения и вылетом электронов. Кроме того, на основе волновой теории не удается найти объяснение *независимости* максимальной скорости электронов от интенсивности облучения, а также ее *зависимости* от частоты колебаний. Невозможно объяснить также и третий закон: если электромагнитная волна содержит в себе энергию и может передавать эту энергию электронам, то непонятно, почему процесс передачи энергии может происходить при одних частотах ($\nu > \nu_0$) и совершенно не может осуществляться при других, даже очень близких частотах ($\nu < \nu_0$).

На основании фотонной теории (согласно которой свет представляет собой поток особых частиц, каждая из которых имеет энергию $h\nu$, зависящую от частоты колебаний) эти законы получают простое объяснение, хотя и с помощью некоторых дополнительных предположений. Первый закон фотоэффекта и в этой теории выражает только постоянство указанного выше «коэффициента полезного действия» фотоэффекта. Остальные законы объясняются на основании предположения, что каждый вылетевший электрон приобретает необходимую ему энергию $w > A$ при «поглощении» *только одного фотона* (вероятность поглощения одним электроном двух фотонов и более в рассматриваемом явлении следует полагать исключительно малой). Тогда на основании закона сохранения энергии можно написать соотношение, предложенное А. Эйнштейном

$$h\nu = A + mv^2/2. \qquad (2.13)$$

Из этой формулы следует, что максимальная скорость электронов, вылетающих из металла, определяется частотой колебаний падающей волны и работой выхода электронов из металла, подвергнутого облучению:

$$v = \sqrt{\frac{2(h\nu - A)}{m}},$$

а минимальное значение частоты ν_0, при которой еще возможен фотоэффект, определяется только работой выхода. Отсутствие фотоэффекта означает $v = 0$, поэтому

$$h\nu_0 = A; \quad \nu_0 = A/h; \quad \lambda_0 = c/\nu_0 = hc/A. \tag{2.14}$$

Частота ν_0 и соответствующая длина волны λ_0 называются *предельными*, *граничными* или *красной границей фотоэффекта*. Последнее название обусловлено тем, что у многих металлов ν_0 и λ_0 лежат в области красного света.

Отсутствие расхождения во времени между началом облучения и появлением вылетающих электронов следует объяснить тем, что процесс «поглощения» фотона электроном происходит в течение малого времени; для согласия с экспериментами это время должно иметь порядок 10^{-9} с.

Интересный опыт, который подтверждает фотонную теорию и не может быть объяснен волновой теорией, был выполнен А. Ф. Иоффе и Н. И. Добронравовым в 1922 г. Схема опыта показана на рис. IV.57. В электрическом поле плоского конденсатора *А—Б* была уравновешена заряженная пылинка из висмута *В*. Нижняя обкладка конденсатора была сделана из очень тонкой алюминиевой фольги. Между этой обкладкой и катодом *К*, испускающим электроны *е*, было приложено высокое напряжение *V*, достаточное для того, чтобы электроны, бомбардирующие фольгу, возбуждали в ней рентгеновское излучение. В каждую секунду из катода на фольгу в среднем падало определенное число электронов, поэтому рентгеновское излучение из фольги должно было состоять примерно из такого же числа импульсов в секунду. Наблюдение за пылинкой показало, что в среднем каждые 30 мин она скачком выходила из равновесия, т. е. теряла заряд. Очевидно, что это явление есть результат фотоэффекта — испускания электронов висмутовой пылинкой под действием падающего на нее рентгеновского излучения от фольги. Согласно волновой теории, энергия рентгеновского импульса должна полагаться равномерно распределенной по фронту волны, т. е. по всем направлениям от той точки, куда попал электрон, вызвавший этот импульс. Поэтому пылинка будет получать от волны очень малую энергию, пропорциональную квадрату диаметра пылинки (площадь сечения) и обратно пропорциональную квадрату расстояния от пылинки до той точки, откуда исходило рентгеновское излучение. Следовательно, придется предположить, что в течение 30 мин поглощенная всей пылинкой энергия волны передается только одному из электронов, который таким образом «накапливает» энергию и тем самым приобретает возможность покинуть пылинку; «непонятным» здесь остается тот процесс, при помощи которого возможно такое сосредоточение поглощенной малыми дозами энергии волны в одном из электронов пылинки. Согласно же фотонной теории, фольга излучает не непрерывную электромагнитную волну, а поток «рентгеновских» фотонов; поэтому следует

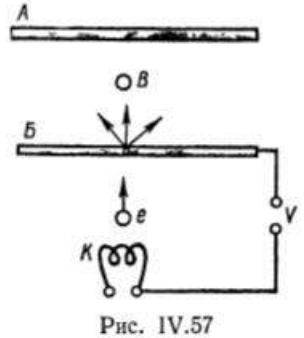


Рис. IV.57

рассчитывать (в соответствии с геометрическими размерами установки и производительностью катода) вероятность попадания в пылинку одного из этих фотонов, испускаемых фольгой; такой расчет оказался в полном соответствии с результатом наблюдений.

ЭФФЕКТ КОМПТОНА

В фотоэлектрическом эффекте фотон «поглощается» электроном проводимости металла и отдает ему всю свою энергию. Однако возможно и такое взаимодействие фотона с электроном, при котором существование фотона не прекращается, т. е. происходит рассеяние фотона на электроне и лишь часть его энергии передается электрону. Схема соответствующего опыта приведена на рис. IV. 58, *а*, где обозначены: *А* — источник монохроматического рентгеновского излучения, *Б* — свинцовые диафрагмы, выделяющие узкий пучок лучей, *В* — тело, содержащее свободные или слабо связанные электроны и прозрачное для

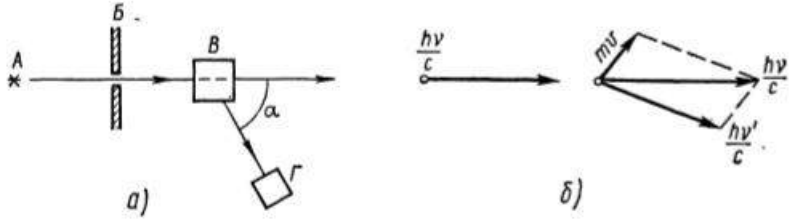


Рис. IV.58

используемого излучения (кусок парафина или графита), *Г* — рентгеновский спектрометр, измеряющий интенсивность и частоту (или длину волны) излучения, рассеиваемого телом *В* под различными углами α. Измерения показывают (эффект Комптона), что рассеянное излучение содержит не только частоту падающей волны ν, но и несколько меньшую частоту ν', причем оказалось, что разность $\nu - \nu'$ зависит от угла α, под которым производилось измерение, и не зависит ни от частоты ν, ни от рассеивающего вещества. Если вместо частот колебаний ν и ν' использовать соответствующие длины волн λ и λ', то, согласно измерениям,

$$\Delta\lambda = \lambda' - \lambda = 0{,}02424\,(1 - \cos\alpha)\cdot 10^{-10}\ \text{м}. \qquad (2.15)$$

По волновой теории, электроны, содержащиеся в рассеивающем веществе, совершая вынужденные колебания под действием электрического поля падающей волны, должны сами излучать электромагнитные волны той же частоты, поэтому в рассеянном излучении других частот не должно было быть. На основании волновых представлений не удается найти также и объяснение других особенностей эффекта Комптона: независимости $\Delta\lambda$ от λ, от рассеивающего вещества и характер зависимости этой величины от угла рассеяния.

Для объяснения эффекта Комптона по фотонной теории предполагается, что рассеяние рентгеновского излучения есть результат упругого столкновения фотонов этого излучения с частицами вещества —

протонами, электронами и т. п. Рассмотрим сначала столкновение фотона, имеющего энергию $h\nu$ и импульс $h\nu/c$, со свободным электроном (или слабо связанным электроном, если энергия, необходимая для освобождения электрона, значительно меньше $h\nu$). В результате удара фотон теряет часть своей энергии и рассеивается под углом α к первоначальному направлению; это означает, что частота рассеянного фотона будет меньше, чем у падающего, а длина волны $\lambda' > \lambda$. Электрон, если до удара полагать его покоящимся, получит после удара импульс mv и энергию $mv^2/2$. На основании законов сохранения энергии и импульса (рис. IV. 58, *б*) имеем:

$$h\nu - h\nu' = \frac{mv^2}{2}; \quad (mv)^2 = \left(\frac{h\nu}{c}\right)^2 + \left(\frac{h\nu'}{c}\right)^2 + 2\frac{h\nu}{c}\frac{h\nu'}{c}\cos\alpha. \quad (2.16)$$

Решая эти уравнения, получим

$$\nu - \nu' = \frac{h}{mc^2}\nu\nu'(1 - \cos\alpha). \quad (2.17)$$

Если перейти к длинам волн, то

$$\Delta\lambda = \lambda' - \lambda = \frac{h}{mc}(1 - \cos\alpha). \quad (2.18)$$

Подставляя значения постоянной Планка, массы электрона и скорости света, получим

$$\Delta\lambda = 0{,}02426\,(1 - \cos\alpha)\cdot 10^{-10}\ \text{м},$$

что совпадает с результатами измерений. В этом расчете мы предположили, что при столкновении «скорость отдачи» электрона v мала по сравнению со скоростью света. При большой энергии ударяющего фотона (вернее, разности энергий $h\nu - h\nu'$) скорость v может оказаться близкой к скорости света и поэтому в согласии с теорией относительности А. Эйнштейна (см. Заключение, § 2) изменение энергии электрона должно записываться в виде $mc^2 - m_0c^2$, где $m = m_0/\sqrt{1 - v^2/c^2}$; если $v \ll c$, то $mc^2 - m_0c^2 \approx mv^2/2$.

Формулы (2.15) и (2.18) показывают, что комптоновское изменение длины волны $\Delta\lambda$ очень мало и не зависит от длины волны падающего излучения. Для возможности наблюдения эффекта важна не величина $\Delta\lambda$, а относительное изменение длины волны $\Delta\lambda/\lambda$; для очень коротких рентгеновских и гамма-лучей (первые опыты производились при $\lambda = 0{,}7126\cdot 10^{-10}$ м) это отношение порядка 5—10%, но для видимого света оно составляет тысячные доли процента и может быть обнаружено только очень чувствительной аппаратурой.

Существование в рассеянном излучении фотонов с первоначальной частотой ν объясняется тем, что часть фотонов соударяется с сильно связанными электронами, находящимися внутри атомов. При таких столкновениях «отдачу» испытывает уже не один электрон, а весь атом в целом; так как масса атома значительно больше массы электрона, то, по закону упругого удара, фотон передаст атому лишь незначительную часть своей энергии и поэтому почти не изменит свою частоту колебаний.

Наблюдая столкновение фотонов с электронами в специальной камере, в которой каждая движущаяся заряженная частица оставляет видимый след («треки частиц» в камере Вильсона), можно определить энергию, полученную электроном при этом столкновении, а также величину и направление его импульса. Результаты измерений также оказались в согласии с приведенными выше расчетами.

Таким образом, полное объяснение всех деталей эффекта Комптона на основе фотонной теории и законов сохранения можно рассматривать, во-первых, как новое подтверждение самого существования фотонов, во-вторых, как доказательство правильности выражений для их энергии ($h\nu$) и импульса ($h\nu/c$).

ФОТОН И ЭЛЕКТРОМАГНИТНАЯ ВОЛНА

Нельзя не заметить, что в приведенных выше объяснениях спектра излучения водорода, фотоэлектрических явлений и эффекта Комптона, совершенно игнорируется непрерывно распределенная в пространстве энергия электромагнитной в о л н ы. Единственная связь между фотонами и электромагнитной волной заключается в частоте колебаний: у фотонов эта частота определяет величину их энергии $h\nu$ и импульса $h\nu/c$, а у электромагнитной волны — длину волны λ, необходимую для расчета дифракционных эффектов (см. ч. IV, § 5). Создается впечатление, что в световой (электромагнитной) волне, которая излучается о д н и м а т о м о м, энергия не распределена по пространству в соответствии со значениями E в каждой точке этой волны. Для объяснения указанных выше опытных данных приходится предполагать, что энергия электромагнитной волны сосредоточена в особых частицах, которые перемещаются вместе с фронтом волны со скоростью света. Другими словами, в световом потоке фотоны должны полагаться «частицами» электромагнитной волны, содержащими в себе энергию этой волны. Однако такое представление вызывает ряд вопросов: существует ли напряженность электромагнитного поля в пространстве между фотонами (если существует, то придется допустить, что величина $w = \varepsilon_0 E^2/2$ не есть энергия); существует ли взаимодействие между самими фотонами и в каких явлениях оно проявляется; каким образом волна «связана» с фотонами и определяет их траекторию и т. д. Эти вопросы обсуждаются и решаются в рамках одного из важных разделов современной теоретической физики — квантовой электродинамики. Здесь же важно подчеркнуть дуалистический характер существующих представлений о свете: необходимость волновой теории для объяснения распространения света и необходимость корпускулярной (фотонной) теории для объяснения элементарных процессов излучения и поглощения света.

Если полагать, что фотон является «элементарной частицей», то излучение или поглощение света атомами, а также фотоэлектрический эффект и комптоновское рассеяние являются, по существу, процессами передачи энергии от одних частиц к другим. В этих процессах строго соблюдаются законы сохранения энергии и импульса взаимодействующих частиц; мы утверждаем, что здесь проявляются к о р п у с к у-

л я р н ы е свойства этих частиц. Однако если нам необходимо объяснить траектории фотонов в явлениях дифракции и интерференции, то мы можем совершенно не интересоваться энергиями и импульсами; эти задачи решаются на основании правила Гюйгенса и принципа суперпозиции векторов **E** и **H** электромагнитных волн. При этом нас не смущает то обстоятельство, что в окрестности какой-либо точки два вектора $\mathbf{E}_1 = \mathbf{E}_2$ и связанные с ними энергии могут «уничтожить» друг друга или, складываясь, образовать из энергии $\varepsilon_0 E_1^2/2$ и $\varepsilon_0 E_2^2/2$ вдвое большую энергию $\varepsilon_0(E_1 + E_2)^2/2$ (см. ч. IV, § 4). Лишь после того, как определены суммарные значения вектора **E** в каждой точке пространства (или экрана), можно перейти от напряженности поля к энергии поля и получить распределение энергии вдоль фронта результирующей волны в согласии с требованием закона сохранения энергии. В необходимости расчета распределения энергии в пространстве или на экране на основе принципа Гюйгенса и принципа суперпозиции мы видим проявление в о л н о в ы х свойств фотонов.

Из формул $\varepsilon_\Phi = h\nu = m_\Phi c^2$, $p_\Phi = h\nu/c = m_\Phi c$, характеризующих корпускулярные и волновые свойства фотонов, может быть получено важное соотношение

$$\lambda = h/m_\Phi c, \tag{2.19}$$

где $\lambda = c/\nu$ — длина электромагнитной волны, «связанной» с данными фотонами.

ФОРМУЛА ДЕ БРОЙЛЯ

В 1924 г. Луи де Бройль высказал предположение, что корпускулярно-волновой дуализм (т. е. одновременное наличие корпускулярных и волновых свойств), возможно, существует и у частиц вещества, в частности у электронов. Предложенная им формула, связывающая эти свойства, была аналогична формуле (2.19), написанной для фотонов:

$$\lambda = h/mv. \tag{2.20}$$

Существенное различие между формулами (2.19) и (2.20) заключается в том, что:

1) у фотонов нет инертной массы п о к о я ($h\nu/c^2$ приравнивается m_Φ на основании общего соотношения между массой и энергией, установленного в теории относительности, т. е. характеризует «запас энергии» в фотоне). Электроны же имеют массу покоя m_0, причем в формуле (2.20)

$$m = m_0/\sqrt{1 - v^2/c^2};$$

2) скорость фотонов (в вакууме) является постоянной величиной, тогда как скорость электронов может иметь различные значения.

Первое подтверждение формулы де Бройля (2.20) было получено Девиссоном и его сотрудниками (1927) при исследовании отражения пучка электронов от поверхности металлов: были обнаружены чередующиеся максимумы и минимумы по-видимому, обусловленные дифракцией некоторой волны, отраженной от плоскостей кристаллической

решетки облучаемого металла. Расчеты длины этой волны оказались в полном согласии с формулой де Бройля. Это совпадение было подтверждено также и другими опытами, поставленными по аналогии с оптическими опытами по дифракции. Впоследствии выяснилось, что волновыми свойствами обладают также и другие частицы вещества — протоны, нейтроны и т. д.; для них также была подтверждена формула де Бройля (2.20).

ДИФРАКЦИЯ ЭЛЕКТРОНОВ

Для того чтобы выяснить сущность *корпускулярно-волнового дуализма в свойствах частиц*, рассмотрим идеальный опыт, имеющий принципиальное значение. Допустим, что на дифракционную решетку с периодом d падает параллельный пучок

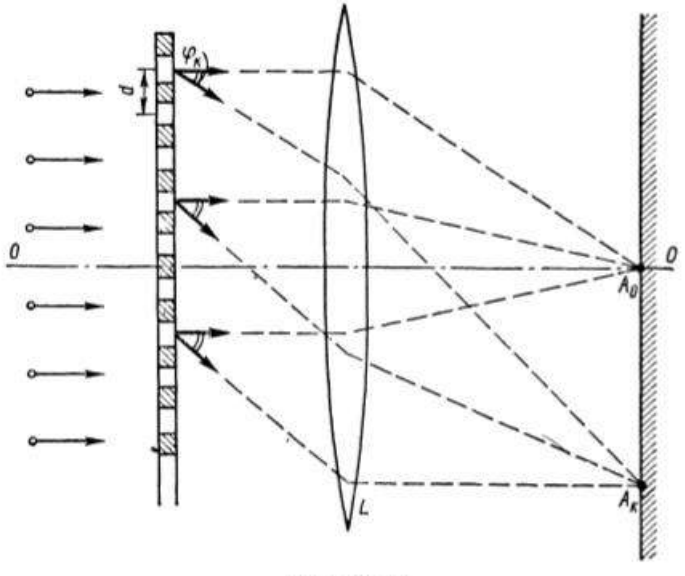


Рис. IV.59

одинаковых частиц, имеющих одинаковые по величине и направлению скорости v (рис. IV. 59). За решеткой поместим прибор L, сортирующий частицы по направлениям (для фотонов таким прибором будет линза). В точке A_0 будут собраны все частицы, которые пройдут через щели со скоростями, параллельными O—O; в точке A_k соберутся все частицы, отклонившиеся на угол φ_k, и т. д. Согласно результатам экспериментов, при прохождении частиц через щели наблюдается отклонение их от первоначального направления (дифракция), а на экране образуется интерференционная картина из чередующихся максимумов (скопление частиц) и минимумов (отсутствие или малое число их). Углы φ_k, соответствующие максимумам, удовлетворяют условию

$$d \sin \varphi_k = k\lambda,$$

где $k = 0, 1, 2, \ldots$, причем в этой формуле λ зависит от массы частиц и их скорости: $\lambda = h/mv$, h — постоянная Планка.

Можно было бы попытаться объяснить дифракцию частиц, проходящих через решетку, тем, что в потоке частиц имеется какой-то волновой процесс (т. е. предположить, что волновые явления обусловлены не свойствами отдельных частиц, а свойствами всего потока частиц). Однако правильность такого объяснения может быть проверена путем дополнительных исследований. В этом вопросе важное значение имеют результаты, полученные В. А. Фабрикантом и его сотрудниками (1949): чередующиеся максимумы и минимумы на экране получаются не только в том случае, когда на решетку падает пучок, содержащий очень большое число электронов, но и при прохождении через решетку *одного электрона за другим через большие промежутки времени*. Эти промежутки были настолько большими, что взаимодействие между электронами, последовательно проходящими через решетку, практически исключалось. Так как дифракционная картина на экране образуется не только заряженными (электроны, протоны), но и нейтральными (нейтроны) частицами, то при объяснении дифракции частиц должно быть исключено также и взаимодействие между ними и атомами вещества в щелях решетки.

Опыты В. А. Фабриканта по дифракции электронов, последовательно проходящих через решетку с большими интервалами времени, можно повторить и для фотонов. Допустим, что на дифракционную решетку ежесекундно падает $3 \cdot 10^8$ фотонов (при длине волны 0,6 микрон это соответствует потоку излучения, равному 0,001 эрг/с). Полагая, что фотоны представляют собою частицы, практически локализованные в точке, а в рассматриваемом потоке они следуют один за другим через равные промежутки времени, найдем расстояние между ними:

$$l = \frac{3 \cdot 10^{10}}{3 \cdot 10^8} = 100 \text{ см}.$$

Это расстояние может превосходить размеры дифракционной установки, поэтому можно считать, что каждый последующий фотон поступает в прибор после того, как предыдущий фотон покинул его и достиг экрана. Вследствие этого взаимодействие между электромагнитными волнами, сопровождающими отдельные фотоны, будет исключено. Появление на экране интерференционных максимумов и минимумов от фотонов, последовательно проходящих через решетку, уже нельзя будет объяснить как результат сложения векторов напряженности какой-то общей электромагнитной волны, прошедшей через разные щели и собранной линзой на экране. Таким образом, можно отметить аналогию между описанием поведения частиц вещества при помощи дебройлевских волн и описанием поведения фотонов при помощи электромагнитных волн.

Изложенные результаты показывают, что *волновыми свойствами обладает каждая частица в отдельности*. Эти результаты не дают основания для з а м е н ы частиц каким-нибудь непрерывным волновым процессом, происходящим в трехмерном пространстве. Выше, на

примере оптичееких явлений, было показано, что отдельные фотоны, испускаемые атомом, не могут быть заменены электромагнитной волной, распространяющейся от этого атома по всем направлениям. Еще большее затруднение возникает при объяснении волновых свойств электронов, протонов, нейтронов и других частиц: здесь нам вообще неизвестны какие-либо волны в трехмерном пространстве, которые могли бы иметь по отношению к этим частицам хотя бы такое же значение, какое имеют электромагнитные волны по отношению к фотонам.

Следует также заметить, что при прохождении волны через многочисленные щели дифракционной решетки эта волна разделяется на части, которые можно (после решетки) развести в различные стороны. Отдельная же частица может пройти только через одну из щелей, поэтому такое «разделение на части», которое возможно для волны, невозможно для частицы. Далее, образование интерференционных максимумов и минимумов на экране объясняется взаимным усилением или ослаблением когерентных частей волны, прошедших через разные щели решетки и собранных прибором L в том или другом месте экрана; для частиц вещества (а впрочем, и для фотонов) такое объяснение («интерференция частиц») недопустимо. Проблема заключается в том, чтобы выяснить, почему частицы «предпочитают» отклоняться после решетки под углами φ_k, соответствующими максимумам, и почему они «избегают» направлений, соответствующих минимумам; при этом, разумеется, должно найти объяснение как существование волны де Бройля, так и зависимость длины этой волны от импульса частиц.

Рассуждения, приведенные выше, показывают всю сложность взаимной связи между корпускулярными и волновыми свойствами всех частиц, включая и фотоны. Следует только подчеркнуть, что корпускулярные и волновые свойства частиц неотделимы друг от друга и несводимы друг к другу. Поведение частиц в любых физических условиях определяется одновременно и корпускулярными, и волновыми свойствами. Однако в одних условиях это поведение в основном определяется корпускулярными свойствами, а волновые играют второстепенную роль, формируя несущественные детали изучаемого процесса. В других условиях главное значение приобретают волновые свойства, определяющие основные черты изучаемого явления; корпускулярные же свойства имеют побочное значение и используются для выяснения некоторых деталей этого процесса.

§ 11. ОСНОВНЫЕ ПОЛОЖЕНИЯ КВАНТОВОЙ ТЕОРИИ. УРАВНЕНИЕ ШРЕДИНГЕРА. СООТНОШЕНИЕ НЕОПРЕДЕЛЕННОСТЕЙ ГЕЙЗЕНБЕРГА

При теоретической обработке результатов наблюдений и измерений всегда используется ряд *предположений*, упрощающих объекты исследований. На основе таких предположений создаются идеализированные *модели* изучаемых предметов и процессов, облегчающие (а иногда и открывающие возможность) применения математических методов исследования. Такие модели широко исполь-

зуются также и в атомной физике. Использование упрощающих предположений и идеализированных моделей неизбежно; они дают первое приближение к объективной природе и поэтому имеют в науке временное значение. В дальнейшем, с развитием науки, они уточняются, а иногда заменяются новыми, более соответствующими объективной природе. Во многих случаях уточнение или замена используемых предположений не сопровождается существенным изменением уже сложившихся физических представлений. Однако в истории физики были отмечены периоды, когда необходимость замены устаревших, потерявших силу предположений и гипотез приводила к коренному изменению общих представлений об устройстве природы и о содержании установленных законов физики. Таким был период (1900—1927) возникновения и развития квантовой физики.

В § 10 было указано, что при описании явлений, в которых участвуют мельчайшие частицы вещества — электроны, протоны, нейтроны и др., на основе представлений и законов классической физики (механики, электродинамики, волновой оптики и т. д.) встретились затруднения, оказавшиеся непреодолимыми. Для объяснения новых явлений (фотоэффект, дифракция электронов и других частиц и т. д.) потребовались новые представления, которые не укладывались в рамки классической физики, явно противоречили ее основным положениям. Со временем отдельные разрозненные предположения и гипотезы, возникшие в различных областях атомной физики, были связаны между собой и привели к формированию единой физической теории, по существу нового физического мировоззрения, получившего название квантовой физики.

Важнейшими свойствами квантовых объектов являются следующие:

1) существование у частиц корпускулярных и волновых свойств, неотделимых друг от друга и несводимых друг к другу;

2) существование у физических систем дискретного спектра устойчивых состояний.

Корпускулярные свойства заключаются в том, что каждая частица имеет некоторую сосредоточенную в малом объеме энергию и импульс; при взаимодействиях частиц между собой соблюдаются законы сохранения энергии и импульса.

Волновые свойства заключаются в том, что траектория частицы определяется некоторой связанной с ней волной, распространение которой подчинено принципу Гюйгенса и для которой соблюдается принцип суперпозиции. Однако это утверждение требует расшифровки: какова природа этих волн; какая физическая величина характеризует эти волны и изменяется в соответствии с волновым уравнением; каким образом поведение частицы связано с значениями этой величины в различных точках пространства, т. е. как взаимодействуют между собой «волна» и «частица». Заметим, что для одной из частиц — фотона — энергия связана с частотой колебаний, а импульс частицы — с длиной волны.

Известно, что фотоны связаны с электромагнитной волной; в уравнение этой волны (см. ч. III, § 29) входят напряженности электрического и магнитного полей и величины ε_0 и μ_0, характеризующие среду,

в которой распространяется эта волна. Величина

$$w = \frac{\varepsilon_0 E^2}{2} + \frac{\mu_0 H^2}{2},$$

которая трактуется как плотность энергии в электромагнитной волне, может рассматриваться также как величина, определяющая число фотонов (а следовательно, и их суммарную энергию) в единице объема. Если через дифракционную решетку проходит один-единственный фотон (и связанное с ним электромагнитное поле), то значение w, рассчитанное для различных точек экрана, следует рассматривать как величину, определяющую вероятность попадания этого фотона в ту или иную точку экрана; тогда при прохождении через решетку очень большого числа фотонов (одновременно или через некоторые интервалы времени) они окажутся распределенными по экрану в соответствии со значениями w.

УРАВНЕНИЕ ШРЕДИНГЕРА

Однако для электронов, протонов, нейтронов и других частиц волны, которые были бы ответственны за дифракцию и в то же время доступны непосредственному изучению при помощи соответствующей физической аппаратуры, не обнаружены. Поэтому для описания волновых свойств этих частиц и, в частности, их дифракции можно ввести некоторую, пока неизвестную, величину ψ, подчиненную стандартному волновому уравнению (см. ч. III, § 29). По аналогии с вектором **E** электромагнитной волны, будем трактовать $\psi^2 dV$, рассчитанную для объема dV, как величину, определяющую вероятность нахождения частицы в пределах этого объема. Для простоты воспользуемся уравнением для амплитуды плоской гармонической волны, распространяющейся вдоль оси OX:

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{4\pi^2}{\lambda^2} \psi = 0. \qquad (2.21)$$

В это «волновое уравнение» необходимо «вложить» физические свойства той частицы, поведение которой мы собираемся описывать величиной ψ. Очевидно, что для этой цели необходимо использовать соотношение, которое связывает волновые и корпускулярные свойства частиц, т. е. формулу де Бройля. Подставив в (2.21) $\lambda = h/mv$, получим

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{4\pi^2 m^2 v^2}{h^2} \psi = 0. \qquad (2.22)$$

Однако это уравнение было бы достаточно только для выяснения траектории свободной частицы, например, при ее прохождении через щель, дифракционную решетку или при отражении от плоскостей кристаллов и т. п. Если же на частицы действуют какие-нибудь внешние силы (например, если электрон движется во внешнем для него электрическом или магнитном поле — в ускорителях, измерительных приборах и т. д.), то поведение частиц будет определяться также и теми внешними полями, которые действуют на данную частицу. Если мы желаем,

чтобы уравнение (2.22) учитывало воздействие на частицы этих внешних полей, то можно поступить следующим образом. Полная энергия E частицы, движущейся в силовом поле, есть сумма кинетической энергии $W = mv^2/2$ и потенциальной энергии $U = e\varphi$, где e — заряд частицы (соответствующий этому полю), а φ — потенциал в данной точке поля:

$$E = W + U.$$

При перемещении частицы в поле все три величины E, W и U будут переменными. Воспользуемся тем, что волновое уравнение (2.22) содержит кинетическую энергию $mv^2/2$, и произведем замену:

$$\frac{mv^2}{2} = E - U.$$

Тогда окончательно

$$\frac{\partial^2\psi}{\partial x^2} + \frac{8\pi^2 m}{h^2}(E - U)\psi = 0.$$

В общем случае, когда решается не плоская задача, т. е. когда необходимо рассчитывать движение частиц в трех измерениях, вместо $\partial^2\psi/\partial x^2$ будет написана сумма $\partial^2\psi/\partial x^2 + \partial^2\psi/\partial y^2 + \partial^2\psi/\partial z^2$, обозначаемая $\Delta\psi$ и называемая «лапласианом» функции ψ:

$$\Delta\psi + \frac{8\pi^2 m}{h^2}(E - U)\psi = 0. \qquad (2.23)$$

Это уравнение позволяет решить ряд важных задач квантовой физики; оно было выведено Э. Шредингером (1926) и носит его имя. Это уравнение применимо для частиц, скорости которых малы по сравнению со скоростями света; для больших скоростей имеются более точные уравнения.

Выбирая величину ψ, мы полагаем ее непрерывной и однозначной функцией от координат x, y и z (вообще говоря, ψ будет функцией также и времени). Так как значение ψ в пределах какого-нибудь элементарного объема dV должно быть связано с вероятностью dw нахождения частицы в этом объеме, то дополнительно полагают, что эта вероятность должна быть пропорциональна квадрату модуля ψ (т. е. $|\psi(x, y, z)|^2$ есть *плотность вероятности* нахождения частиц в окрестности интересующей нас точки с координатами x, y, z). Поэтому если нам известно, что частица находится в пределах объема V, то суммарная вероятность для этого объема должна быть равна единице:

$$dw = |\psi|^2 dV; \quad \int_V |\psi|^2 dV = 1. \qquad (2.24)$$

Второе соотношение называется *нормировочным условием* для функции ψ.

Подчеркнем важное обстоятельство: согласно формуле (2.24), вероятность нахождения частицы в определенной точке ($dV = 0$) равна нулю. В квантовой физике имеет смысл определять вероятность нахождения частиц только в некоторой, не равной нулю области пространства; при $|\psi|^2 = \text{const}$ или для очень малых объемов ΔV эта вероятность будет пропорциональна объему.

РЕШЕНИЯ УРАВНЕНИЯ ШРЕДИНГЕРА

При решении многих задач требуется выяснить стационарные состояния частиц и физических систем и установившиеся течения изучаемых процессов; в этом случае изменение величины ψ (а следовательно, и «плотности вероятности» $|\psi|^2$) со временем не представляет интереса. Для решения же нестационарных задач следовало бы воспользоваться формулой плоской и монохроматической волны в более общем виде, содержащем время:

$$\Psi(x, t) = \Psi_0 e^{-i(kx - \omega t)} = \Psi_0 e^{-ikx} e^{i\omega t}$$

($k = 2\pi/\lambda$), и в соответствующем дифференциальном уравнении учесть как корпускулярно-волновые свойства частиц, так и действие на них внешних (нестационарных) силовых полей. В уравнении же Шредингера (2.23) ψ означает только амплитуду волны

$$\psi = \Psi_0 e^{-ikx}$$

и поэтому с его помощью решаются только «стационарные задачи».

Частица, например электрон, может: 1) перемещаться в свободном пространстве, в котором отсутствуют силовые поля, 2) двигаться в заданных внешних полях и, наконец, 3) входить в состав физической системы и участвовать в процессах, происходящих внутри ее. Решая уравнение Шредингера для этих задач, можно рассчитать функцию $\psi(x, y, z)$ в каждой точке пространства и тем самым определить вероятность нахождения электрона в любом интересующем нас месте.

СВОБОДНЫЙ ЭЛЕКТРОН

Рассмотрим каждую из этих задач в отдельности. В первой задаче скорость электрона остается по величине постоянной, потенциальная энергия отсутствует ($U = 0$, $E = W = mv^2/2$) и общее решение уравнения Шредингера (2.23) для амплитуды плоской волны, распространяющейся вдоль оси OX (направления движения электрона), будет иметь вид

$$\psi = Ae^{ikx} + Be^{-ikx}, \qquad (2.25)$$

где $k = \frac{2\pi}{h}\sqrt{2mW}$ — волновое число. В этом решении первый член соответствует волне, распространяющейся в положительном направлении оси OX, а второй член — волне, распространяющейся в обратном направлении. Так как в рассматриваемом примере отсутствуют электроны, движущиеся в противоположном направлении, то в решении необходимо оставить только один член:

$$\psi = Ae^{ikx}.$$

Если на пути электрона окажется дифракционная решетка, то, применяя для волны ψ принцип Гюйгенса и принцип суперпозиции, можно найти распределение $|\psi|^2$ на экране и тем самым определить вероятность попадания электрона в различные участки экрана; при прохождении через решетку большого числа электронов (одновременно или последовательно через любые промежутки времени) на экране образуется суммарная дифракционная картина — чередующиеся максимумы и минимумы.

ПОТЕНЦИАЛЬНЫЙ БАРЬЕР

Допустим теперь, что электрон движется в поле, в котором потенциальная энергия не равна нулю. В наиболее простом случае, когда $U = \text{const}$ и напряженность поля равна нулю, решение уравнения Шредингера будет отличаться от (2.25) только значением волнового числа:

$$k = \frac{2\pi}{\lambda} = \frac{2\pi}{h}\sqrt{2m(E - U)}.$$

Задача усложняется, если $U = U(x, y, z)$, т. е. напряженность поля отлична от нуля и частица движется с ускорением. Однако некоторые задачи после под-

ходящей идеализации могут быть решены с помощью уравнения Шредингера и в том случае, когда скорость частицы изменяется.

Рассмотрим простой пример — движение заряженной частицы вдоль оси OX в тормозящем электрическом поле и допустим, что торможение происходит на некотором малом участке пути, протяженность которого равна d (рис. IV.60, *а*). Будем полагать, что до этого участка поле отсутствует ($U = 0$), частица движется равномерно со скоростью v_0 и имеет энергию $E = mv_0^2/2$. Далее допустим, что в точках $x > x_0$ потенциал поля постоянен ($U = \text{const}$) и тормозящие силы опять отсутствуют. Согласно классическим представлениям, при $E > U$ частица пройдет этот «потенциальный барьер» и после него будет двигаться со скоростью

$$v_1 = \sqrt{2m(E-U)}.$$

При $E < U$ частица остановится в некоторой точке участка d и затем будет двигаться в обратном направлении; выйдя из этого участка она вновь будет иметь скорость v_0 и энергию E. Таким образом, согласно классическим представлениям, частица *либо* преодолеет барьер (при $E_1 > U$), *либо* же отразится от него ($E_2 < U$) (рис. IV, 60, *а*).

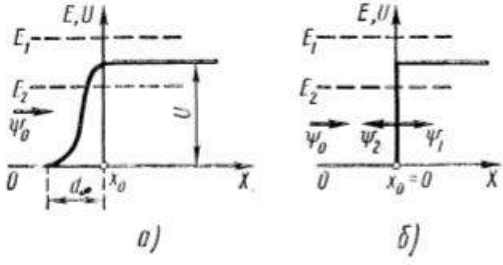


Рис. IV.60

Однако наличие у частиц волновых свойств и необходимость описания ее поведения волновым уравнением Шредингера не позволяют делать такие категорические утверждения. Для упрощения рассуждений идеализируем эту задачу допущением, что в точке $x = x_0$ (рис. IV,60, *б*) потенциал «скачком» возрастает от нуля до U (т. е. ширина участка d равна нулю). Ввиду этого уравнение Шредингера следует решать отдельно для областей $x < x_0$ и $x > x_0$, но в самой точке $x = x_0$ оба решения должны совпадать. Общие решения для этих областей запишем в виде:

$$\psi_0 = A_0 e^{ik_1 x},$$

$$\psi = A_1 e^{ik_2 x} + A_2 e^{-ik_2 x} = \psi_1 + \psi_2.$$

Для первого участка мы исключили второй член решения (2.25), так как, по условию задачи, частицы до барьера движутся только в одном направлении. Однако для второго участка этот член должен быть сохранен, пока не будет выяснено «волновое» поведение частицы в потенциальном барьере. Если, например, на барьер ежесекундно поступает N_0 частиц, то возможно, что из них N_1 пройдут через барьер, а N_2 отразятся от него. Лишь в том случае, если будет доказано, что число отраженных частиц N_2 при $E > U$ и число проходящих частиц N_1 при $E < U$ в точности равны нулю, можно исключать соответствующий член в выражении для ψ (только в этом случае классическое и квантовое решения задачи будут совпадать). Для выяснения этого вопроса выберем некоторую площадку S, перпендикулярную направлению движения частиц; тогда числа частиц N_0, N_1 и N_2, ежесекундно проходящих через эту площадку, могут быть выражены через их скорости (v_0 и v_1) и через числа частиц в единице объема (n_0, n_1 и n_2):

$$N_0 = n_0 v_0 S; \quad N_1 = n_1 v_1 S; \quad N_2 = n_2 v_0 S.$$

Так как $N_0 = N_1 + N_2$, то, разделив на $n_0 v_0$, получим

$$\frac{n_1 v_1}{n_0 v_0} + \frac{n_2}{n_0} = 1. \tag{2.26}$$

Первый член этого равенства, который обозначим через $\tau = n_1 v_1 / n_0 v_0 = N_1/N_0$, есть «коэффициент пропускания» («прозрачности») барьера; второй член $\rho = n_2/n_0 = N_2/N_0$ есть «коэффициент отражения» частиц от барьера. Для их вычисления вос-

пользуемся тем, что числа частиц в единице объема определяются квадратами амплитуд волновых функций и поэтому

$$\frac{n_1}{n_0}=\frac{|\psi_1|^2}{|\psi_0|^2}=\frac{A_1^2}{A_0^2};\quad \frac{n_2}{n_0}=\frac{|\psi_2|^2}{|\psi_0|^2}=\frac{A_2^2}{A_0^2}.$$

Далее, скорости частиц v_0 и v_1 могут быть выражены через длины дебройлевских волн в рассматриваемых областях ($\lambda_0=h/mv_0$, $\lambda_1=h/mv_1$) или через соответствующие волновые числа ($k_0=2\pi/\lambda_0$, $k_1=2\pi/\lambda_1$); тогда

$$v_1/v_0=\lambda_0/\lambda_1=k_1/k_0.$$

Воспользуемся теперь условием, что при $x=x_0$ оба решения волнового уравнения должны совпадать, т. е. $\psi_0=\psi=\psi_1+\psi_2$. Для упрощения положим, что барьер расположен в начале координат, т. е. $x_0=0$; тогда очевидно, что

$$A_0=A_1+A_2;\quad \frac{A_1}{A_0}+\frac{A_2}{A_0}=1.$$

Полученные соотношения позволяют рассчитать коэффициенты пропускания и отражения частиц от потенциального барьера. Выразим эти коэффициенты через скорости частиц:

$$\tau=\frac{4v_0v_1}{(v_0+v_1)^2}=4\,\frac{v_1/v_0}{(1+v_1/v_0)^2};\quad \rho=\left(\frac{v_1-v_0}{v_1+v_0}\right)^2=\frac{(v_1/v_0-1)^2}{(v_1/v_0+1)^2}.$$

Таким образом, отражение от барьера ($\rho\neq 0$) будет происходить и в том случае, когда $E>U$, тогда как по классическим представлениям при этом условии все частицы должны были бы преодолеть барьер и выйти во вторую область. Так как отношение скоростей равно

$$\frac{v_1}{v_0}=\frac{\sqrt{2m(E-U)}}{\sqrt{2mE}}=\sqrt{1-\frac{U}{E}},$$

то коэффициент отражения будет зависеть от относительной высоты барьера U/E; например, при $U=0{,}5E$ $\rho\approx 0{,}03$. Если $E\gg U$, то $\rho\approx 0$.

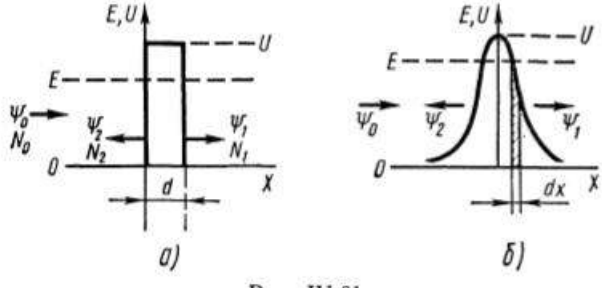


Рис. IV.61

Выше было предположено, что ширина барьера $d=0$; в этом случае отражение частиц происходит в точках с координатами $x=x_0$. Если же барьер имеет конечную ширину, то отражение будет происходить на протяжении всего участка d с постепенно возрастающим значением коэффициента ρ.

При $E<U$, когда по классическим представлениям должно иметь место полное отражение всех N_0 частиц от барьера, расчет волновых функций показывает, что непосредственно за барьером волна ψ_1, соответствующая «проходящей» части волны ψ, не равна нулю, но ее интенсивность $|\psi_1|^2$ очень быстро убывает по экспоненциальному закону по мере удаления от границы барьера. Ввиду этого если барьер имеет некоторую малую толщину d, как это показано на рис. IV.61, *а*, то из N_0 падающих на барьер частиц некоторая часть N_1 преодолеет такой барьер и только $N_2=N_0-N_1$

частиц будут отражены от него. Коэффициент прозрачности оказывается в зависимости от E, U и d:

$$\tau = e^{-\frac{4\pi}{h} d \sqrt{2m(E-U)}}.$$

Если, например, $m = 9 \cdot 10^{-31}$ кг (электрон), $U - E = 5{,}4 \cdot 10^{-19}$ Дж, $d = 0{,}527 \cdot 10^{-10}$ м, $\tau \approx 0{,}368$, т. е. около 37 % электронов смогут преодолеть этот барьер; однако если толщину барьера увеличить в 10 раз, то число прошедших электронов уменьшится почти в миллион раз.

Прохождение частиц через потенциальный барьер при $E < U$ (когда по классическим представлениям оно невозможно), называется *туннельным эффектом*. Если барьер имеет не прямоугольную форму, как это показано на рис. IV.61, *б*, но изменяется плавно, т. е. $U = U(x)$, то частицы будут проникать внутрь барьера и в каждом участке шириной dx будет происходить разделение частиц на проходящие dN_1 и отраженные dN_2. Соотношение между ними будет зависеть от $U(x)/E$ и толщины dx.

ЧАСТИЦА В ПОТЕНЦИАЛЬНОЙ ЯМЕ

Рассмотрим теперь некоторые особенности физических систем, обусловленные корпускулярно-волновыми свойствами частиц, входящих в их состав. Простейшей будет система, состоящая из одной частицы, движущейся между двумя непроницаемыми для нее плоскостями (например, непреодолимыми для нее потенциальными

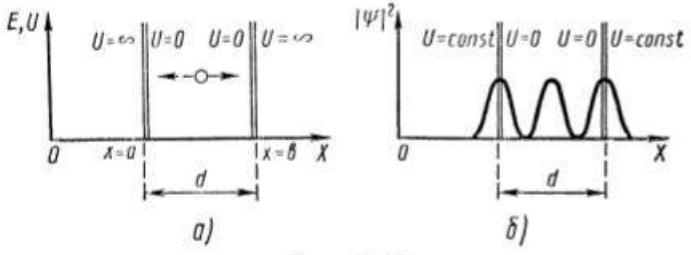


Рис. IV.62

барьерами), расстояние между которыми равно d (рис. IV.62, *а*). Для простоты допустим, что частица движется вдоль линии, перпендикулярной этим плоскостям (ось OX). Далее будем полагать, что между отражающими плоскостями поля нет (а за пределами участка d потенциал поля равен бесконечности).

По классическим представлениям в этих условиях частица будет перемещаться с постоянной скоростью от одной плоскости к другой и обратно. Важно подчеркнуть, что это движение возможно при *любой* скорости, т. е. на величину скорости и кинетической энергии частицы никаких ограничений не накладывается. Если для характеристики такого движения воспользоваться понятием вероятности, то можно утверждать, что по классическим представлениям вероятность нахождения частицы на ее траектории распределена равномерно.

Перейдем к описанию движения частицы в этих условиях при помощи уравнения Шредингера. Решение этого уравнения должно дать распределение вероятности нахождения частицы, обусловленное ее волновыми свойствами. Прежде всего заметим, что за пределами области лежащей между отражающими плоскостями, значение волновой функции ψ будет равно нулю, а внутри этой области — отлично от нуля. Ввиду непрерывности этой функции на самих плоскостях значение ψ также должно быть нулевым, следовательно,

$$\psi_{x=a} = 0; \quad \psi_{x=b} = 0.$$

Таким образом, плоская монохроматическая волна ψ, «сопровождающая частицу», должна отражаться от плоскостей, имея на концах отрезка d нулевые значения.

Ввиду этого на длине d должно укладываться целое число дебройлевских полуволн, т. е. между d и λ возможны только следующие соотношения:

$$d = \lambda/2, \quad \lambda, \quad 3\lambda/2, \quad 2\lambda, \quad \ldots, \quad n\lambda/2,$$

где $n = 1, 2, 3, \ldots$ Это означает, что движение частицы между отражающими плоскостями возможно только с определенной длиной волны $\lambda = 2d/n$, а следовательно, и с определенной скоростью

$$v = \frac{h}{m\lambda} = n\frac{h}{2dm}. \tag{2.27}$$

При других скоростях движение частицы между этими плоскостями *невозможно* (разумеется, такое категорическое утверждение связано с идеализацией задачи: реальные отражающие потенциальные барьеры всегда имеют некоторую протяженность).

Таким образом, между рассматриваемыми плоскостями (т. е. отражающими барьерами бесконечно малой протяженности) частица должна иметь одно из следующих значений кинетической энергии:

$$E = \frac{mv^2}{2} = n^2\frac{h^2}{8md^2}, \tag{2.28}$$

где $n = 1, 2, 3, \ldots$ Разрешаемые значения энергии зависят от массы частицы и расстояния между отражающими плоскостями. Если в этой области окажется частица с другим (не «разрешенным») значением энергии, то она, очевидно, будет вынуждена излучать избыток энергии и довести свою энергию до ближайшего значения, соответствующего в о з м о ж н о м у движению. Этот вывод означает, что промежуточные значения энергии частицы, отличающиеся от дозволенных, соответствуют н е у с т о й ч и в ы м (неравновесным, нестационарным) состояниям частицы. Формула (2 28) для энергии устойчивых движений частицы может быть выведена также путем решения уравнения Шредингера с использованием свойства непрерывности и однозначности волновой функции ψ и тех значений, которые она по условиям задачи имеет на границах области d.

Анализ формулы (2.28) для энергии частицы, вынужденной двигаться между двумя отражающими плоскостями, показывает, что спектр дозволенных значений этой энергии существенно зависит от расстояния d. Если d велико, то соседние значения энергии в этом спектре будут расположены очень близко друг от друга:

$$\Delta E = E_{n+1} - E_n = \frac{h^2}{8md^2}[(n+1)^2 - n^2] = \frac{h^2}{8md^2}(2n+1).$$

Например, если электрон вынужден существовать в области $d = 10^{-3}$ м, то ΔE имеет порядок 10^{-30} Дж. При больших d, соответствующих обычным размерам измерительной аппаратуры, в пределах которой движется электрон, спектр возможных значений его энергии будет почти непрерывным и можно утверждать, что электрон в этих приборах может перемещаться с любыми скоростями. Если же размеры области, в которой вынужден существовать электрон, очень малы и имеют, например, порядок размеров атома $d \approx 10^{-10}$ м, то ΔE будет порядка 10^{-16} Дж и спектр будет явно линейчатым. Заметим, между прочим, что предполагаемое отсутствие электронов внутри атомных ядер может быть обусловлено еще и тем, что для размеров ядра ($d \approx 10^{-15}$ м) величины E и ΔE достигают очень больших значений, обеспечивающих их вылет из пределов ядра.

Если потенциальный барьер на границах участка d изменяется на величину конечную, т. е. если за пределами этого участка потенциал U не равен бесконечности, а имеет некоторое конечное значение, то на этих границах возможен «туннельный эффект». Вероятность обнаружения частицы за пределами участка будет отличаться от нуля, причем, как показывают расчеты, квадрат модуля волновой функции $|\psi|^2$ весьма быстро (экспоненциально) убывает с увеличением расстояния от плоскостей (рис. IV.62. *б*). При наличии большого числа N частиц, движущихся между пластинками, некоторая часть из них будет «просачиваться» через барьеры и выйдет за пределы области d.

В более общем случае потенциальная энергия U может быть непрерывной функцией от координаты частицы, плавно возрастающей по мере удаления от середины

(или какой-нибудь другой точки) участка d. Если, например функция $U(x)$ имеет такой же вид, как и у линейного гармонического осциллятора ($U = kx^2/2$, где $k = \omega_0^2 m$ — коэффициент возвращающей силы $F = -kx$, $\omega_0 = 2\pi\nu_0 = \sqrt{k/m}$ — угловая частота колебаний), то решение уравнения Шредингера, имеющего для данной задачи вид

$$\frac{\partial^2\psi}{\partial x^2} + \frac{8\pi^2 m}{h^2}\left(E - \frac{kx^2}{2}\right)\psi = 0,$$

приводит к выводу, что «дозволенные» (стационарные) значения энергии частицы должны определяться формулой, отличной от (2.28):

$$E = \frac{hk}{2\pi m}\left(n + \frac{1}{2}\right) = h\nu_0\left(n + \frac{1}{2}\right). \qquad (2.29)$$

Эта формула показывает, что, во-первых, гармонический осциллятор должен иметь дискретный спектр значений энергии E_1, E_2, ... и, во-вторых, он не может иметь энергию, равную нулю. Наименьшее значение энергии осциллятора равно $E_0 = \frac{1}{2} h\nu_0$.

АТОМ ВОДОРОДА. КВАНТОВЫЕ ЧИСЛА

В атоме водорода электрон движется в поле протона, который можно рассматривать как заряженную частицу сферической формы. Тогда потенциальная энергия будет выражаться формулой

$$U = -\frac{e^2}{4\pi\varepsilon_0 r} = -k\frac{e^2}{r}.$$

Уравнение Шредингера для такой задачи пишется в трех измерениях:

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} + \frac{8\pi^2 m}{h^2}\left(E + k\frac{e^2}{r}\right)\psi = 0.$$

Решение этой задачи приводится в пособиях по квантовой физике; здесь же перечислим важнейшие результаты:

1) электроны могут находиться в атоме водорода, только имея определенные значения энергии:

$$E_n = -\frac{2\pi^2 m e^4}{n^2 h^2} \qquad (2.30)$$

($n = 1, 2, 3, \ldots$). Состояния с другими значениями, как и в рассмотренных выше примерах, следует трактовать как неустойчивые (нестационарные), поэтому при $E \neq E_n$ электрон должен излучать (или поглощать) энергию, с тем чтобы перейти в одно из состояний, «дозволенных» для него при данных условиях. Формула (2.30) совпадает с выражением для энергии атома водорода, полученным в § 9 на основе постулатов Бора;

2) состояние электрона в атоме водорода характеризуется набором значений трех «квантовых чисел»:

главного квантового числа n, определяющего энергию электрона в атоме согласно формуле (2.30);

орбитального квантового числа l, определяющего момент количества движения по орбите L:

$$L = \frac{h}{2\pi}\sqrt{l(l+1)}; \qquad (2.31)$$

число l принимает значения 0, 1, 2, ..., $n-1$, т. е. для данного значения главного квантового числа n существует n значений орбитального числа l;

магнитного квантового числа m, определяющего «дозволенные» (стационарные, устойчивые) ориентации орбитального момента L в пространстве, например их проекции на направление внешнего магнитного поля. Это число имеет значения 0, ± 1, ± 2, ..., $\pm(l-1)$. Таким образом, вектор момента количества движения электрона в атоме может иметь в пространстве $2l+1$ различных ориентаций.

В первоначальной теории Н. Бора предполагалось, что электроны движутся вокруг ядра по круговым или эллиптическим орбитам, причем число n определяет энергию (или скорость) электрона по орбите, число l определяет размеры орбиты (радиус окружности или полуоси эллипса), а следовательно, и момент количества движения (mvr) электрона на орбите, а число m определяет возможные ориентации этого момента в пространстве. Однако в отличие от таких модельных интерпретаций квантовых чисел решение уравнения Шредингера для атома водорода допускают также и состояния с $l = 0$, т. е. с нулевым значением момента количества движения $L = 0$. По боровским представлениям это означает отсутствие орбиты. Для таких состояний распределение вероятности нахождения электрона в атоме оказывается сферически симметричным, т. е. $|\psi|^2$ зависит только от расстояния r до центра атома и поэтому одинаково вдоль поверхности сферы. На этой сферической поверхности, радиус которой равен радиусу боровской орбиты, соответствующей значению $l = 0$, $|\psi|^2$ достигает максимума, а при удалении от нее быстро убывает, поэтому электрон оказывается «запертым» в некотором кольцевом слое. Заметим, между прочим, что вероятность нахождения электрона на самой поверхности этой сферы равна нулю, так как вероятность нахождения пропорциональна объему пространства; из $d\omega = |\psi|^2 dV = |\psi|^2 4\pi r^2 dr$ следует, что при $dr = 0$ $d\omega = 0$.

Для состояния с $l \neq 0$ распределение $|\psi|^2$ уже не является сферически симметричным; для этих состояний боровские орбиты оказываются теми линиями, вдоль которых $|\psi|^2$ имеет максимумы.

Каждое из значений энергии E_n, соответствующих определенному устойчивому состоянию электрона, атома или любой другой физической системы, называется *уровнем энергии*. Совокупность уровней энергии, определяемая различными значениями главного квантового числа $n = 1, 2, 3, \ldots$, называется *энергетическим спектром* частицы или системы. Ниже дана схема различных состояний электрона в атоме водорода, соответствующая всем возможным значениям квантовых чисел n, l и m. Дополнительно указаны состояния, определяемые ориентацией собственного момента количества движения электрона («спи-

на») (см. ч. III, § 20) относительно орбитального момента L; это четвертое (*спиновое*) квантовое число имеет только два значения: $s = \pm 1/2$.

Первый основной уровень $n = 1$ состоит только из двух подуровней (рис. IV. 63).

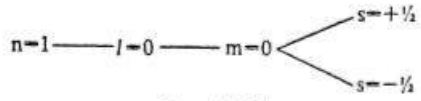


Рис. IV.63

Второй основной уровень $n = 2$ состоит из восьми подуровней (рис. IV. 64).

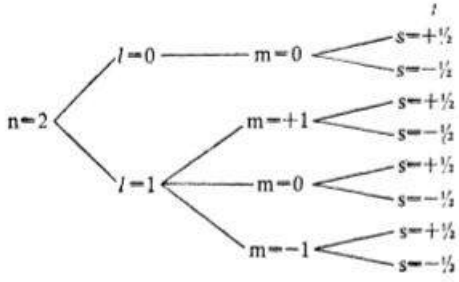


Рис. IV.64

Заметим, что этот уровень $n = 2$ распадается на две части; первая содержит два подуровня такого же типа ($l = 0$, $m = 0$), какие имеются на первом основном уровне $n = 1$; вторая содержит шесть новых уровней ($l = 1$; $m = +1, 0, -1$). Третий, основной, уровень $n = 3$ повторяет подуровни $l = 0$ и $l = 1$ и дополнительно содержит десять подуровней, объединенных квантовым числом $l = 2$ (рис. IV. 65).

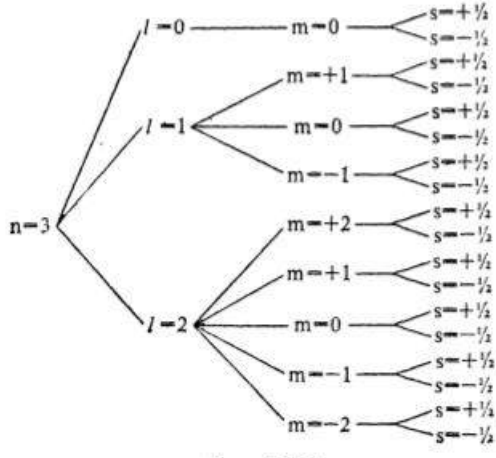


Рис. IV.65

Эти схемы могут быть продолжены также и для последующих значений n. Заметим, что одному значению главного квантового числа n, т. е. определенному значению энергии атома водорода E_n, соответствует

$$z = \sum_{l=0}^{n-1} 2(2l+1) = 2n^2 \qquad (2.32)$$

различных устойчивых состояний электрона в атоме. Если атом изолирован от внешних воздействий (разреженный газ), то его энергия изменяется только при переходе электрона из состояний, имеющих одно значение n, к состояниям с другим значением n.

Таким образом, уравнение Шредингера, примененное к изолированному атому водорода, позволяет определить спектр уровней энергии и спектр состояний без каких-либо дополнительных предположений. Постулаты Бора, использованные в § 9 для вывода сериальных формул водорода, уже содержатся в этом уравнении.

СООТНОШЕНИЕ НЕОПРЕДЕЛЕННОСТЕЙ

Существование корпускулярных и волновых свойств у элементарных частиц и связь между ними имеет в квантовой физике еще одно выражение. В. Гейзенберг (1927) показал, что эти свойства ограничивают точность о д н о в р е м е н н ы х измерений координат и импульсов. Если, например, измерительная установка позволяет точно зафиксировать координату x движущейся частицы в какой-нибудь определенный момент времени, то измерение точного значения импульса этой частицы в направлении той же оси для того же самого момента времени оказывается из-за волновых свойств частиц н е в о з м о ж н ы м. Если же измерительная установка позволяет определить точное значение импульса частицы, то ее волновые свойства не разрешают найти точное местонахождение частицы в тот же самый момент времени. Имеется несколько записей этих «соотношений неопределенности» В. Гейзенберга, например

$$\Delta p_x \Delta x \geqslant \hbar/2; \quad \Delta p_y \Delta y \geqslant \hbar/2; \quad \Delta p_z \Delta z \geqslant \hbar/2, \qquad (2.33)$$

где $\hbar = h/2\pi$; Δx, Δy, Δz — неточности («неопределенности») в измерении координат частицы, а Δp_x, Δp_y и Δp_z — неточности в измерении составляющих импульса вдоль соответствующих координатных осей. Эти соотношения показывают, что повышение точности измерения одной из этих «сопряженных» величин сопровождается обязательным понижением точности в определении другой величины. Заметим, что например, $\Delta x = 0$ означает только $\Delta p_x \to \infty$; точные же измерения других составляющих импульса (p_y и p_z) при этом возможны.

Иногда из соотношений неопределенности (2.33) делают вывод, что одновременные точные значения координат и импульса не существуют вообще. Полагают, что в данный момент времени существует некоторое непрерывное множество значений этих величин, имеющих различные вероятности реализации с очень острым максимумом для одного значения. На рис. IV. 66 показаны графики так называемых

«волновых пакетов» для координаты x и импульса p_x в некоторый момент времени; по оси ординат отложены «плотности вероятности» для каждой из этих величин. По этим функциям можно рассчитать средние квадратичные отклонения $<\Delta x>^2$ и $<\Delta p_x>^2$ (на рис. IV. 66 эти отклонения обозначены Δx и Δp_x) от средних значений x_0 и p_0. Тогда соотношения Гейзенберга записывают в виде связи между этими средними квадратичными отклонениями (флуктуациями), например вдоль оси OX

$$\langle \Delta p_x^2 \rangle \langle \Delta x^2 \rangle \geqslant \hbar^2/4. \tag{2.34}$$

Корпускулярно-волновые свойства частиц накладывают ограничения также и на одновременное измерение двух других «сопряженных»

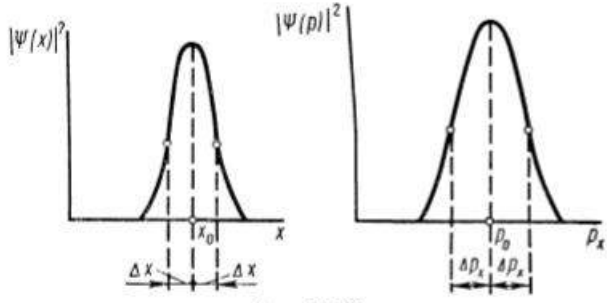


Рис. IV.66

физических величин: энергии и времени. Соответствующее соотношение неопределенностей имеет вид

$$\Delta E\, \Delta t \geqslant \hbar/2. \tag{2.35}$$

Величину Δt можно интерпретировать как время, в течение которого частица имеет некоторое (наивероятное) значение энергии E с разбросом, равным ΔE. Более строгий вывод соотношения между неточностями в определении энергии (частицы или системы) и временем пребывания в этом состоянии также приводит к формуле (2.35).

Приведем несколько иллюстраций к соотношениям неопределенностей:

1) допустим, что электрон движется в трубке катодного осциллографа или телевизора со скоростью $v = 10^6$ м/с и эта скорость может быть определена (вычислена или измерена) с точностью до 0,01%, т. е. $\Delta v = 100$ м/с. Тогда по соотношению неопределенностей координаты электрона могут быть определены с точностью до

$$\Delta x = \frac{h}{2\pi \cdot \Delta (mv)} = \frac{6{,}63 \cdot 10^{-34}}{6{,}28 \cdot 9{,}1 \cdot 10^{-31} \cdot 10^2} \approx 10^{-6} \text{ м}.$$

Если сравнивать эту величину с размерами самого прибора, то следует полагать, что точность определения координаты электрона достаточно велика и понятие траектории движения электронов в этих приборах имеет смысл. Однако по сравнению с предполагаемыми размерами самого электрона, радиус которого имеет порядок 10^{-12} м, эта неточность чрезвычайно велика;

2) если рассматривается движение электрона внутри атома, то наибольшая неточность в определении его координаты может иметь порядок размеров атома, т. е. $\Delta x \approx 10^{-10}$ м. Тогда по соотношению неопределенностей скорость электрона может

быть определена с точностью до

$$\Delta v = \frac{h}{2\pi m\,\Delta x} = \frac{6{,}63\cdot 10^{-34}}{6{,}28\cdot 9{,}1\cdot 10^{-31}\cdot 10^{-10}} \approx 10^{6}\ \text{м/с}.$$

Однако, согласно расчетам, приведенным в § 9, величина скорости электрона в атоме имеет такой же порядок, следовательно, одновременное задание (или измерение) точного значения скорости электрона и точной координаты его внутри атома не имеет смысла;

3) рассмотрим движение альфа-частицы вблизи ядра (например, в опытах Резерфорда по рассеянию альфа-частиц атомами). Допустим, что скорость этих частиц определена с точностью до $\Delta v \approx 10^{4}$ м/с. Неопределенность в их координате будет иметь порядок (масса альфа-частицы равна $\approx 6{,}7\cdot 10^{-27}$ кг)

$$\Delta x = \frac{h}{2\pi m\,\Delta v} = \frac{6{,}63\cdot 10^{-34}}{6{,}68\cdot 6{,}7\cdot 10^{-27}\cdot 10^{4}} \approx 1{,}47\cdot 10^{-12}\ \text{м}.$$

Сравнивая этот результат с размерами атома ($\approx 10^{-10}$ м) и атомного ядра ($\approx 10^{-15}$ м), мы видим, что внутри атома взаимодействие между альфа-частицей и ядром можно описывать по закону Кулона для точечных зарядов и что понятие траектории альфа-частицы внутри атома имеет смысл;

4) допустим, что атом находится в возбужденном состоянии, и по формулам, приведенным в § 9, имеет энергию E_n. Согласно соотношению неопределенности (2.35), энергия частицы или физической системы и время пребывания в состоянии с этой энергией не могут быть заданы (известны или измерены) точно. Допустим, что неточность в значении энергии равна ΔE; тогда фотон, испускаемый атомом при переходе в нормальное состояние, будет иметь не одну определенную частоту ν, а некоторый интервал частот $\Delta\nu = \Delta E/h$. Величину $\Delta\nu$ можно приравнять естественной ширине спектральной линии, которая может быть измерена. Если, например, $\Delta\nu = 10^{10}$ Гц, то время пребывания атома в возбужденном состоянии может быть оценено по соотношению неопределенностей:

$$\Delta t = \frac{h}{2\pi\,\Delta E} = \frac{1}{2\pi\,\Delta\nu} \approx 0{,}6\cdot 10^{-10}\ \text{с}.$$

Заметим, что в нормальном состоянии атом может существовать бесконечно долго, поэтому энергия этого состояния может быть известна точно.

Таким образом, если размеры области, в пределах которой движется данная частица, велики (по сравнению с длиной дебройлевской волны), то можно пользоваться классическим («траекторным») описанием движения частицы, т. е. применять к ней понятия и законы классической физики. Если же область существования частицы или амплитуда ее колебаний малы по сравнению с длиной дебройлевской волны, то классическое описание поведения этой частицы оказывается весьма грубым и даже может оказаться вообще не имеющим смысла. В этих условиях главными факторами, определяющими поведение частицы, являются уже их волновые свойства.

§ 12. ЧАСТИЦЫ И ФИЗИЧЕСКИЕ СИСТЕМЫ. СТАТИСТИЧЕСКИЕ РАСПРЕДЕЛЕНИЯ ЧАСТИЦ ПО СОСТОЯНИЯМ. ПОНЯТИЕ ТЕМПЕРАТУРЫ

ЧАСТИЦА

Частицей обычно называется физический объект, свойства которого сосредоточены в пределах относительно малых объемов (по сравнению с объемом всей системы). Линейные размеры этого объема должны быть настолько малы, чтобы весьма важные в физике понятия «координаты частицы» и «расстояние между частицами» имели определенный смысл.

Однако указание на относительную малость размеров объекта не является существенной частью определения понятия «частица»; необходим перечень ф и з и ч е -

с к и х требований, которым должен удовлетворять объект для того, чтобы его можно было назвать *частицей*. Ни в классической, ни в квантовой физике полный перечень таких требований не приводится, поэтому *частицей* называют и электрон или протон, и атом или атомное ядро. В период господства классической механики (конец XIX в.), основными объектами которой являлись «материальная точка» и «абсолютно твердое тело», простейшими физическими телами могли считаться частицы, не имеющие никакой внутренней структуры, абсолютно твердые, абсолютно упругие и т. п. Современная физика придает понятию «частица» более свободное определение, вследствие чего *элементарными* называются частицы с весьма различными свойствами (фотоны, электроны, нейтроны и т. д.).

Перечень требований, которым должны удовлетворять частицы, очевидно, не может быть исчерпывающим и будет отражать определенный уровень физических знаний; наиболее же важным является следующее условие:

физический объект можно назвать частицей (входящей в состав некоторой системы), если во всех физических явлениях, в которых он участвует, изменения в его внутреннем состоянии не обнаруживаются, или если влиянием его внутреннего состояния на общий ход физических процессов в системе можно пренебречь.

При этом условии понятие «частица» приобретает относительный смысл. Если в некоторой области физических явлений, изучаемой при помощи определенной аппаратуры, данное тело не обнаруживает своей внутренней структуры и обладает постоянными физическими свойствами, то в этой области и по отношению к этой аппаратуре рассматриваемое тело можно назвать *частицей*. В других физических условиях и по отношению к другой (более чувствительной) аппаратуре это тело может вести себя как сложное, т. е. может обнаружиться состав и внутренняя структура частицы, влияние внутренних изменений на ее отношение к действию внешних сил и т. п. Как правило, при относительно малых энергиях и слабых воздействиях поведение многих тел определяется в основном только внешними воздействиями, т. е. тела не обнаруживают (точнее, почти не обнаруживают) своего внутреннего состояния; при этих условиях все свойства частиц можно полагать неизменными. В области же высоких энергий и сильных внешних воздействий тела показывают свойства, свидетельствующие о наличии у них состава, структуры и спектра внутренних состояний. Так, например, ядро какого-нибудь тяжелого элемента при столкновении с протоном, имеющим малую энергию, ведет себя как одна частица с определенным зарядом и массой; если же протон имеет достаточно большую энергию, то при взаимодействии с ним ядро обнаруживает свой сложный состав, структуру и внутреннее состояние. Мы называем молекулу *частицей газа*, несмотря на наличие у нее состава и структуры, а также нормального и возбужденных состояний. При этом мы исходим из предположения, что изменения во внутреннем состоянии одной молекулы не влияют на ход процессов, происходящих в газе. Однако возможно, что это влияние исчезающе мало в каждом отдельном акте взаимодействия данной молекулы с остальными, но если число таких актов чрезвычайно велико, то после суммирования оно может оказаться заметным.

ФИЗИЧЕСКАЯ СИСТЕМА

Понятие «физическая система», как и любое другое понятие, также должно быть по возможности точно определено. Это определение может быть лаконичным или подробным, но оно должно охватывать существенные особенности физических объектов, называемых *системами*, отличающие их от объектов, названных *частицами*. Возможно, что нахождение достаточно четких и однозначных определений для физической системы будет трудной задачей, однако независимо от их качества, они необходимы.

Физической системой называют локализованную в некотором ограниченном объеме устойчивую совокупность взаимодействующих частиц, которая в физических процессах ведет себя как один объект, обладающий определенными свойствами.

Свойства физической системы (так же как и одной частицы) могут быть выражены числовыми значениями определенных величин (массы, электрического заряда, спина и т. д.), а также характеристическими функциями, т. е. формулами для расчета различных функций состояния: энергии, энтропии, спектра излучения и т. п.

Основные отличительные особенности физических систем, которые должны быть подчеркнуты в определениях этого понятия, заключаются в следующем:

1) устойчивость по отношению к различным внешним воздействиям;

2) наличие качественно новых свойств по сравнению с составными частями;

3) существование спектра возможных состояний, среди которых одно или несколько состояний выделяются как наиболее устойчивые (стационарные, равновесные);

4) существование тенденции к переходу в равновесное состояние;

5) наличие хаотического (теплового) движения частиц внутри системы, оцениваемое температурой.

Свойства физических систем формируются из свойств составных частиц и определяются характером их взаимодействия между собой. Прежде всего следует заметить одно интересное обстоятельство: система обладает устойчивостью и обнаруживает качественно новые свойства лишь в том случае, когда она укомплектована из частиц, обладающих *различными* (полярными) свойствами. Наиболее простым и наглядным примером является атом водорода; каждый протон или электрон в свободном (изолированном) состоянии может иметь любые значения энергии, но составленный из них атом может существовать только в состояниях с определенными значениями энергии; одни только протоны или одни только электроны устойчивой (прочной) системы не образуют. Ниже мы перечислим все важнейшие свойства физических систем атомного масштаба:

1. Масса (инертная) физической системы M всегда меньше суммы масс m_i составных частей этой системы (измеренных в их свободном состоянии). Разность

$$\Delta m = \Sigma m_i - M$$

называемая *дефектом массы* этой системы, будучи умноженной на квадрат скорости света, показывает ту энергию, которая выделилась при образовании данной системы из свободных частиц. Такой же величины энергию $\Delta W = \Delta m c^2$ необходимо затратить, чтобы разрушить систему, т. е. удалить ее составные части на столь большие расстояния, при которых их взаимодействием можно было бы пренебречь. ΔW называется *энергией связи* данной системы и характеризует его устойчивость (прочность по отношению к различным внешним воздействиям). Для некоторых систем трудно измерить Δm (ввиду относительной малости), но легко измерить ΔW.

Физические системы образуются и существуют благодаря силам взаимодействия между ее составными частями. На относительно больших расстояниях всегда должны действовать силы притяжения (без них система не могла бы образоваться), но на малых расстояниях либо эти силы должны уравновешиваться, вызывая, например, центростремительные ускорения (как в солнечной системе или боровской модели атома), либо же должны появиться силы отталкивания. Можно указать на физические системы, которые существуют весьма малое время только вследствие того, что между ее составными частицами при их сближении не возникают силы отталкивания (например, позитроний). Силы отталкивания, если они существуют, определяют форму образовавшихся систем (например, молекул из атомов, атомных ядер из нуклонов). Энергия связи системы определяется интенсивностью сил притяжения (см. рис. II.2 и соответствующий текст).

Характерными примерами возникновения физических систем с выделением энергии связи являются: конденсация паров в жидкость, кристаллизация, соединение атомов в молекулы, образование атомов водорода из протонов и электронов, образование атомных ядер из протонов и нейтронов.

Обычно выделяющуюся при образовании (или затрачиваемую при разрушении) энергию ΔW делят на число частиц и находят среднюю энергию, необходимую для удаления одной частицы из состава системы. Однако, ввиду того что энергия связи пропорциональна массе всей системы, более подходящей характеристикой системы следует считать *удельную энергию связи*, приходящуюся на единицу массы. Приведем

таблицу измеренных значений удельной энергии, Дж/кг, для перечисленных выше процессов образования систем:

Конденсация водяного пара в жидкость (при температуре в 100 °С и нормальном давлении)	$2{,}26 \cdot 10^6$
Образование кристалла льда из воды (при 0 °С)	$3{,}35 \cdot 10^5$
Соединение атомов водорода и кислорода в молекулу воды	$1{,}35 \cdot 10^7$
Соединение протона и электрона в атом водорода	$1{,}28 \cdot 10^9$
Соединение протона и нейтрона в ядро дейтерия	$1{,}06 \cdot 10^{14}$

2. В физических системах обнаруживаются волновые свойства составляющих ее частиц. Выше упоминалось о появлении квантовых свойств у атома водорода; то же самое имеет место и у атомных ядер, сложных атомов, молекул и т. п. В этих системах волновые свойства частиц имеют важное значение в формировании поведения всей системы в целом. Вследствие этого процессы, в которых участвуют так называемые «микросистемы» (ядра, атомы, молекулы), описываются понятиями и законами квантовой физики, тогда как для «макросистем» могут быть использованы с удовлетворительной точностью понятия и законы классической физики.

3. Физическая система характеризуется набором (*спектром*) возможных для нее состояний. У некоторых систем спектр возможных состояний совпадает с «энергетическим спектром», т. е. одному уровню энергии соответствует одно-единственное состояние. У других систем каждому уровню в энергетическом спектре может соответствовать несколько (иногда очень много) различных состояний. Например, у атома водорода (см. § 10—11) одному значению главного квантового числа n, определяющего энергию атома E_n, соответствует $2n^2$ различных состояний, отличающихся значениями остальных квантовых чисел. Уровень энергии, который охватывает некоторое множество различных состояний, называется *вырожденным*; *степенью вырождения* называется число различных состояний, имеющих данное значение энергии.

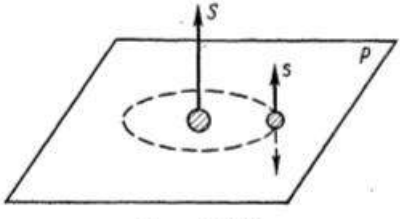


Рис. IV.67

При помощи внешнего воздействия можно устранить вырождение, т. е. изменить величину энергии для тех различных состояний, которые ранее (при отсутствии внешнего воздействия) имели одинаковые энергии. Рассмотрим пример — идеализированный атом водорода по первичной теории Бора. Внесем небольшое уточнение; учтем, что протон имеет (см. ч. III, § 20) собственный механический (и магнитный) момент S, и примем направление вектора S за ось всего атома (рис. IV. 67). Проведем плоскость P, перпендикулярную этой оси. Допустим, что атом находится в возбужденном состоянии, соответствующем $n = 3$ и энергии $E_3 = 1{,}5$ эВ. При этой энергии атом может находиться в $2n^2 = 2 \cdot 3^2 = 18$ различных состояниях. Эти состояния отличаются: 1) размерами орбит (определяемыми числами l); 2) разрешаемыми ориентациями плоскостей каждой орбиты относительно плоскости P (числа m) и, наконец, 3) двумя возможными ориентациями вектора собственного механического (и магнитного) момента электрона — спина — относительно плоскости орбиты (число s).

Если ввести такой атом в электрическое поле, то на протон и электрон будут действовать различно направленные силы, вследствие чего атом будет несколько деформирован. Эта деформация будет различной для каждого из 18 состояний, охватываемых $n = 3$, так как она будет зависеть от размеров орбит, ориентаций их плоскостей и направления спина электрона относительно внешнего поля. Ввиду этого вместо одного уровня E_3 образуются 18 близко расположенных подуровней, которые можно обозначить через E_{31}; E_{32}, ... Различие между энергиями этих подуровней будет зависеть от напряженности внешнего поля, но оно оказывается значительно меньше, чем различие между соседними уровнями энергии атома. Существование такого расщепления, очевидно, можно обнаружить, если исследовать излучение атома, помещенного в достаточно сильное поле. При переходе электрона из одного подуровня, принадле-

жащего уровню n, на какой-нибудь подуровень, входящий в состав уровня k, будет излучаться фотон с частотой

$$\nu_{\alpha\beta} = (E_{n\alpha} - E_{k\beta})/h.$$

Следовательно, вместо одной спектральной линии с частотой $\nu = (E_n - E_k)/h$, которая излучалась бы изолированным атомом, в этом случае будет наблюдаться некоторое множество близких друг к другу спектральных линий.

Расщепление спектральных линий (т. е. энергетических уровней) атомов, помещенных в магнитном поле, было обнаружено Зееманом (1896), а в электрическом поле — Штарком (1913).

Заметим, что существование «вырождения» энергетических уровней у атома водорода следует только из первоначальной модели Бора. В этой модели предполагалось, что энергия атома состоит т о л ь к о из кинетической энергии электрона и потенциальной энергии взаимодействия протона с электроном, рассматриваемых как точечные заряды (или равномерно заряженные сферы). Существование других видов взаимодействия между протоном и электроном, например взаимодействие, обусловленное наличием у них магнитных моментов, не предполагалось. Однако если учесть, что в атоме имеется определенное направление, выделенное моментом протона, то в общем случае следует полагать, что энергия взаимодействия должна определяться не только размерами орбиты, но и ориентацией плоскостей орбит и спина электрона относительно этого направления. Эта дополнительная энергия взаимодействия ничтожно мала по сравнению с кулоновской потенциальной энергией, однако если эту энергию учесть, то вырождение будет отсутствовать. В частности, для каждого из состояний с $l = 0$ (см. § 11), когда, по квантовой теории, представление об орбите теряет смысл, уровень энергии оказывается расщепленным на два подуровня: один из них соответствует энергии атома при одинаковых направлениях собственных магнитных моментов протона и электрона, другой — при противоположных направлениях.

У физических систем следует различать:

1) спектр состояний самой *системы в целом*. Обозначим уровни энергии, соответствующие устойчивым (стационарным) состояниям системы, через $E_1, E_2, \ldots$ Например, у идеального газа этот спектр непрерывный, у атома водорода имеет дискретную структуру, приведенную в § 9;

2) спектр состояний, в которых могут находиться *отдельные составные части* этой системы. Например, у атома кислорода ядро имеет свой спектр устойчивых состояний, электроны же, окружающие ядро, имеют другой, существенно отличный от первого спектр значений кинетической энергии. Заметим также, что и потенциальная энергия взаимодействия электрона с ядром также имеет дискретный набор «дозволенных» значений. Уровни энергии какой-нибудь составной частицы системы будем обозначать через $\varepsilon_1, \varepsilon_2, \ldots$ (большим значениям индекса соответствуют большие энергии).

В некоторых физических системах имеется ограниченное количество «мест», которые могут быть заняты частицами с определенной энергией ε_i; допустим, что в системе имеется:

1) z_1 мест для частиц с энергией ε_1;
2) z_2 » » » » ε_2;
3) z_3 » » » » ε_3

и т. д. Число z_i равно произведению числа α_i различных состояний частицы, соответствующих данному уровню ε_i, на число мест β_i в структуре системы (*в его объеме*), где находятся эти уровни: $z_i = \alpha_i\beta_i$. Оба числа α и β могут быть различными для различных ε. Следовательно, если система содержит N одинаковых частиц, то для их распределения внутри системы имеется $Z = \Sigma z_i$ мест. Это число может быть конечным, если конечна каждая из составляющих z_i и если спектр частиц внутри системы ограничен какими-то (минимальным и максимальным) значениями энергии уровней ε. Если же верхнего предела для энергии уровней нет, то общее число мест в системе для пребывания в них рассматриваемых частиц может быть неограниченным. Возникает вопрос: каким образом распределяются частицы по этим местам, или, как принято выражаться, в каком порядке «заполняются» частицами возможные для них энергетические уровни. Ответ на этот вопрос связан с весьма важной тенденцией, которая имеется у всех физических систем.

4. В физических системах существует тенденция к преимущественному заполнению уровней с малыми значениями энергии. Под действием этой «упорядочивающей тенденции» частицы в первую очередь занимают места с малыми значениями энергии («низшие уровни») и лишь после того, как все места в них будут заняты, начинается заполнение состояний с большими значениями энергии («высшие», или «верхние», уровни). Если число частиц данного сорта в системе N меньше полного числа возможных (устойчивых для них) состояний, то незаполненными («вакантными») должны оставаться состояния, принадлежащие высшим уровням.

Коэффициентом заполнения уровня ε_1 в системе будем называть отношение числа ΔN_i частиц, занимающих такие уровни, к соответствующему числу мест в системе с энергией ε_i:

$$\gamma_i = \Delta N_i / z_i.$$

Для заполненных уровней $\Delta N_i = z_i$; $\gamma_i = 1$.

РОЛЬ ТЕПЛОВОГО ДВИЖЕНИЯ ЧАСТИЦ

Однако указанная тенденция к преимущественному заполнению низших уровней может обеспечить строгий порядок в распределении частиц по состояниям (местам) только при отсутствии в системе каких-либо других факторов, которые могли бы влиять на это распределение. Таким фактором может быть хаотическое тепловое движение (если в нем участвуют рассматриваемые частицы системы). При этом движении возможны передача энергии от одних частиц к другим и, следовательно, их переход из одного состояния в другое. Поэтому упорядоченное распределение частиц по состояниям, при котором низшие уровни заполнены и вакансии имеются только у высших уровней, может быть осуществлено т о л ь к о при абсолютном нуле температуры. При отличных от нуля температурах частицы имеют возможность получить в результате тепловых взаимодействий энергию порядка $\Delta\varepsilon = kT$ (k— постоянная Больцмана). Однако, для того чтобы какая-нибудь частица, имеющая энергию ε_i, могла покинуть свое место и перейти в другое, необходимо, чтобы полученная ею энергия $\Delta\varepsilon$ была равна разности между ε_i и энергией ближайшего вакантного места. Для частиц, занимающих низшие уровни, такие переходы потребуют больших $\Delta\varepsilon$ (так как все соседние уровни с близкими значениями энергии заняты), поэтому п р и н и з к и х т е м п е р а т у р а х эти частицы будут оставаться на своих местах. Переходы будут совершаться только частицами, находящимися на верхних уровнях. По мере повышения температуры в эти переходы будут постепенно вовлекаться и частицы, находящиеся на низших уровнях (тем более что в ближайших верхних уровнях появляются вакансии). При очень высоких температурах возможно, что переходы, обусловленные тепловыми взаимодействиями, затронут все существующие в системе уровни энергии для частиц.

ФУНКЦИИ РАСПРЕДЕЛЕНИЯ ПО СОСТОЯНИЯМ

Таким образом, дополнительно к задаче нахождения чисел возможных состояний для каждого значения энергии возникает новая проблема: определить коэффициенты заполнения каждого из уровней для заданной температуры системы. Числа частиц ΔN, находящихся на уровнях с энергией ε, представляются в зависимости от общего числа частиц N и некоторой функции от ε и температуры T:

$$\Delta N = NBf(\varepsilon, T); \quad \Delta N/N = Bf(\varepsilon, T), \tag{2.36}$$

где B — величина, не зависящая от ε и T; для различных систем она может иметь различные значения.

В статистической физике важное значение имеют три вида *функции распределения*:

$$f_1(\varepsilon, T) = e^{(\varepsilon_0 - \varepsilon)/kT}; \quad f_2(\varepsilon, T) = \frac{1}{e^{(\varepsilon - \varepsilon_0)/kT} - 1}; \quad f_3(\varepsilon, T) = \frac{1}{e^{(\varepsilon - \varepsilon_0)/kT} + 1}. \tag{2.37}$$

Первая из них называется *распределением* (или *статистикой*) *Максвелла—Больцмана* и применяется к системам, у которых постоянное число частиц распределяется

по заданному энергетическому спектру. Вторая формула есть *распределение Бозе—Эйнштейна* и относится к системам, у которых: 1) число частиц не остается постоянным и зависит от температуры (т. е. в системе с повышением температуры могут «рождаться», а с понижением — «исчезать» частицы; такой системой является, например, фотонный газ, находящийся в равновесии с тепловыми излучателями; см. § 15) и 2) на каждом энергетическом уровне может находиться любое число частиц. Третья формула — *распределение Ферми—Дирака* — применяется к системам, у которых на каждом уровне может находиться только одна частица. Такие системы обладают дополнительным свойством: при каждой температуре лишь некоторая доля общего числа частиц принимает участие в тепловом движении; с повышением температуры в тепловое движение постепенно вовлекаются и остальные частицы (такой системой является, например, электронный газ в металлах).

ТЕМПЕРАТУРА

Понятие температуры в термодинамике и статистической физике является фундаментальным. Однако его строгое определение оказывается сложным и несколько необычным. Ниже мы рассмотрим одно из возможных определений температуры, связанное со статистическим распределением частиц системы по состояниям. Для этого сначала вернемся к термодинамическим системам, состоящим из N одинаковых частиц, из которых N_1 имеют энергию ε_1, N_2 — энергию ε_2 и т. д., поэтому общая энергия системы равна

$$U = N_1\varepsilon_1 + N_2\varepsilon_2 + \ldots = \sum N_i\varepsilon_i = N\varepsilon_{\text{ср}}, \tag{2.38}$$

где $\varepsilon_{\text{ср}}$ — средняя энергия, приходящаяся на одну частицу. Очевидно, что состояние системы не может быть охарактеризовано одним только значением средней энергии; если же дополнительно указать функцию распределения частиц по энергиям:

$$f(\varepsilon_i) = N_i/N; \quad N_i = Nf(\varepsilon_i), \tag{2.39}$$

то описание состояния системы будет более детальным.

У идеального газа температура и давление связываются со средней энергией молекул (см. ч. II, § 9) $\varepsilon_{\text{ср}} = ikT/2$:

$$T = 2\varepsilon_{\text{ср}}/ik; \quad p = nkT,$$

где i — число степеней свободы одной молекулы.

Заметим, что давление и температура газа *не зависят* от функции распределения молекул по энергиям (скоростям) и не определяют ее вид. Действительно, миниатюрные термометры и манометры, установленные в различных местах объема газа, будут отмечать только среднюю энергию поступательного движения молекул и не могут фиксировать числа N_1, N_2, ... молекул, имеющих различные энергии. При одинаковой средней энергии молекул термометры не отличают максвелловского распределения молекул по скоростям от другого, при котором, например, все скорости молекул одинаковые. Поэтому описание состояния термодинамических систем при помощи давления и температуры не является полным и соответствует описанию, при котором задается только средняя энергия молекул; при этом описании функция распределения молекул по энергиям может иметь различный вид.

Очевидно, что более удобным было бы определение температуры в зависимости от *вида функции распределения* частиц по энергетическим уровням. В этом случае задание температуры системы означало бы одновременное задание функции распределения и поэтому описание системы было бы более полным. Выше было указано на существование в системе двух противоположных тенденций:

1) упорядочивающей тенденции, обеспечивающей преимущественное заполнение низших уровней энергии, и

2) теплового (хаотического) движения, направленного к равномерному распределению энергии системы между частицами (точнее, между степенями свободы частиц).

При температуре абсолютного нуля будет существовать только первая тенденция и частицы системы окажутся распределенными по состояниям в определенном порядке: не может быть вакансий на низших уровнях, если есть частицы, занимающие

более высокий уровень. Это означает, что энергия системы при абсолютном нуле вовсе не должна быть нулевой; она будет равна

$$U_0 = N_1\varepsilon_1 + N_2\varepsilon_2 + \ldots + N_n\varepsilon_n,$$

т. е. зависит от того распределения частиц по энергетическим уровням, которое устанавливается при $T = 0$. Следовательно, и средняя энергия, приходящаяся на одну частицу системы, будет при абсолютном нуле температуры отличной от нуля. Ввиду этого определение понятия температуры через среднюю энергию одной частицы теряет смысл.

Если у изучаемой системы и у термометра (который также является физической системой) частицы размещены по состояниям т о л ь к о в соответствии с «упорядочивающей тенденцией», т. е. если эти оба объекта находятся при температуре абсолютного нуля, то между ними никакого взаимодействия не будет. Если же термометр внести в систему с температурой T, то между частицами системы и термометра будет происходить взаимодействие (обмен энергиями); распределение частиц по уровням в термометре будет нарушаться до тех пор, пока оно не станет таким же, какое существует в системе. Другими словами, выравнивание температуры представляется как переход к одинаковому характеру распределения частиц по состояниям, т. е. к одинаковому виду функции распределения $f(\varepsilon)$.

При повышении температуры изменяется не только распределение частиц по уровням $f(\varepsilon)$, но и внутренняя энергия самой системы U. Разность энергии $U - U_0 = \Delta U$ будет пропорциональна температуре и, разумеется, также может быть использована для ее определения. Однако между ΔU и температурой T не существует простой зависимости. Такая зависимость (и притом линейная) имеется только у идеального газа. Таким образом,

---

для характеристики состояния любой термодинамической системы ее температуру следует связывать только с видом функции распределения частиц системы по уровням энергии.

В частности, например, если в системе устанавливается распределение Максвелла—Больцмана $N_i/N = Ae^{-\varepsilon/kT}$ и если установлено, что на уровне с энергией ε_i находится N_i частиц из общего числа N, то температура определится из соотношения

$$\frac{1}{T} = \frac{k}{\varepsilon_i} \ln\left(A\,\frac{N}{N_i}\right).$$

Для систем, у которых частицы распределены согласно статистикам Бозе—Эйнштейна и Ферми—Дирака, связь между температурой и видом функции распределения оказывается еще более сложной. Однако в теоретических исследованиях обычно приходится определять не температуру по известной функции распределения, а, наоборот, находить заполнение энергетических уровней по известной (измеренной) температуре.

Определенная таким образом температура может принимать для «квантовых систем» не только положительные, но и отрицательные значения. Отрицательные температуры будут соответствовать такому распределению частиц по энергиям, при котором коэффициенты заполнения верхних уровней больше, чем нижних (например, если заполнены верхние уровни, когда имеются незаполненные низшие уровни).

Связь между температурой и функцией распределения может быть выражена и в более простом виде. Допустим, что в системе имеется N_i частиц с энергией ε_i и $N_i + \Delta N_i$ частиц с энергией $\varepsilon_i + \Delta\varepsilon_i$. При переходе от одного значения энергии к другому изменение числа частиц на единичный интервал энергии будет равно $\Delta N_i/\Delta\varepsilon_i$. Эта величина может быть положительной или отрицательной в зависимости от того, увеличивается или уменьшается число N_i при переходе к большим энергиям; найденная для всех возможных значений энергии частиц, она является некоторой характеристикой состояния системы (так как связана с функцией распределения частиц по энергиям). Однако числа N_i зависят от размеров системы (от общего числа частиц). Поэтому более объективной (применимой к любым системам) характеристикой будет относительная величина

$$\frac{1}{N_i}\,\frac{\Delta N_i}{\Delta\varepsilon_i} = \theta. \qquad (2.42)$$

В статистической физике установлено, что θ — отрицательная величина, которая зависит только от температуры и обратно пропорциональна ей; мы получим *основную формулу статистической физики*, если подставим $\theta = -1/kT$:

$$\frac{1}{N_i}\frac{dN_i}{d\varepsilon_i} = -\frac{1}{kT}; \quad N_i = Ae^{-\varepsilon_i/kT}. \tag{2.41}$$

Постоянная A может быть вычислена из условия, что общее число частиц равно N:

$$N = \sum N_i; \quad A = \frac{N}{\sum e^{-\varepsilon_i/kT}}; \quad \frac{N_i}{N} = \frac{e^{-\varepsilon_i/kT}}{\sum e^{-\varepsilon_i/kT}}. \tag{2.42}$$

Однако от температуры зависит также и поведение о д н о й частицы, которая может переходить от одного уровня к другому, получая от других частиц или отдавая им энергию при тепловых взаимодействиях. Допустим, что среднее время пребывания частицы на уровне ε_i равно τ_i, а на уровне $\varepsilon_i + \Delta\varepsilon_i$ равно $\tau_i + \Delta\tau_i$. Величина

$$\frac{1}{\tau_i}\frac{\Delta\tau_i}{\Delta\varepsilon_i} = \theta \tag{2.43}$$

будет показывать относительное изменение времени пребывания на уровнях при переходе частицы от малых энергий к большим. Если в системе имеется много частиц N, то число частиц, находящихся на уровне ε_i в данный момент времени, будет пропорционально N и времени пребыва-

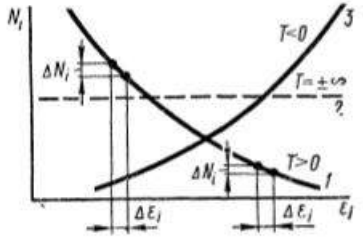


Рис. IV.68

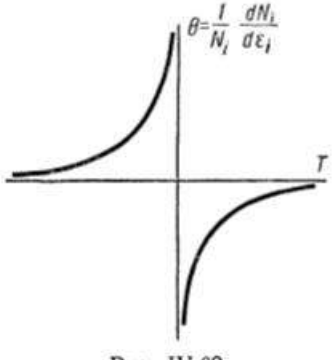


Рис. IV.69

ния на этом уровне, т. е. $N_i = \text{const} \cdot N \cdot \tau_i$. Подставляя в (2.43) значение τ_i, выраженное через N_i, мы получим соотношение (2.40).

Таким образом, температура оказалась величиной, показывающей, с *какой относительной плотностью заполнены энергетические уровни системы в различных местах спектра*: чем выше температура, тем равномернее распределены частицы по уровням энергии. При бесконечно большой температуре $\theta = 0$ и $dN_i/d\varepsilon_i = 0$, т. е. N_i одинакова для всех значений энергии. Это соответствует равномерному распределению энергии между частицами данной системы (или между ее степенями свободы, одинаково участвующими в общем обмене энергией). Заметим также, что в формуле (2.40) нет величин, характеризующих вещество, из которого состоит данное тело или система, что означает универсальность введенного таким образом понятия температуры.

Формула (2.40) поясняет также смысл понятия «отрицательная температура»; если при положительной температуре переход к большим энергиям сопровождается уменьшением плотности заполнения уровней $dN_i/d\varepsilon_i$, то при отрицательных температурах эта плотность увеличивается. На рис. IV.68 показаны графики зависимости N_i от ε_i при положительной (линия *1*) и отрицательной (линия *3*) температурах; прямая *2*, параллельная оси абсцисс, соответствует бесконечно большим температурам (N_i одинаково для всех значений ε_i). На рис. IV.69 показана зависимость относительной плотности заполнения уровней от температуры для классического распределения частиц по уровням (2.42). К идеальному газу, у которого температура свя-

зывается с средней кинетической энергией частиц, понятие отрицательной температуры неприменимо.

Величина

$$Z = \sum e^{-\varepsilon_i/kT} \tag{2.44}$$

называется *статистической суммой* и является характеристикой системы. В статистической термодинамике эта величина связывается с важнейшими термодинамическими функциями состояния; в частности, функция $\Psi = U - TS$, содержащая внутреннюю энергию и энтропию системы S, равна

$$\Psi = -k \ln Z.$$

Функция распределения частиц системы по ее энергетическому спектру, которую мы запишем в виде

$$\psi(\varepsilon) = Be^{-\varepsilon/kT}, \tag{2.45}$$

монотонно убывает при переходе от малых энергий к большим (при $T > 0$). Однако у некоторых очень важных в физике систем функции распределения (например, молекул по скоростям, см. ч. III, § 9; энергии излучения по спектру, см. ч. IV, § 15, и др.) имеют *максимумы* при некоторых значениях энергии, зависящих от температуры. Покажем, что наличие таких максимумов не только не противоречит указанному свойству функции $\psi(\varepsilon)$, но и следует из нее.

ЭНЕРГЕТИЧЕСКИЙ СПЕКТР СИСТЕМЫ

Допустим, что энергетический спектр системы является очень «плотным», т. е. состоит из очень большого числа очень близко расположенных уровней. Разделим такой спектр на равные участки $\Delta\varepsilon$, охватывающие несколько (Δn) уровней. Так как эти уровни очень близки друг к другу, то число частиц на каждом из них будет почти одинаковым и равным $N_i = N\psi(\varepsilon_i)$, поэтому суммарное число частиц на всех Δn уровнях будет равно (индекс опускаем):

$$\Delta N \approx N\psi(\varepsilon)\,\Delta n. \tag{2.46}$$

Число уровней на единичном интервале энергии $\Delta n/\Delta\varepsilon$ может быть различным в различных местах спектра; в пределе, когда уровни расположены бесконечно близко друг к другу (т. е. когда энергетический спектр является непрерывным), это отношение можно представить в виде функции от энергии:

$$\frac{dn}{d\varepsilon} = \omega(\varepsilon).$$

Тогда формулу (2.46) можно переписать в виде

$$dN = N\psi(\varepsilon)\,\omega(\varepsilon)\,d\varepsilon; \quad f(\varepsilon) = \frac{dN}{N} = Be^{-\varepsilon/kT}\,\omega(\varepsilon)\,d\varepsilon. \tag{2.47}$$

Здесь важно подчеркнуть, что в этой формуле ε означает не энергию отдельного уровня, а *среднюю энергию*, взятую в пределах участка $d\varepsilon$, а dN — число частиц не на одном (взятом внутри $d\varepsilon$) уровне, а на всех уровнях, энергии которых лежат в пределах ε, $\varepsilon + d\varepsilon$. В этом заключается существенное различие между функциями распределения $f(\varepsilon)$ и $\psi(\varepsilon)$. При переходе от малых значений энергии к большим функция $\psi(\varepsilon)$ *всегда убывает*, тогда как поведение $f(\varepsilon)$ определяется произведением двух функций: $e^{-\varepsilon/kT}$ и $\omega(\varepsilon)$. Если у какой-нибудь системы $\omega(\varepsilon)$ есть возрастающая функция, то $f(\varepsilon)$ будет иметь максимум при некотором значении ε.

Переход от систем с дискретным энергетическим спектром к системам с непрерывным спектром можно представить как постепенное увеличение плотности, с какой расположены уровни в энергетическом спектре. Однако в пределе, когда энергетический спектр становится непрерывным, функция $\omega(\varepsilon) = dn/d\varepsilon$ теряет свою наглядную трактовку как *число* уровней в единичном интервале энергии, но остается весьма важной характеристикой энергетического спектра системы. Кроме того,

изменяется также и «статистическая сумма» (2.44), которая состоит из конечных величин и при увеличении числа слагаемых делается бесконечно большой. Ввиду этого, согласно (2.42), становится бесконечно малым и число частиц N_i на каждом определенном уровне; в пределе, когда число уровней бесконечно большое, N_i равно нулю. Этот результат имеет важное значение при вероятностной трактовке статистических явлений как одно из основных положений (см. также § 10):

вероятность того, что какая-либо определенная («меченная») частица системы в данный момент времени имеет определенное значение энергии (т. е. находится на определенном уровне спектра), равна нулю; равна нулю также вероятность определенного размещения частиц системы по ее энергетическому спектру.

Ввиду этого имеет смысл находить только вероятность того, что данная частица имеет энергию, лежащую в пределах ε, $\varepsilon + d\varepsilon$. Так как эта вероятность, очевидно, будет пропорциональна ширине интервала $d\varepsilon$ и, кроме того, может быть различна в различных местах спектра, то ее записывают в виде

$$d\omega = f(\varepsilon)\, d\varepsilon.$$

Тогда число частиц, имеющих эти значения энергии, будет равно

$$dN = N\, d\omega = Nf(\varepsilon)\, d\varepsilon.$$

Таким образом, функция вероятности $f(\varepsilon)$ совпадает с функцией $\psi(\varepsilon)\omega(\varepsilon)$ в формуле распределения частиц по непрерывному спектру (2.47). Ввиду того что статистическая сумма Z для непрерывных спектров теряет смысл, в формуле (2.47) постоянная B (которая у дискретных спектров равнялась $1/Z$) должна быть вычислена заново:

$$\int_0^\infty \frac{dN}{N} = 1; \quad \int_0^\infty Be^{-\varepsilon/kT}\, \omega(\varepsilon)\, d\varepsilon = 1; \quad \frac{1}{B} = \int_0^\infty e^{-\varepsilon/kT}\, \omega(\varepsilon)\, d\varepsilon. \qquad (2.48)$$

Величина

$$Z = \int_0^\infty e^{-\varepsilon/kT}\, \omega(\varepsilon)\, d\varepsilon \qquad (2.49)$$

называется *статистическим интегралом* и является характеристикой систем с непрерывным энергетическим спектром. Так же как и у систем с дискретным спектром, эта величина связывается с термодинамическими функциями.

РАСПРЕДЕЛЕНИЕ МОЛЕКУЛ ПО СКОРОСТЯМ

Для иллюстрации рассмотрим распределение молекул идеального газа по скоростям. Соберем начала векторов скоростей всех молекул в одной точке O; тогда концы одинаковых векторов v будут расположены на сферической поверхности $4\pi v^2$. Каждый такой конец вектора будем называть *точкой*; очевидно, что точка изображает одну молекулу, имеющую определенную по величине и направлению скорость. Выберем в этом «пространстве скоростей» кольцевой слой (рис. IV. 70, *а*) толщиной dv и объемом $4\pi v^2 dv$. В этом слое будет находиться некоторое число (dN) точек, равное числу молекул, скорости которых лежат в пределах $v, v + dv$. Представим это число в виде

$$dN = \rho 4\pi v^2\, dv,$$

где ρ — «плотность», с какой этот слой заполнен точками. Проблема заключается в том, чтобы установить (экспериментально или теоретически — на основании дополнительных предположений), каковы значения ρ в различных местах «пространства скоростей».

Допустим сначала, что энергетический спектр молекул газа имеет *дискретный* характер и в рассматриваемом слое толщиной Δv находится Δn уровней; энергии этих уровней будут заключены в пределах $\varepsilon, \varepsilon + \Delta\varepsilon$. На рис. IV.70, *б* эти уровни условно изображены черточками, перпендикулярными линии $O\varepsilon$, на которой отло-

жены энергии молекул. Согласно основной формуле статистической физики, число молекул на каждом уровне с энергией ε равно

$$N_\varepsilon = NBe^{-\varepsilon/kT}.$$

Следовательно, число молекул в рассматриваемом слое будет равно

$$\Delta N = N_\varepsilon \Delta n = NBe^{-\varepsilon/kT}\Delta n.$$

Для расчета Δn должно быть известно: 1) либо число уровней, приходящихся на единичный интервал скоростей:

$$\sigma(v) = \frac{\Delta n}{\Delta v}; \quad \Delta n = \sigma \Delta v;$$

2) либо число уровней, приходящихся на единичный интервал энергии:

$$\omega(\varepsilon) = \frac{\Delta n}{\Delta \varepsilon}; \quad \Delta n = \omega \Delta \varepsilon.$$

Максвелловское распределение молекул по абсолютным скоростям (см. формулы (2.6) и (2.7), ч. II, § 9) может быть получено, если полагать, что ω пропорционально

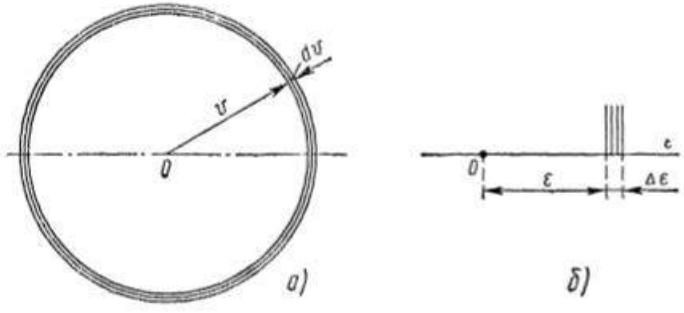


Рис. IV.70

скорости или σ пропорционально квадрату скорости. Аналогичный результат будет получен, если Δn пропорционально объему слоя $\Delta V = 4\pi v^2 \Delta v$. Однако все эти три условия есть дополнительные предположения к зависимости N_ε от энергии и температуры; они означают, что уровни энергии в «пространстве скоростей» (рис. IV.70, *а*) или вдоль «линии энергии» (рис. IV.70, *б*) должны быть расположены неравномерно. Естественное предположение, что плотность уровней σ одинакова для всех значений скоростей (т. е. что «пространство скоростей» однородное), не приводит к максвелловскому распределению.

§ 13. ВЗАИМОДЕЙСТВИЕ АТОМОВ И МОЛЕКУЛ. ЭНЕРГИЯ СВЯЗИ АТОМОВ

В § 12 было указано, что образование устойчивой физической системы возможно, если между ее частицами на больших (относительно) расстояниях действуют силы притяжения, а на малых — силы отталкивания. Природа (происхождение) этих сил, их зависимость от расстояния, а также потенциальная энергия взаимодействия системы обусловлены составом системы и свойствами частиц. Однако объяснение свойств системы на основании свойств ее составных частей оказывается, как правило, трудной задачей. Основное и принципиальное затруднение заключается в том, что истинная природа сил, действующих между объектами атомного масштаба, нам неизвестна и мы вынуждены пред-

полагать, что между ними действуют знакомые и хорошо изученные нами силы: гравитационные, электрические и магнитные. Лишь после того, как попытки объяснить свойства систем действием «классических» сил оказываются безуспешными, появляется необходимость в предположениях о существовании новых видов взаимодействия между частицами и соответствующих им сил. Так были введены в физику представления о «силах молекулярного взаимодействия», при помощи которых объясняется различие в уравнениях состояния идеальных и реальных газов; так появились «ядерные силы», действующие между протонами и нейтронами на близких расстояниях, и т. д. Однако следует отметить, что главные особенности сил, действующих в атомах и молекулах, могут быть объяснены, если полагать, что их природа — электрическая; существенно новое, внесенное квантовой теорией, — это необходимость учета волновых свойств частиц. Поэтому многие вопросы теории атомов и молекул требуют использования уравнения Шредингера и, хотя решение этого уравнения оказывается очень громоздким, во многих случаях удается получить приближенные количественные и ценные качественные результаты.

Как и в других областях физики, в теории атомов и молекул широко используются различные идеализированные модели. Следует отметить еще один прием, облегчающий исследование: так как между силами взаимодействия и потенциальной энергией существует непосредственная и простая связь (см., например, ч. I, § 12), то вместо недоступных измерению и контролю сил взаимодействия F определяют потенциальную энергию U, которую или легко вычислить, или найти экспериментально; если, например, выяснен вид функции $U(r)$, то $F = -\mathrm{d}U/\mathrm{d}r$.

КУЛОНОВСКОЕ ВЗАИМОДЕЙСТВИЕ АТОМОВ

Рассмотрим несколько примеров электрического взаимодействия двух частиц — атомов или молекул. Сначала предположим, что расположение электрических зарядов (в каждой из этих частиц) отличается

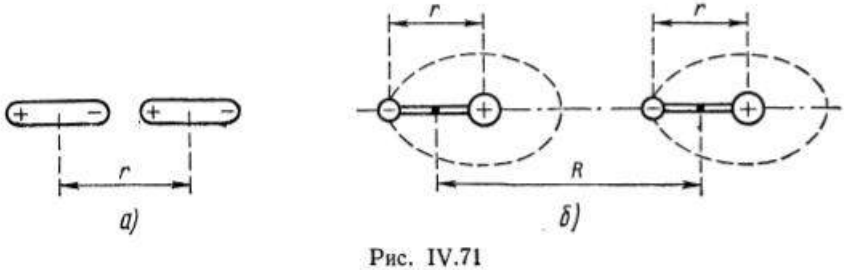


Рис. IV.71

от сферически симметричного, поэтому каждая из них имеет некоторый электрический момент (рис. IV.71, *а*). Допустим, что эти моменты одинаковы и равны p. При температуре абсолютного нуля эти частицы будут располагаться так, как это показано на рис. IV.71, *а*. Тогда их потенциальная энергия будет обратно пропорциональна кубу расстоя-

яния между ними (см. ч. III, § 4). Однако тепловое движение будет нарушать такое упорядоченное расположение взаимодействующих диполей, и если рассчитать среднее значение энергии взаимодействия при высоких температурах, то оказывается, что

$$U = -\frac{p^2}{24\pi^2\varepsilon_0^2 kT}\frac{1}{r^6}, \qquad (2.50)$$

следовательно, сила электрического притяжения между частицами будет обратно пропорциональна седьмой степени расстояния. Разумеется, для удержания этих частиц на расстоянии r необходимы силы отталкивания, но их происхождение должно быть выяснено отдельно.

Допустим теперь, что частицы в свободном состоянии не имеют электрического момента, но при сближении на малые расстояния r они, действуя друг на друга своими электрическими полями, вызывают перераспределение зарядов и тем самым возбуждают («индуцируют») появление дипольных моментов. При этом, как бы ни перемещались частицы друг относительно друга, направление индуцированного электрического момента всегда будет совпадать с линией, соединяющей частицы. Вследствие этого энергия взаимодействия (при фиксированном r) не зависит от наличия теплового движения. Характеристикой таких поляризующихся частиц является отношение приобретаемого дипольного момента p к напряженности внешнего поля E: $\alpha = p/E$. Согласно формуле Дебая, в этом случае

$$U = -\frac{\alpha p^2}{8\pi^2\varepsilon_0^2}\frac{1}{r^6}. \qquad (2.51)$$

Если взаимодействующие частицы представляют собой системы, подобные атому водорода (один заряд вращается относительно другого), то можно показать, что при согласованном вращении сумма сил притяжения между разноименными зарядами оказывается несколько больше суммы сил отталкивания одноименных зарядов. Действительно, в состоянии, показанном на рис. IV.71, *б*, разность этих сил равна

$$\Delta F = \sum (F_{\text{прит}} - F_{\text{отт}}) = \frac{e^2}{(R+r)^2} + \frac{e^2}{(R-r)^2} - \frac{e^2}{R^2} - \frac{e^2}{R^2} \approx 6\frac{e^2 r^2}{R^4}.$$

В других положениях сила взаимодействия таких «синхронно вращающихся» диполей будет иной, но ее среднее значение (за один оборот) будет отлично от нуля.

ВЗАИМОДЕЙСТВИЕ МОЛЕКУЛ

Некоторые факты указывают на существование сил притяжения также и между нейтральными атомами и молекулами, которые не имеют электрического момента и не приобретают его при взаимодействиях на малых расстояниях (например, взаимодействие атомов «благородных» элементов в их жидком и твердом состояниях и др.). Происхождение таких сил объясняется следующим образом. Внутри атомных систем существуют осциллирующие частицы, энергия которых, согласно квантовой теории (см. § 11), равна $E = h\nu_0 (n + {}^1/_2)$, где n — квантовое число, указывающее на наличие возбуждения. Однако в возбужден-

ном состоянии осциллятор, если он представляет собой электрическую систему, будет иметь дипольный момент; нам же необходимо выяснить природу сил взаимодействия между частицами, не имеющими этого момента. Поэтому, полагая $n = 0$, имеем $E_0 = h\nu_0/2$ («нулевая энергия осциллятора»).

Взаимодействие двух осцилляторов приводит к некоторому изменению их частот колебаний: один из них приобретает частоту $\nu_1 = \nu_0 + \Delta\nu_1$, другой — $\nu_2 = \nu_0 - \Delta\nu_2$. Расчет значений $\Delta\nu_1$ и $\Delta\nu_2$ показывает, что суммарная «нулевая энергия» обоих осцилляторов, равная до взаимодействия $h\nu_0/2 + h\nu_0/2 = h\nu_0$, при наличии этого взаимодействия оказывается меньше на величину ΔE, обратно пропорциональную шестой степени расстояния между центрами осцилляторов. Полагая, что ΔE есть появившаяся потенциальная энергия взаимодействия этих осцилляторов, получим

$$U = \Delta E = -\frac{\text{const}}{r^6}. \tag{2.52}$$

Отрицательный знак потенциальной энергии означает, что осцилляторы будут притягиваться друг к другу. Если взаимодействующие атомы или молекулы содержат несколько осцилляторов, то необходимо найти равнодействующую этих сил.

Таким образом, существование «нулевой энергии» у систем с квантовыми свойствами позволяет объяснить появление сил притяжения, не связанных с наличием постоянного или индуцированного дипольного момента. Поэтому эти силы, как дополнительные, будут действовать и в том случае, когда такие моменты имеются.

Силы, возникающие между атомами и молекулами вследствие наличия «нулевой энергии» у взаимодействующих осцилляторов, называются силами Ван-дер-Ваальса. Так как все атомы и молекулы представляют собой квантовые осцилляторы, то силы Ван-дер-Ваальса присутствуют во всех их взаимодействиях. При наличии у атомов и молекул электрического дипольного момента появляются два других вида взаимодействия: «ориентационное», описываемое формулой (2.50), и «индукционное», описываемое формулой (2.51).

Однако расчеты показывают, что взаимодействия только таких типов недостаточны для объяснения известных связей между атомами и молекулами. Например, при помощи этих сил невозможно объяснить наблюдаемую *прочность* атомных кристаллов, двухатомных молекул водорода, кислорода и других газов. Для этой цели понадобилось более детальное изучение роли электронов во взаимодействии атомов и молекул. Допустим, что взаимодействующие атомы состоят из ядра и одного валентного электрона, находящегося в атоме в основном состоянии (при котором вероятность нахождения электрона вокруг ядра имеет сферическую симметрию; см. § 11). Если расстояние r между атомами велико, то взаимодействие между ними будет очень слабым и потенциальную энергию взаимодействия можно полагать равной нулю. При малых расстояниях r электрон одного атома будет испытывать кулоновское притяжение к ядру другого атома и отталкиваться от связанного с ним электрона. Возникает проблема: найти распределе-

ние вероятностей нахождения обоих электронов в пространстве, окружающем оба ядра, т. е. в суммарном электрическом поле всех четырех зарядов. Решение этой задачи при помощи уравнения Шредингера приводится в пособиях по квантовой физике; укажем лишь важнейшие результаты:

1) вероятность нахождения электрона в единице объема в различных точках вокруг атомов (т. е. величина $|\psi|^2$) *не равна* сумме вероятностей, рассчитанных отдельно для каждого изолированного атома.

Вследствие интерференции дебройлевских волн обоих электронов в рассматриваемой системе «электронное облако» (которое ранее окружало каждое изолированное ядро в виде сферически симметричного кольцевого слоя) оказывается деформированным и теряет сферическую симметрию;

2) если спины электронов направлены в противоположные стороны, то эта вероятность *увеличивается* между атомами и *уменьшается* за их пределами;

3) более частое пребывание обоих электронов (с противоположными спинами) в пространстве между ядрами приводит к появлению потенциальной энергии у всей системы и вследствие этого — к появлению сил взаимодействия между ядрами.

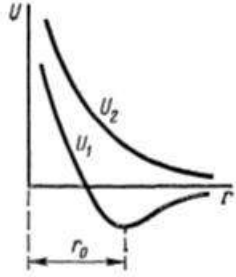


Рис. IV.72

Зависимость этой энергии U_1 от расстояния r между ядрами показана на рис. IV.72; на расстоянии $r = r_0$ эта энергия имеет минимум, поэтому при $r < r_0$ ядра будут отталкиваться друг от друга, а при $r > r_0$ — притягиваться; теоретический расчет для атома водорода ($U_1 = 4{,}37$ эВ и $r_0 = 0{,}735$ Å) оказывается очень близким к экспериментально найденным ($U_1 = 4{,}38$, $r_0 = 0{,}753$);

4) если спины электронов направлены в одну сторону, то потенциальная энергия взаимодействия в системе U_2 всегда оказывается положительной и непрерывно убывающей с расстоянием r, что соответствует силам отталкивания. Соответствующая зависимость U_2 от r показана на рис. IV.72.

Таким образом, благодаря волновым свойствам электронов возможны появления сил притяжения и отталкивания между электрически нейтральными атомами и молекулами. Существенно, что эти силы вызваны *двумя* электронами, поэтому такая связь называется *двухэлектронной*. Иногда эту связь называют *валентной* (так как в появлении этой связи участвуют валентные электроны) или *обменной*, имея в виду то обстоятельство, что «индивидуальная принадлежность» данного электрона определенному ядру в этой картине теряет смысл; электроны в суммарном электронном облаке оказываются «обобщенными» и одновременно «обслуживают» оба ядра. Эту связь также называют *гомеополярной*.

Важным свойством гомеополярной (валентной) связи является *насыщенность*. Если на незаполненной наружной оболочке атома имеется несколько валентных электронов, то он может иметь двухэлек-

тронные связи только с определенным числом соседних атомов. Это обстоятельство имеет весьма важное значение в химии, в теории кристаллических структур и т. д.

ИОННАЯ СВЯЗЬ

Для объяснения связи между атомами в молекулах типа NaCl, KJ и т. д. привлекается электрическое взаимодействие между противоположно заряженными ионами, которые образуются при переходе валентных электронов от одного атома к другому. Здесь действует «упорядочивающая тенденция» (см. § 12) к преимущественному замещению энергетических уровней, соответствующих малым энергиям. Переход валентного электрона, например, от натрия к хлору означает уменьшение энергии системы, переход ее в низшее квантовое состояние. Для объяснения прочности такой системы предполагается, что противоположно заряженные ионы притягиваются по закону Кулона и соответствующая этому потенциальная энергия равна

$$U_1 = -\frac{e^2}{4\pi\varepsilon_0 r}.$$

Однако при малых расстояниях появляются силы отталкивания, для которых, по М. Борну, потенциальная энергия должна быть представлена в виде

$$U_2 = \frac{B}{r^n}, \tag{2.53}$$

где B и n могут рассматриваться как постоянные величины. Этот вид связи называется *ионным*.

РОЛЬ ЭЛЕКТРОННОГО ГАЗА В МЕТАЛЛАХ

Однако в кристаллической структуре металла связь между атомами не может быть ионной или валентной. Ионы металла, расположенные в узлах решетки, имеют одноименные электрические заряды, а число валентных электронов явно недостаточно для образования двухэлектронной связи каждого атома со всеми окружающими его соседями. Поэтому приходится предположить, что при сближении атомов металла и образовании кристаллической решетки слабо связанные валентные электроны «обслуживают» уже не один определенный «канал связи», как это было в рассмотренной выше двухатомной молекуле, а принимают участие в образовании сил притяжения между многими атомами. Это означает, что вероятность нахождения любого электрона в единице объема $|\psi|^2$ должна быть *одинаковой* во всем объеме металлического кристалла, за исключением, разумеется, поверхностного слоя и ближайших окрестностей самих атомов (где расположены их электронные оболочки). Валентные электроны, покинувшие свои места в атомах, становятся *общими* («коллективными») для всех атомов. Они составляют так называемый «электронный газ», плотность которого в пределах металла одинакова и зависит от среднего расстояния между ио-

нами (т. е. узлами кристаллической решетки). Связь между ионами металла, обусловленная электронным газом, называется металлической.

Для объяснения появления свободных электронов рассмотрим систему из N изолированных атомов (или молекул), каждый из которых пред тавляет небольшую физическую систему с определенным дискретным спектром уровней энергии $\varepsilon_1, \varepsilon_2, \ldots$ В общем случае эти уровни могут быть вырожденными, т. е. уровень ε_i может охватывать z_i различных состояний. В случае атома водорода уровни энергии ε определяются главным квантовым числом n и каждый из них содержит $z_n = 2n^2$ различных состояний с одинаковыми энергиями ε_n. В сложных атомах энергия определяется двумя квантовыми числами n и l, поэтому одному значению энергии соответствует $2(2l + 1)$ различных состояний.

При соединении этих атомов в физическую систему появляется взаимодействие между ними и вследствие этого несколько изменяется структура спектра устойчивых состояний каждого атома, а также значение энергии уровней. Полученная система будет иметь свой спектр уровней $E_1, E_2, \ldots$, который, очевидно, может быть обнаружен в физических явлениях, в которых эта система участвует. В первую очередь нас будет интересовать лишь основное состояние с наименьшей энергией E_1; необходимо выяснить, в каких «внутренних» состояниях могут находиться составные части системы, если энергия всей системы равна E_1.

Изменение энергетического спектра данного атома мы объясняем воздействием на него электрического и магнитного поля других атомов. Поэтому, для того чтобы определить это изменение в общих чертах, следует выяснить, какие изменения в спектре состояний атома могут быть вызваны внешними (однородными) полями. Допустим, что атом, имеющий в своей оболочке Z электронов, помещен в электрическое поле с напряженностью E. Энергия каждого электрона, связанного в атоме, изменится на некоторую величину w. Для внутренних электронов, у которых энергия связи W очень велика, дополнительная энергия w будет значительно меньше W и поэтому изменением уровней энергии внутренних электронов можно пренебречь. У внешних электронов энергия связи W значительно меньше и поэтому добавление w может заметно изменить их состояние.

Однако величина w оказывается различной для различных состояний, охватываемых одним уровнем, так как она зависит от ориентации плоскости орбиты и спина электрона относительно внешнего поля. Следовательно, те состояния с различными квантовыми числами m и l, которые в изолированном атоме принадлежали одному уровню энергии, будут во внешнем электрическом поле иметь различные (хотя и очень близкие) значения энергии. Это означает (см. § 12) устранение вырождения атомных уровней энергии, т. е. *расщепление* каждого уровня энергии на $2(2l + 1)$ близких подуровней. Очевидно, что разность энергий двух соседних подуровней будет увеличиваться с увеличением напряженности внешнего электрического поля. Следует заметить, что влияние ориентации спина электрона на значение энергии оказывается слабым, поэтому в некоторых рассуждениях и расчетах спиновое расщепление не принимается во внимание.

В кристаллической решетке на данный атом действует электрическое поле соседних атомов. В этом случае также произойдет смещение энергии уровней на величину w, которая опять-таки различна для различных состояний (охватываемых одним уровнем). Вследствие этого произойдет расщепление уровней и появление множества подуровней. Рассмотрим изменение энергии у одного из валентных электронов; может оказаться, что полученная под действием электрического поля соседних атомов энергия w равна или больше энергии связи W. В этом случае валентный электрон сможет покинуть свой атом и свободно перемещаться по объему металла. Так образуется электронный газ в металлических кристаллах. Очевидно, что свободные электроны могут быть получены и в тех кристаллах, в которых $W > w$, но только благодаря «туннельному эффекту» (см. § 11); плотность электронного газа в таких кристаллах будет зависеть от «ширины» потенциального барьера (равной расстоянию между атомами в решетке) и от ее «высоты», равной $U = W - w$.

Заметим, что малые значения энергий w и W не означают, что валентные электроны могут быть удалены от атома слабыми электрическими полями. Рассчитаем, например, напряженность E внешнего электрического поля, которая необходима, чтобы «разорвать» атом водорода (удалить валентный электрон в бесконечность). Для этого, очевидно, силы $\pm eE$, действующие на протон и электрон в противоположных направлениях, должны быть больше кулоновской силы притяжения между ними. Если электрон находится на первой орбите ($r = 0{,}53 \cdot 10^{-10}$ м), то

$$eE = k\frac{e^2}{r^2}; \quad E = \frac{ke}{r^2} = \frac{9 \cdot 10^9 \cdot 1{,}6 \cdot 10^{-19}}{0{,}28 \cdot 10^{-20}} \approx 5{,}7 \cdot 10^{11} \text{ В/м}.$$

Такая исключительно большая напряженность поля может существовать только вблизи зарядов малых размеров (протоны, ядра атомов, ионы). Сама же энергия электрического взаимодействия валентного электрона ввиду малости его заряда оказывается незначительной:

$$W = \int_r^\infty eE\,\mathrm{d}r = k\frac{e^2}{r} = 9 \cdot 10^9 \frac{(1{,}6 \cdot 10^{-19})^2}{0{,}53 \cdot 10^{-10}} \approx 4{,}35 \cdot 10^{-18} \text{ Дж} \approx 27 \text{ эВ}.$$

Выяснение структуры энергетического спектра системы взаимодействующих атомов и молекул представляет собой важную, но трудную задачу. Однако для объяснения свойств этих систем знание спектра уровней совершенно необходимо.

Перечисленные выше виды взаимодействия между атомами и молекулами характеризуются величиной «энергии связи», т. е. потенциальной энергией взаимодействия, приходящейся на один моль вещества. Наиболее универсальная (действующая между любыми атомами и молекулами) связь типа Ван-дер-Ваальса оказывается слабой и имеет энергию связи порядка 10^3 Дж/моль; валентная и металлическая — 10^5 Дж/моль, ионная — около 10^7 Дж/моль.

Глава 3

ИЗЛУЧЕНИЕ И ПОГЛОЩЕНИЕ ЭНЕРГИИ АТОМАМИ И МОЛЕКУЛАМИ

§ 14. АТОМНЫЕ И МОЛЕКУЛЯРНЫЕ СПЕКТРЫ ИЗЛУЧЕНИЯ; СПОСОБЫ ВОЗБУЖДЕНИЯ

КВАНТОВЫЕ ЧИСЛА

Различные состояния сложных атомов (как изолированных, так и помещенных во внешнее магнитное поле) характеризуются, как и у атома водорода, четырьмя квантовыми числами. Однако в отличие от чисел n, l, m и s, применяемых для нумерации состояний идеализированного атома водорода (модель Н. Бора), более удобной оказывается другая совокупность квантовых чисел:

1) *главное* квантовое число n, как и в атоме водорода, определяет энергетический уровень E_n, который, однако, состоит из некоторого набора близко расположенных друг к другу энергетических подуровней. Главное квантовое число имеет значения $n = 1, 2, 3, \ldots$;

2) *орбитальное* квантовое число l учитывает дискретные значения момента импульса L_{op} электронной орбиты ($L_{op} = l \cdot h/2\pi$). В частности, при $l = 0$ этот момент отсутствует, так как в этом состоянии вероятность нахождения электрона в атоме имеет сферическую симметрию и понятие орбиты теряет смысл. Для заданного числа n орбитальное квантовое число принимает значения $l = 0, 1, 2, \ldots, n - 1$;

3) *внутреннее* квантовое число j выражает дискретные значения суммарного (L) момента импульса электронной орбиты L_{op} и самого электрона (спина). Так как при данном $\mathbf{L}_{op}$ спин электрона $s = 1/2$ может иметь только две возможные ориентации (параллельно и антипараллельно направлению $\mathbf{L}_{op}$), то $\mathbf{L} = \mathbf{L}_{op} + \mathbf{s}$ и поэтому третье квантовое число имеет только два различных значения:

$$j = l \pm \frac{1}{2}.$$

Для $l = 0$ и $L = 0$ суммарный момент импульса будет состоять только из спина электрона; в этом случае взаимная ориентация векторов L_{op} и s теряет смысл и поэтому j имеет только одно значение, равное 1/2;

4) *магнитное* квантовое число m учитывает пространственное квантование состояний атома во внешнем магнитном поле H. Полный вектор момента импульса атома L может иметь только определённые проекции на направление поля: от $+L$ (соответствующего параллельности L и H) до $-L$ (соответствующего антипараллельности этих векторов); промежуточные значения этой проекции отличаются на $h/2\pi$:

$$+L; \quad L - (h/2\pi); \quad L - 2\,(h/2\pi); \ldots; \quad -L,$$

поэтому четвертое квантовое число имеет значения:

$$m_j = +j;\quad j-1;\quad j-2;\ \ldots;\quad -j,$$

причем для каждого из двух значений: $j = l + 1/2$ и $j = l - 1/2$ — должен быть составлен отдельный набор значений m_j (важная роль внешнего магнитного поля объясняется тем, что и орбиты электронов и сами электроны имеют магнитные моменты, пропорциональные соответствующим моментам импульсов L_{op} и s, и поэтому взаимодействуют с внешним полем.

Приведем таблицу различных значений квантовых чисел для $n = 1$, 2 и 3:

n	l	j	m_j
1	0	1/2	−1/2; +1/2
2	0	1/2	−1/2; +1/2
	1	3/2	−3/2; −1/2; +1/2; +3/2
		1/2	−1/2; +1/2
3	0	1/2	−1/2; +1/2
	1	3/2	−3/2; −1/2; +1/2; +3/2
		1/2	−1/2; +1/2
	2	5/2	−5/2; −3/2; −1/2; +1/2; +3/2; +5/2
		3/2	−3/2; −1/2; +1/2; +3/2

СТРУКТУРА ЭЛЕКТРОННОЙ ОБОЛОЧКИ

Совокупность всех электронов, расположенных (или, по модельным представлениям, вращающихся) вокруг ядра, называется *электронным облаком* или *электронной оболочкой* атома. Эта оболочка делится на «слои», соответствующие энергетическим уровням E_n; число n означает номер слоя, начиная с ближайшего к ядру. Приняты условные названия: $n = 1$ — K-слой; $n = 2$ — L-слой; $n = 3$ — M-слой, $n = 4$ — N-слой и т. д.

Каждый слой разделяется на группы и подгруппы, характеризуемые числами l и j (т. е. каждый энергетический уровень E_n содержит некоторое число подуровней, энергия которых определяется этими числами). Подуровни, соответствующие числам $l = 0$, обозначаются буквой s и содержат только одно состояние ($j = 1/2$). Подуровни, для которых $l = 1, 2, 3, 4, \ldots$, обозначаются соответственно буквами p, d, f, g, ... и содержат по два состояния. Например, если электрон находится в состоянии с $n = 2$, $l = 1$ и $j = 5/2$, то этот уровень энергии условно записывается в виде $2p_{5/2}$ (первая цифра означает главное

квантовое число, буква — подуровень, а дробь справа — число j). Формирование электронных оболочек при переходе от одних атомов периодической системы к другим подчиняется «упорядочивающей тенденции» (см. § 12): по мере увеличения числа протонов в ядре, каждый новый электрон занимает то место в электронной оболочке, при котором приращение энергии атома минимальное. Расчеты показывают, что при малых значениях n разность энергии электрона в двух соседних слоях E_n и E_{n+1} превосходит разность энергии подуровней, отличающихся при данном n значениями l. Поэтому при переходе от одного атома к последующему заполняются сначала уровни энергии с малыми значениями n. Однако начиная с калия обнаруживаются отступления от этого правила: валентный электрон калия занимает не свободный уровень в M-слое ($n = 3$), а переходит в N-слой ($n = 4$) в состоянии 4s. Это означает, что с возрастанием числа протонов в ядре и общего числа электронов вокруг них спектры энергий, соответствующих двум соседним значениям n, накладываются друг на друга.

Таким образом, можно заметить некоторые сходства и различие между схемой уровней энергии атома водорода и распределением элементов в периодической системе Менделеева. Если пронумеровать все состояния атома водорода по возрастающим значениям энергии, то соответствующие им квантовые числа n, l, m и s будут аналогичны квантовым числам n, l, j и m_j, которые определяют положение электронов в сложных атомах. Различия в этих схемах вызваны взаимодействием между электронами, которое отсутствует в атоме водорода, аналогия же обусловлена корпускулярно-волновыми свойствами электронов и общей «упорядочивающей тенденцией», направленной (у сложных систем) к созданию структур с наименьшей возможной энергией.

Химические и оптические свойства веществ определяются электронами, находящимися на периферийных орбитах; они слабее других связаны с остальной частью атома.

Вещества, у которых в атомах полностью заполнены K- и L-слои, а также s- и p-состояния ($l = 0, 1$) в других слоях, химически инертны (например, у гелия заполнен K-слой, у неона — K- и L-слои, у аргона K- и L-слои, а также s- и p-состояния M-слоя и т. д.). Если для удаления валентного электрона у натрия необходима энергия 5,14 эВ, то для удаления электрона с заполненной оболочки соседнего неона необходима энергия 21,5 эВ.

СПЕКТРЫ ИЗЛУЧЕНИЯ АТОМОВ

Сведения о структуре электронной оболочки атомов получены при анализе испускаемого ими спектра, когда внешнее воздействие на атомы сведено к минимуму (излучение разреженного газа). Дополнительные сведения получаются при изучении спектра излучения атомов, помещенных во внешнее электрическое и магнитное поле. Рассмотрим сначала спектр излучения свободных атомов; пронумеруем дискретные значения энергии возбужденного атома, начиная с наименьшего (не возбужденного) состояния:

$$E_1,\ E_2,\ E_3,\ \ldots,\ E_n,\ \ldots,$$

где E_1 — энергия основного состояния атома, остальные соответствуют возбужденным состояниям. Нас может интересовать:

1) *состав спектра* (какие частоты колебаний или длины волн имеются в изучении данного атома);

2) *относительная интенсивность* спектральных линий;

3) *ширина* каждой спектральной линии (см. § 9);

Состав спектра излучения определяется возможными переходами из состояний с большей энергией в какое-нибудь состояние с низшей энергией; излучаемые атомом частоты определяются по формуле

$$h\nu_{nk} = E_k - E_n.$$

Однако выяснилось, что не все переходы осуществляются с одинаковой вероятностью; некоторые переходы реализуются настолько редко, что соответствующие им спектральные линии имеют чрезвычайно малую интенсивность или вовсе не регистрируются; такие переходы называются *запрещенными*. Квантовая физика, рассчитывающая вероятности нахождения электрона в различных местах атома, позволяет решать и задачу о вероятности перехода. Например, эта вероятность оказалась большой для переходов, при которых квантовое число l изменяется на единицу: $\Delta l = \pm 1$ — или если внутреннее квантовое число изменяется на $\Delta j = 0; \pm 1$ (*правило отбора* квантовых переходов в атомах). Таким образом, не только состав спектра, но и наблюдаемые относительные интенсивности спектральных линий получают теоретическое объяснение.

Некоторые атомы имеют структуру, отличающуюся от атома водорода только зарядом и массой ядра (например, ионизированные атомы гелия и лития с одним электроном). Спектр излучения таких «водородоподобных» атомов по структуре не отличается от спектра водорода, а формула для энергии состояний содержит число элементарных зарядов ядра Z:

$$E_n = -\frac{1}{n^2}\frac{Z^2me^4}{8h^2\varepsilon_0^2}, \qquad (2.54)$$

где n — номер возбужденного состояния.

Сравнительно легко можно группировать спектральные линии в излучении таких атомов, которые состоят из ядра, окруженного «компактной» электронной оболочкой (с заполненными уровнями l), и единственного электрона, вращающегося на относительно больших расстояниях от ядра и оболочки (щелочные элементы). Энергия возбужденных состояний таких атомов представляется в виде

$$E_n = -\frac{R}{(n+s)^2}, \qquad (2.55)$$

где R — постоянная величина, а s — некоторая поправка, учитывающая взаимодействие наружного электрона с ядром и остальными электронами.

Ширина спектральных линий атомов определяется временем пребывания их в возбужденных состояниях. Если излучающий атом не подвергается внешним воздействиям, то ширина спектральной линии (называемая *естественной*) может быть вычислена при помощи соот-

ношения неопределенностей Гейзенберга (см. § 10). Так как значения энергии уровней могут быть известны только с неточностью $\Delta E \geqslant \geqslant h/2\pi\Delta t$, где Δt — время пребывания атома в возбужденном состоянии, то частота колебаний будет иметь естественный разброс, равный $\Delta\nu = \Delta E/h = 1/2\pi \cdot \Delta t$. У атомов, находящихся во внешнем электрическом или магнитном поле или под воздействием соседних атомов, время пребывания Δt в возбужденных состояниях уменьшается, и это вызывает дополнительное увеличение ширины спектральных линий.

СПЕКТРЫ ИЗЛУЧЕНИЯ МОЛЕКУЛ

Спектр излучения молекул определяется также структурой их энергетических уровней. В свободном состоянии энергия каждой молекулы может быть представлена в виде трех компонентов, соответствующих трем различным видам движения в молекуле:

1) энергия движения всех электронов (как «принадлежащих» отдельным атомам, так и «обобщенных»);

2) энергия колебательного движения атомов внутри молекулы;

3) энергия вращения атомов вокруг центра массы молекулы.

В этот перечень не входит энергия, заключенная внутри атомных ядер, так как связанное с ней излучение (гамма-лучи) не включается в рассматриваемый спектр. Таким образом, энергия свободной молекулы

$$E = E_{эл} + E_{кол} + E_{вр}. \qquad (2.56)$$

Первая компонента содержит совокупность электронных уровней энергии отдельных атомов, деформированную вследствие взаимодействия между ними (уровни несколько смещены и расширены; см. § 13). Вторая и третья компоненты, согласно квантовой теории, также оказываются состоящими из дискретного набора уровней; колебательная энергия (см. § 11, формула 2.29) состоит из «нулевой энергии» основного состояния и ряда возбужденных состояний, соответствующих квантовым числам $n_k = 1, 2, 3, \ldots$:

$$E_{кол} = h\nu_0\left(\frac{1}{2} + n_k\right).$$

Энергия вращательного движения состоит из уровней, соответствующих дискретному набору моментов количества движения $L = I\omega$ (I — момент инерции, ω — угловая скорость вращения):

$$L = \frac{h}{2\pi}\sqrt{l_{вр}(l_{вр}+1)}; \quad E = \frac{I\omega^2}{2} = \frac{h^2}{8\pi^2 I} l_{вр}(l_{вр}+1), \qquad (2.57)$$

где $l_{вр} = 0, 1, 2, \ldots$ Структура спектра излучения молекулы будет определяться совокупностью этих уровней и разрешаемыми переходами (правилами отбора: изменение квантовых чисел для колебательного и вращательного движения должно быть не более $\Delta n_k = \pm 1$; $\Delta l_{вр} = \pm 1$).

Частоты спектральных линий, испускаемых молекулами, могут соответствовать переходам от одного электронного уровня энергии к другому или от одного колебательного (или вращательного) к другому. Однако возможны и переходы от уровней с одними значениями $E_{кол}$ и $E_{вр}$ к уровням, имеющим другие значения энергии всех трех компонент, вследствие чего возникают линии, соответствующие комбинированным («колебательно-вращательным» или «электронно-колебательным») спектрам.

Таким образом, спектр молекул оказывается очень сложным. В некоторых участках спектральные линии оказываются расположенными настолько густо, что с трудом разрешаются. В других участках они перекрываются, образуя более или менее широкие полосы с непрерывным распределением частот колебаний *(полосатые спектры)*. При очень тесной связи между атомами, какая имеется в жидком и твердом состояниях, весь излучаемый спектр оказывается непрерывным.

В нормальном состоянии любой физической системы каждая частица занимает состояние с наименьшей возможной энергией E_0. Если в системе имеется N одинаковых частиц, которые не могут находиться в одинаковых состояниях, то они располагаются в системе, занимая уровни с возрастающим значением энергии, без пропусков (вакансий). Для того чтобы перевести систему в возбужденное состояние, необходимо сообщить энергию либо одной, либо одновременно нескольким частицам, т. е. перевести их на более высокие вакантные уровни. При этом, очевидно, система может поглощать энергию только определенными дозами, т. е. ей невозможно сообщить энергию, большую или мёньшую, чем эти дозы. Ввиду этого внешний источник энергии должен либо иметь в точности такие дозы, либо же располагать непрерывным спектром значений энергии, из которых система могла бы выбрать требуемые количества энергии.

Возбуждение атомов и молекул может быть произведено различными способами:

1) нагреванием до температур, при которых энергия хаотического теплового движения частиц (порядка kT) сделается достаточной для возбуждения;

2) облучением электромагнитными волнами (фотонами), начиная от радиотехнического и оптического диапазона и кончая рентгеновскими и гамма-лучами;

3) бомбардировкой частицами, обладающими достаточной энергией.

Первый способ является универсальным; все тела при температурах выше абсолютного нуля возбуждаются тепловым движением и излучают. Спектр излучения определяется составом и состоянием (газообразным или конденсированным) вещества, а интенсивность излучения существенно зависит от температуры. Второй способ у некоторых веществ сопровождается свечением, среди которых важное значение имеет люминесценция. Третий способ используется для получения излучения очень малых длин волн и прежде всего рентгеновских лучей. Перечисленные способы возбуждения рассматриваются в следующих параграфах.

Важные сведения о колебательных процессах, происходящих в атомах и молекулах твердых и жидких тел, были получены при изучении света, рассеянного различными веществами.

Если на вещество действует монохроматическое излучение, имеющее частоту ν_0, то в рассеянном излучении кроме этой частоты обнаруживаются весьма слабые излучения с частотами

$$\nu_1' = \nu_0 + \Delta\nu_1', \qquad \nu_2' = \nu_0 - \Delta\nu_2',$$
$$\nu_1'' = \nu_0 + \Delta\nu_1'', \ \ldots; \qquad \nu_2'' = \nu_0 - \Delta\nu_2'', \ \ldots,$$

т. е. в веществе происходит комбинация частоты внешнего воздействия ν_0 с какими-то частотами колебаний самого вещества. Такое рассеяние называется *комбинационным*; оно было обнаружено Л. И. Мандельштамом и Г. С. Ландсбергом в твердых телах и Раманом — в жидких (1928).

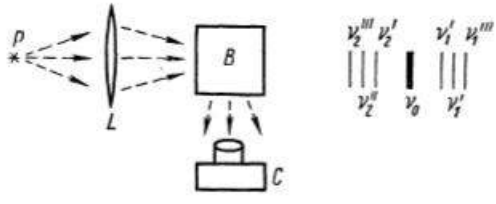


Рис. IV.73

Схема опыта для наблюдения комбинированного рассеяния показана на рис. IV.73, где P — источник монохроматического света, L — линза, фокусирующая излучение на исследуемое вещество B, C — спектрограф, определяющий частоты колебаний (или длины волн) в рассеянном излучении и их интенсивность; справа схематически показано расположение новых линий относительно ν_0 (или λ_0). Линии, имеющие ме́ньшую частоту колебаний (большую длину волны), условно называются «красными» спутниками; линии, имеющие бо́льшую частоту колебаний, — «фиолетовыми».

Изучение спектра комбинационного рассеяния показало, что:

1) линии спутников (ν_1', ν_1'', ... ; ν_2', ν_2'', ...) располагаются симметрично по обе стороны от линии падающего излучения (ν_0), т. е.

$$\nu_{кр} = \nu_0 - \Delta\nu; \quad \nu_0 - \nu_{кр} = \nu_{фиол} - \nu_0; \quad \nu_{фиол} = \nu_0 + \Delta\nu;$$

2) интенсивность «красных» (длинноволновых) спутников почти не зависит от температуры рассеивающего вещества, тогда как интенсивность «фиолетовых» (коротковолновых) спутников быстро увеличивается с повышением температуры.

При нормальных температурах интенсивность «красных» спутников больше, чем интенсивность фиолетовых;

3) разности частот $\Delta\nu' = \nu_0 - \nu_{кр}'$; $\Delta\nu'' = \nu_0 - \nu_{кр}''$ и т. д., а также $\Delta\nu' = \nu_{фиол}' - \nu_0$; $\Delta\nu'' = \nu_{фиол}'' - \nu_0$ и т. д. не зависят от частоты ν_0 действующего на вещество излучения, а определяются только рассеивающим веществом, т. е. характеризуют его состав и структуру. Число появляющихся при рассеивании спутников также определяется веществом.

Явление комбинационного рассеяния можно объяснить, рассматривая его как результат взаимодействия падающего светового фотона

с молекулой; при этом фотоны могут изменять направление своего движения, что и приводит к рассеянию света. В большинстве случаев имеет место упругое соударение, а так как масса молекулы много больше массы фотона, то энергия падающего фотона не изменяется и в рассеянном спектре наблюдаются линии падающего света, которые мы называем основными. Однако наряду с таким рассеянием возможно и рассеяние с частичным изменением энергии фотона. При столкновениях фотон может или отдать часть своей энергии молекуле и перевести ее из нормального состояния в возбужденное, или же получить энергию от молекулы, переходящей из возбужденного состояния в нормальное. В первом случае рассеянный фотон будет обладать меньшей частотой — образуется «красный» спутник; во втором случае — большей частотой и образуется «фиолетовый» спутник. При этом следует иметь в виду, что энергия молекулы может изменяться только дискретно.

Допустим, что на возбуждение молекулы требуется затратить некоторую энергию A; такое же количество энергии молекула отдает, переходя в нормальное состояние. Если часть энергии падающего фотона расходуется на возбуждение молекулы, то энергия рассеянного («красного») фотона

$$h\nu_{кр} = h\nu_0 - A.$$

Если же падающий фотон получает энергию от ранее возбужденной молекулы, то энергия рассеянного («фиолетового») фотона определится из равенства

$$h\nu_{ф} = h\nu_0 + A.$$

Из этих равенств следует, что частоты рассеянного света будут соответственно равны:

$$\nu_{кр} = \nu_0 - \frac{A}{h}; \quad \nu_{ф} = \nu_0 + \frac{A}{h},$$

т. е. спутники расположены симметрично относительно линий первичного спектра. Так как число возбужденных молекул значительно меньше числа молекул невозбужденных, то интенсивность «фиолетовых» спутников меньше, чем «красных». При нагревании число возбужденных молекул быстро растет, в связи с чем растет и интенсивность «фиолетовых» спутников.

Молекула может находиться в различных энергетических состояниях, поэтому в рассеянном свете могут появиться различные спутники с большими и меньшими частотами по сравнению с частотой падающего света. Так возникают различные «комбинационные» частоты: $\nu_1 = \nu_0 \pm \Delta\nu'$; $\nu_2 = \nu_0 \pm \Delta\nu''$ и т. д.

Комбинационное рассеяние широко используется в научных и технических целях: 1) спектры комбинационного рассеяния столь характерны для рассеивающего вещества, что по ним возможно отличать не только гомологические, но и изомерные органические соединения, для анализа которых химия часто не имеет средств; 2) комбинационное рассеяние является мощным средством для изучения строения моле-

кул, их внутренних колебаний и вращательных движений; 3) комбинационное рассеяние используется для установления индивидуального состава бензинов, масел, продуктов нефтепереработок и др.

§ 15. ТЕПЛОВОЕ ИЗЛУЧЕНИЕ

Излучение атомов и молекул может происходить при одновременном действии различных способов возбуждения. *Тепловое излучение есть только та часть излучения данного тела, для которого источником энергии является хаотическое тепловое движение частиц вещества*; это излучение имеется при всех температурах, отличных от абсолютного нуля, и существенно зависит от величины температуры.

Напомним две особенности теплового движения в термодинамических системах (см. ч. II, § 1, 3): 1) тенденция к достижению равновесного состояния с одинаковой температурой по объему системы и 2) тенденция к равновесному распределению энергии между частицами и между их степенями свободы. Обе эти тенденции действуют и в том случае, когда обмен энергиями между телами, их частицами и степенями свободы частиц осуществляется не при тепловых столкновениях, а путем излучения и поглощения фотонов. Вследствие этого в замкнутой системе (если, например, тела находятся внутри некоторого объема с идеально отражающими стенками) со временем устанавливается определенное р а в н о в е с н о е распределение энергии по спектру. Экспериментальное и теоретическое излучение этого распределения имело весьма важное значение в физике.

Состав излучения от какого-нибудь тела характеризуется не только набором частот (или длин волн) в спектре этого излучения, но и энергиями, которые приходятся на отдельные участки спектра. Поэтому при экспериментальном изучении спектров различных тел пользуются приборами, регистрирующими энергию очень узких участков спектра. Простейшими из них являются: 1) фотографическая пластинка (пленка), почернение которой пропорционально поглощенной энергии излучения; 2) миниатюрные термопары, тонкий спай которых, поглощая излучение, нагревается и возбуждает термоэлектродвижущую силу, пропорциональную поглощенной энергии; 3) болометры — тонкие полоски зачерненной платины (или другого тела, хорошо поглощающего излучение всех длин волн), нагревание которых определяется по изменению их электрического сопротивления, и др. Разлагая излучение в спектр и перемещая датчики вдоль спектра, можно определить распределение энергии излучения по его отдельным, почти монохроматическим участкам. При этом должно учитываться также поглощение энергии в приборе, разлагающем излучение в спектр, и отражение излучения самими датчиками; оба эти фактора могут быть различными для различных участков спектра.

ЗАКОН КИРХГОФА

Излучение, которое находится в равновесии с нагретыми телами, должно иметь определенное распределение энергии по спектру. Докажем это; допустим, что внутри объема V, стенки которого не пропу-

скают излучения, находятся тела A, B, C, ... и «фотонный газ», содержащий фотоны различных энергий: $h\nu_1$, $h\nu_2$, ... («телами» являются также стенки, ограничивающие объем V). Обозначим через $N(\nu)$ число фотонов, которые падают в единицу времени на единичную поверхность одного из тел. Часть из них ($N_{\text{погл}}$) поглотится, а другая часть ($N_{\text{отр}}$) отразится. Заметим, что коэффициенты поглощения $\alpha = N_{\text{погл}}/N$ и отражения $\rho = N_{\text{отр}}/N$ могут быть различными для фотонов различных частот и, кроме того, могут зависеть от вещества и температуры тела, на которое падают эти фотоны, т. е. $\alpha = \alpha(\nu, T)$, $\rho = \rho(\nu, T)$. Однако рассматриваемое тело, имея некоторую температуру, само испускает такие фотоны.

Обозначим через $P(\nu)$ число фотонов частоты (ν), испускаемых этим телом в единицу времени с единицы поверхности. При равновесном обмене фотонами между данным телом и «фотонным газом» должно наблюдаться равенство:

$$N_{\text{погл}}(\nu) = P(\nu); \quad \alpha(\nu) N(\nu) = P(\nu); \quad \frac{P(\nu)}{\alpha(\nu)} = N(\nu).$$

Числа $P(\nu)$ определяются не только веществом, но и температурой данного тела. Однако отношение $P(\nu)/\alpha(\nu)$ может не зависеть от температуры. В этом случае условие равновесия между «фотонным газом» и телами A, B, C, ... имело бы вид

$$N(\nu) = \left| \frac{P(\nu, T_1)}{\alpha(\nu, T_1)} \right|_A = \left| \frac{P(\nu, T_2)}{\alpha(\nu, T_2)} \right|_B = \ldots,$$

т. е. тела могли бы иметь различные температуры. Но если отношение $P(\nu, T)/\alpha(\nu, T)$ есть функция от температуры тела, то $N(\nu)$ также будет зависеть от T и тогда переход к равновесному состоянию внутри объема V приведет к выравниванию температур ($T_1 = T_2 = \ldots = T$) *независимо* от вещества тел, находящихся внутри этого объема. Таким образом, в объеме V, содержащем различные тела с различными температурами, равновесное состояние установится со временем только при условии, если отношение $P(\nu, T)/\alpha(\nu, T)$ для всех тел (независимо от их вещества) есть *одинаковая функция* от частоты ν и температуры тела T.

Эти рассуждения можно повторить и на основе волновой теории излучения. Допустим, что на единичную поверхность одного из находящихся внутри объема V тел падает монохроматический участок спектра, ограниченный длинами волн λ, $\lambda + d\lambda$ и содержащий поток излучения

$$d\Phi_{\lambda T} = \Phi_e \psi(\lambda, T)\, d\lambda, \tag{2.58}$$

где $\psi(\lambda, T)$ — функция распределения мощности (энергии) по спектру падающего излучения (см. формулу (1.3)). Часть этого потока $(d\Phi_{\lambda T})_{\text{погл}}$ поглотится, вторая часть $(d\Phi_{\lambda T})_{\text{отр}}$ отразится с поверхности тела. Коэффициенты поглощения (α) и отражения (ρ) могут быть функциями от длины волны и температуры, различными для различных веществ:

$$\alpha_{\lambda T} = \frac{(d\Phi_{\lambda T})_{\text{погл}}}{d\Phi_{\lambda T}}; \quad \rho_{\lambda T} = \frac{(d\Phi_{\lambda T})_{\text{отр}}}{d\Phi_{\lambda T}}. \tag{2.59}$$

Допустим, что рассматриваемое тело при данной температуре T с единицы поверхности в единицу времени излучает монохроматический поток $\mathrm{d}R_{\lambda T}$, который можно представить в виде

$$\mathrm{d}R_{\lambda T} = r_{\lambda T}\,\mathrm{d}\lambda. \tag{2.60}$$

Функция от длины волны и температуры $r_{\lambda T}$ называется *спектральной плотностью энергетической светимости* (а также *спектральной плотностью излучательности*) данного тела. Условием равновесия между излучением, находящимся в объеме V, и собственным излучением данного тела будет равенство

$$(\mathrm{d}\Phi_{\lambda T})_{\text{погл}} = \mathrm{d}R_{\lambda T}; \quad \alpha_{\lambda T}\Phi_e\psi(\lambda, T)\,\mathrm{d}\lambda = r_{\lambda T}\,\mathrm{d}\lambda. \tag{2.61}$$

Повторив рассуждения, приведенные выше для тел, взаимодействующих с «фотонным газом», можно прийти к тому же выводу: в объеме V установится равновесное состояние с одинаковыми температурами тел, если только отношение $r_{\lambda T}/\alpha_{aT} = \Phi_e\psi(\lambda, T)$ будет для всех тел одинаковой функцией от длины волны и температуры. Обозначим эту функцию через $u_{\lambda T}$:

$$\frac{r_{\lambda T}}{\alpha_{aT}} = u_{\lambda T}. \tag{2.62}$$

Полученный вывод, подтверждаемый измерениями, называется з а к о н о м К и р х г о ф а:

отношение спектральной плотности энергетической светимости к спектральному коэффициенту поглощения есть для всех тел одна и та же функция от температуры и длины волны.

Из этого закона следует, что тела, имеющие в какой-либо области спектра большой коэффициент поглощения, обладают в этой области также и большой испускательной способностью.

Среди тел, находящихся внутри рассматриваемого объема V, могут быть тела, *полностью поглощающие все участки спектра падающего излучения*, т. е. такие, у которых $\alpha_{\lambda T}$ равно единице для всех длин волн и при любой температуре. Такие идеально поглощающие тела называются *абсолютно черными*. Согласно закону Кирхгофа, испускательная способность таких тел должна быть больше, чем у любых других тел (у которых $\alpha_{\lambda T} < 1$), поэтому абсолютно черное тело называют также *полным излучателем*.

Напишем закон Кирхгофа (2.62) для различных тел (A, B, ...) и для полного излучателя (O):

$$\left(\frac{r_{\lambda T}}{\alpha_{\lambda T}}\right)_A = \left(\frac{r_{\lambda T}}{\alpha_{\lambda T}}\right)_B = \ldots = \left(\frac{r_{\lambda T}}{1}\right)_O = u_{\lambda T}. \tag{2.63}$$

Отсюда следует, что функция $u_{\lambda T}$ есть спектральная плотность энергетической светимости абсолютно черного тела. Для других (реальных) тепловых излучателей

$$r_{\lambda T} = \alpha_{\lambda T} u_{\lambda T}. \tag{2.64}$$

Таким образом, согласно закону Кирхгофа, тепловое излучение *любого* тела ($\alpha_{\lambda T} < 1$) в любой области спектра всегда *меньше*, чем у аб-

солютно черного тела. Если в пределах всего спектра (или достаточно широкого участка) коэффициент поглощения оказывается постоянным, то тело называется *серым*.

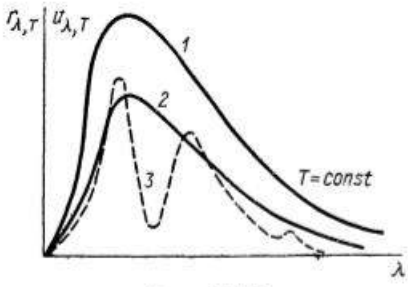


Рис. IV.74

На рис. IV.74 даны опытные кривые распределения энергии в спектре абсолютно черного *1*, серого *2* и произвольного *3* тела. Кривая спектрального распределения для серого тела может быть получена из кривой распределения для абсолютно черного тела путем умножения ординат последней на постоянный множитель, меньший единицы и равный коэффициенту поглощения серого тела. Таково приблизительно излучение вольфрамовой проволоки в электрических лампах.

Излучение некоторых тел является *селективным* (избирательным). Кривая излучения *3* таких тел может иметь несколько максимумов и минимумов, но для всех тел она лежит ниже кривой излучения абсолютно черного тела, как следует из закона Кирхгофа.

На рис. IV.75 графически изображена зависимость спектрального коэффициента поглощения некоторого тела от длины волны при данной температуре T; для сравнения показаны соответствующие графики для абсолютно черного ($\alpha = 1$) и серого ($\alpha = \text{const} < 1$) тел. При изменении температуры характер кривой $\alpha = f(\lambda, T)$ может измениться; лучи, сильно поглощающиеся при одной температуре, могут слабо поглощаться при другой температуре, и наоборот.

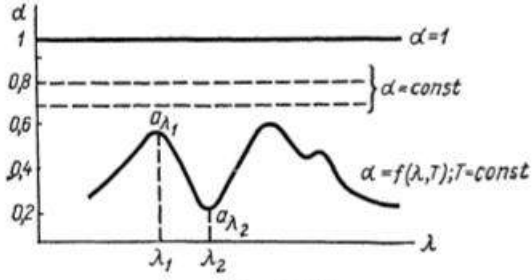


Рис. IV.75

В таблице даны коэффициенты поглощения для некоторых веществ при различных температурах (при условии, что толщина тела достаточно велика и та часть излучения, которая прошла через границу тела, полностью поглощается):

Вещество	t, °C	Коэффициент поглощения	Вещество	t, °C	Коэффициент поглощения
Железо окисленное	100	0,736	Платина полированная	225—625	0,054—0,104
Золото полированное	225—635	0,018—0,035	Картон асбестовый	24	0,96
Медь полированная	80—115	0,018—0,023	Вода	0—100	0,95—0,963
Медь расплавленная	1075—1275	0,16—0,13	Кирпич	20	0,93
Никель полированный	225—375	0,07—0,087	Сажа	95—270	0,952
			Бархат черный	20	0,996

Так как на поверхности тела часть излучения отражается, то приведенный в таблице коэффициент поглощения равен $1 - \rho$, где ρ — коэффициент отражения.

Кроме спектральных коэффициентов отражения и поглощения, а также — спектральной излучательности каждое тело характеризуется также и соответствующими интегральными величинами.

Энергетической светимостью (излучательностью) данного тела называют суммарную лучистую энергию, испускаемую единицей поверхности тела в единицу времени; эта величина будет функцией от спектрального состава излучения данного тела и его температуры:

$$R_e(T) = \int_0^\infty dR_{\lambda T} = \int_0^\infty r_{\lambda T}\, d\lambda. \qquad (2.65)$$

Допустим, что на тело падает излучение сложного состава, характеризуемое некоторой функцией распределения; коэффициенты отражения и преломления, полученные для такого излучения, также будут зависеть от состава падающего излучения и температуры самого тела; обозначим их через $\rho(T)$ и $\alpha(T)$. Повторяя рассуждения, приведенные выше для монохроматических участков спектра, можно получить интегральный закон Кирхгофа: *отношение энергетической светимости тела к его коэффициенту поглощения есть для всех тел одинаковая функция от температуры*:

$$\frac{R_e(T)}{\alpha(T)} = U(T). \qquad (2.66)$$

Для абсолютно черного тела $\alpha(T) = 1$, поэтому универсальная функция $U(T)$ есть энергетическая светимость полного излучателя.

Пропорциональность между испускательными способностями тел и их коэффициентами поглощения можно показать на примере сажи или платиновой черни, которые имеют большой коэффициент поглощения и большую плотность излучательности. Внесем внутрь нагретой полости (например, внутрь муфельной печи) фарфоровый черепок, часть которого зачернена сажей, платиновой чернью или тушью. Когда черепок нагреется и примет температуру полости, отличить темные места черепка от светлых, незачерненных, оказывается невозможным. Темные места больше поглощают, но и больше излучают, а светлые места меньше излучают, зато больше отражают. Вынув черепок из печи, мы заметим яркое свечение зачерненных мест, так как теперь они больше излучают, чем светлые места (отраженное излучение отсутствует, так как нет падающего излучения).

Получение света от пламени горящей свечи основано на той же пропорциональности между испускательной и поглощательной способностями тел. В пламени имеются частицы сажи, обладающие большим поглощением; они и дают яркий свет. Если пламя не содержит частиц сажи (например, пламя газовой горелки), оно не будет светиться. Следует заметить, что не всякое черное тело полностью поглощает падающее на них излучение; в невидимой области спектра (ультрафиолетовой или инфракрасной) они могут иметь большие коэффициенты отражения.

Абсолютно черных тел (полностью поглощающих излучение всех длин волн) в природе нет, но можно указать на тело, которое по своим свойствам практически не будет отличаться от абсолютно черного. Такой моделью абсолютно черного тела является полость с очень малым отверстием (рис. IV. 76). Луч (любой длины волны), попавший внутрь такой полости, может выйти из нее обратно только после многократных отражений. При каждом отражении от стенок полости часть энергии луча

поглощается и лишь ничтожная доля энергии лучей, попавших в отверстие, сможет выйти обратно, поэтому коэффициент поглощения отверстия оказывается весьма близким к единице. Такая «модель» абсолютно черного тела может быть нагрета до высоких температур; тогда из отверстия в полости выходит интенсивное излучение

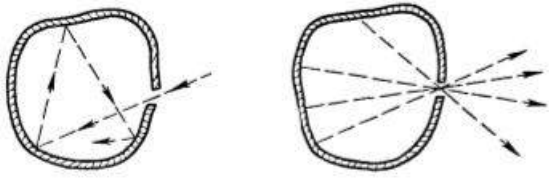


Рис. IV.76

и отверстие будет ярко светиться (при этом оно по-прежнему остается абсолютно поглощающим). Топочное устройство с «глазком» в плавильных или коксовых печах, муфельные печи с отверстием, зрачок глаза являются такого же рода «моделями», весьма близкими к абсолютно черному телу.

ФОРМУЛА ПЛАНКА

Законы Кирхгофа для теплового излучения тел (монохроматического или интегрального) позволяют определить спектральные и интегральные характеристики *излучения* различных тел, если экспериментально измерены соответствующие коэффициенты *поглощения*. При этом должно быть известно равновесное распределение энергии в спектре полного излучения (абсолютно черного тела), т. е. вид функции $u(\lambda, T)$. Так как упомянутые выше «модели» не являются достаточно точным приближением к полному (идеальному) излучателю, то были сделаны попытки теоретического вывода этой функции с последующим сопоставлением с результатами измерений. Эта задача была успешно решена М. Планком (1900), предложившим идею о квантовом характере излучения атомов и молекул. Для спектральной плотности энергетической светимости полного излучателя им была получена формула (в функции от ν или λ)

$$u_{\nu T} = \frac{2\pi\nu^2}{c^2} \frac{h\nu}{e^{h\nu/kT} - 1}; \qquad u_{\lambda T} = \frac{2\pi hc^2}{\lambda^5} \frac{1}{e^{hc/\lambda kT} - 1}, \tag{2.67}$$

которая хорошо согласуется с результатами измерений, проведенных на «моделях» абсолютно черного тела. Пользуясь формулами (2.67), можно рассчитать ту часть энергии, которая содержится в единице объема и приходится на монохроматический участок спектра. Эта энергия связана с $u_{\lambda T}$ соотношением $dU_{\nu T} = \frac{4}{c} u_{\nu T} d\lambda$ и равна:

$$dU_{\nu T} = \frac{8\pi\nu^2}{c^3} \frac{h\nu}{e^{h\nu/kT} - 1} d\nu; \quad dU_{\lambda T} = \frac{8\pi hc}{\lambda^5} \frac{1}{e^{hc/\lambda kT} - 1} d\lambda. \tag{2.68}$$

Имеется несколько различных подходов к выводу этой формулы; приведем вывод, основанный на использовании статистики Бозе —

Эйнштейна и идеи Планка о квантовом характере излучения. Воспользуемся представлением о «фотонном газе» и отметим различие между ним и идеальным газом:

1) в идеальном газе функция распределения молекул по энергиям переходит к равновесному виду благодаря столкновениям. В фотонном газе обмена энергиями между фотонами не происходит и их распределение по энергиям изменяется только теми телами, которые поглощают и испускают фотоны. Ввиду этого тела («тепловые излучатели») навязывают фотонному газу определенное распределение энергии по спектру, соответствующее распределению «осцилляторов» самого тела по их энергетическим уровням;

2) число молекул идеального газа при нагревании и охлаждении остается постоянным, тогда как число фотонов в равновесном тепловом излучении (в единице объема) увеличивается с повышением температуры;

3) молекулы идеального газа имеют одинаковые массы, но различные скорости, импульсы и энергии; фотоны же имеют одинаковые скорости, но различные массы, импульсы и энергии.

Ввиду этого вместо «пространства скоростей», которое было использовано для молекул идеального газа (см. рис. IV.70, § 12), для фотонного газа воспользуемся «пространством импульсов», в котором каждая «точка» A будет обозначать один фотон, движущийся в направлении OA и имеющий импульс $h\nu/c$ (рис. IV.77). Число фотонов, энергия которых заключена в интервале $\varepsilon_1 = h\nu$, $\varepsilon_2 = h(\nu + d\nu)$, будем полагать пропорциональным объему кольцевого слоя с радиусом $h\nu/c$ и толщиной $d(h\nu/c)$:

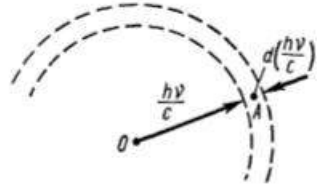


Рис. IV.77

$$dN_\nu = \rho 4\pi (h\nu/c)^2 \, d(h\nu/c),$$

где ρ — функция от частоты колебаний и температуры тела — есть «плотность», с какой заполнен «точками» этот слой. Это число необходимо удвоить ввиду того, что в каждом направлении могут распространяться два фотона одинаковой энергии с противоположными направлениями поляризации.

Плотность ρ, с которой заполнено такими «точками» пространство импульсов, обусловлена той плотностью, с которой «частицы-осцилляторы» в тепловых излучателях заполняют соответствующие участки энергетического спектра. Воспользуемся простым приемом: 1) каждый атом, испускающий набор частот $\nu_1, \nu_2, \ldots, \nu_z$, можно заменить z атомами, каждый из которых испускает фотоны только одной частоты; 2) атом, испускающий фотоны n раз в секунду, можно заменить «подсистемой», имеющей в своем составе n возбужденных «частиц-осцилляторов», излучающих один раз в секунду. Благодаря этому тепловой излучатель может быть уподоблен системе, рассмотренной в § 12. В формуле (2.37) положим $\varepsilon_0 = 0$, $\varepsilon = h\nu$ (фотон либо имеет энергию $h\nu$,

либо не существует); относительное число возбужденных осцилляторов с энергией $h\nu$ есть относительное число ежесекундно испускаемых фотонов, имеющих эту же энергию, поэтому

$$\rho = \frac{1}{e^{h\nu/kT} - 1}.$$

Умножив dN_ν на два и на $h\nu$ и подставив значение ρ, получим

$$\frac{8\pi\nu^2}{c^3} h^3 \frac{h\nu}{e^{h\nu/kT} - 1} d\nu,$$

которое в h^3 раз больше, чем $dU_{\nu T}$. Этот результат означает, что «пространство импульсов» для фотонов необходимо разделить на равные клетки, имеющие объем h^3; в каждой такой клетке может находиться только одна точка, изображающая один фотон. Этот прием широко применяется в квантовой физике.

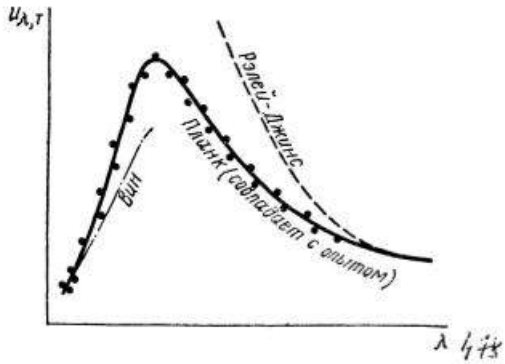


Рис. IV.78

В области очень коротких ($h\nu \gg kT$) и очень длинных ($h\nu \ll kT$) участков спектра формулу Планка можно заменить приближенными формулами:

$$u_{\lambda T} = a\lambda^{-5} e^{-b/\lambda T} \quad \text{(формула Вина)},$$

$$u_{\lambda T} = 8\pi\lambda^{-4} kT \quad \text{(формула Релея — Джинса)}.$$

На рис. IV.78 даны графики, соответствующие этим формулам; точками изображены результаты измерений, подтверждающие формулу Планка.

ЗАКОНЫ ИЗЛУЧЕНИЯ ЧЕРНОГО ТЕЛА

Установлены следующие законы излучения абсолютно черного тела (полного излучателя).

Закон Стефана—Больцмана. *Энергетическая светимость абсолютно черного тела пропорциональна четвертой степени его абсолютной температуры*:

$$R = \sigma T^4. \tag{2.69}$$

Величина σ называется *постоянной Стефана—Больцмана*:

$$\sigma = 5{,}6703 \cdot 10^{-8} \text{ Дж} \cdot \text{м}^{-2} \cdot \text{с}^{-1} \cdot \text{К}^{-4}.$$

Энергия, испускаемая за время t абсолютно черным телом с излучающей поверхностью S при постоянной температуре T,

$$W = \sigma T^4 S t. \tag{2.70}$$

Если же температура тела изменяется со временем, т. е. $T = T(t)$,

$$W = \int_0^t \sigma T^4(t)\, S\, dt.$$

Закон Стефана—Больцмана указывает на чрезвычайно быстрый рост мощности излучения с возрастанием температуры. Например, при повышении температуры от 800 до 2400 К (т. е. от 527 до 2127° С) излучение абсолютно черного тела возрастает в 81 раз.

Если абсолютно черное тело окружено средой с температурой T_0, то оно будет поглощать энергию, излучаемую самой средой.

В этом случае разность между мощностью испускаемого и поглощаемого излучений можно приближенно выразить формулой

$$U = \sigma (T^4 - T_0^4).$$

К реальным телам закон Стефана—Больцмана неприменим, так как наблюдения показывают более сложную зависимость R от температуры, а также от формы тела и состояния его поверхности.

Закон смещения Вина. *Длина волны λ_0, на которую приходится максимум спектральной плотности энергетической светимости абсолютно черного тела, обратно пропорциональна абсолютной температуре тела*:

$$\lambda_0 = \frac{b}{T}, \quad \text{или} \quad \lambda_0 T = b. \quad (2.71)$$

Константа b, называемая *постоянной закона Вина*, равна $b = 2{,}8978 \cdot 10^{-3}$ м·К (λ выражена в метрах).

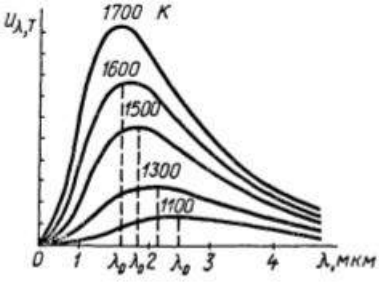


Рис. IV.79

Таким образом, при повышении температуры растет не только полное излучение, но, кроме того, изменяется распределение энергии по спектру. Например, при малых температурах тела излучают главным образом инфракрасные лучи, а по мере повышения температуры излучение делается красноватым, оранжевым и, наконец, белым. На рис. IV.79 показаны эмпирические кривые распределения энергии излучения абсолютно черного тела по длинам волн при разных температурах: из них видно, что максимум спектральной плотности излучения при повышении температуры смещается в сторону коротких волн.

Закон Стефана—Больцмана получается из формулы Планка путем интегрирования по всем частотам (или длинам волн) спектра при постоянной температуре:

$$U_T = \int_0^\infty u_{\lambda T}\, d\lambda = \sigma T^4,$$

причем для постоянной σ получается выражение

$$\sigma = 1{,}08 \frac{12\pi h}{c^2} \left(\frac{k}{h}\right)^4. \quad (2.72)$$

Подстановка входящих в эту формулу констант дает величину, совпадающую с эмпирическим значением σ.

Закон смещения Вина и его константу можно получить из формулы Планка нахождением максимума функции $u_{\lambda T}$, для чего берется про-

изводная от $u_{\lambda T}$ по λ и приравнивается нулю. Вычисление приводит к формуле

$$\lambda_0 T = \frac{ch}{4{,}96k} = b.$$

Расчет постоянной b по этой формуле также дает результат, совпадающий с эмпирическим значением постоянной Вина.

ПРИМЕНЕНИЕ ЗАКОНОВ ТЕПЛОВОГО ИЗЛУЧЕНИЯ

Рассмотрим важнейшие применения законов теплового излучения.

Тепловые источники света. Большинство искусственных источников света является тепловыми излучателями (электрические лампы накаливания, обычные дуговые лампы и т. д.). Однако эти источники света не являются достаточно экономичными.

В § 2 было сказано, что глаз обладает чувствительностью только к очень узкому участку спектра (от 380 до 770 нм); все остальные волны не оказывают зрительного ощущения. Максимальная чувствительность глаза соответствует длине волны $\lambda = 0{,}555$ мкм. Исходя из этого свойства глаза, следует требовать от источников света такого распределения энергии в спектре, при котором максимальная спектральная плотность излучения падала бы на длину волны $\lambda = 0{,}555$ мкм или около нее. Если в качестве такого источника взять абсолютно черное тело, то по закону смещения Вина можно вычислить его абсолютную температуру:

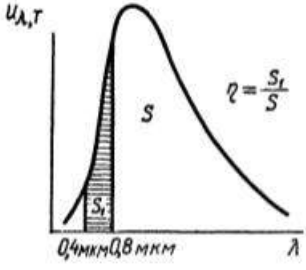


Рис. IV.80

$$T = \frac{0{,}002898 \text{ м} \cdot \text{К}}{0{,}555 \cdot 10^{-6} \text{ м}} \approx 5200 \text{ К}.$$

Таким образом, наиболее выгодный тепловой источник света должен иметь температуру 5200 К, что соответствует температуре солнечной поверхности. Такое совпадение является результатом биологического приспособления человеческого зрения к распределению энергии в спектре солнечного излучения. Но и у этого источника света коэффициент полезного действия (отношение энергии видимого излучения к полной энергии всего излучения) будет невелик. Графически на рис. IV.80 этот коэффициент выражается отношением площадей S_1 и S; площадь S_1 выражает энергию излучения видимой области спектра, S — всю энергию излучения.

Расчет показывает, что при температуре около 5000—6000 К световой к. п. д. равен всего 14—15% (для абсолютно черного тела). При температуре же существующих искусственных источников света (≈ 3000 К) этот к. п. д. составляет всего около 1—3%.

Такая невысокая «световая отдача» теплового излучателя объясняется тем, что при хаотическом движении атомов и молекул возбуждаются не только световые (видимые), но и другие электромагнитные волны, которые не оказывают светового воздействия на глаз. Поэтому невозможно избирательно заставить тело излучать только те волны, к которым чувствителен глаз: обязательно излучаются и невидимые волны.

Важнейшие из современных температурных источников света — это электрические лампы накаливания с вольфрамовой нитью. Температура плавления вольфрама равна 3655 К. Однако нагрев нити до температур выше 2500 К опасен, так как вольфрам при этой температуре очень быстро распыляется и нить разрушается. Для уменьшения распыления нити было предложено наполнять лампы инертными газами (аргон, ксенон, азот) при давлении около 0,5 атм. Это позволило поднять температуру нити до 3000—3200 К. При этих температурах максимум спектральной плотности излучения лежит в области инфракрасных волн (около 1,1 мкм), поэтому все современные лампы накаливания имеют к. п. д., немногим больший 1%.

Оптическая пирометрия. Изложенные выше законы излучения черного тела позволяют определять температуру этого тела, если известна длина волны λ_0, соответствующая максимуму $u_{\lambda T}$ (по закону Вина), или если известна величина интегральной плотности излучения (по закону Стефана—Больцмана). Эти методы определения температуры тела по его тепловому излучению называются *оптической пирометрией*; они особенно удобны при измерении очень высоких температур. Так как упомянутые законы применимы только к абсолютно черному телу, то оптическая пирометрия, основанная на них, дает хорошие результаты только при измерении температур тел, близких по своим свойствам к абсолютно черному. На практике таковыми являются заводские печи, лабораторные муфельные печи, топки котлов и т. п. Рассмотрим три способа определения температуры тепловых излучателей:

1. *Метод, основанный на законе смещения Вина.* Если нам известна та длина волны, на которую приходится максимум спектральной плотности излучения, то температура тела может быть вычислена по формуле (2.71).

В частности, таким способом определяется температура на поверхности солнца, звезд и т. д.

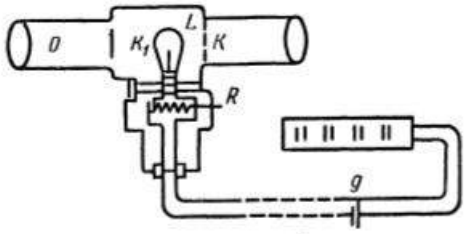


Рис. IV.81

Для нечерных тел этот способ не дает истинную температуру тела; если в спектре излучения имеется один максимум и мы рассчитаем T по формуле (2.71), то расчет дает нам температуру абсолютно черного тела, имеющего почти такое же распределение энергии в спектре, как и испытуемое тело. При этом цветность излучения абсолютно черного тела будет одинакова с цветностью исследуемого излучения. Такая температура тела называется *цветовой температурой*. Цветовая температура нити лампы накаливания равна 2700—3000 К, что очень близко к ее истинной температуре.

2. *Радиационный способ измерения температур* основан на измерении энергетической светимости тела R и вычислении его температуры по закону Стефана—Больцмана. Соответствующие приборы называются *радиационными пирометрами*.

Естественно, что если излучающее тело не является абсолютно черным, то радиационный пирометр не дает истинной температуры тела, а покажет ту температуру абсолютно черного тела, при которой энергетическая светимость последнего равна энергетической светимости испытуемого тела. Такая температура тела называется *радиационной* или *энергетической* температурой.

Из недостатков радиационного пирометра укажем на невозможность его применения для определения температур небольших объектов, а также на влияние среды, находящейся между объектом и пирометром, которая поглощает часть излучения.

3. *Яркостный метод определения температур.* Принцип действия его основан на визуальном сравнении яркости раскаленной нити лампы пирометра с яркостью изображения накаленного испытуемого тела. Схема прибора дана на рис. IV.81. Он представляет собой зрительную трубу с помещенной внутри электрической лампой, питаемой от аккумулятора. На рисунке : O — объектив, L — лампа, K — светофильтры с узкой полосой пропускания (в области длины волны 0,665 мкм), R — реостат, g — шкала. Равенство яркостей, наблюдаемое через монохроматический фильтр, определяется по исчезновению изображения нити на фоне изображения раскаленного тела. Накал нити регулируется реостатом, а температура определяется по шкале амперметра, градуированного прямо на температуру.

Если температура тела очень велика, то на пути лучей помещается дымчатый фильтр K_1.

§ 16. РЕНТГЕНОВСКИЕ ЛУЧИ. ХАРАКТЕРИСТИЧЕСКИЙ РЕНТГЕНОВСКИЙ СПЕКТР ЭЛЕМЕНТОВ

Рентгеновское излучение возбуждается различными способами, в частности при столкновении электронов (обладающих большими энергиями) с атомами тяжелых элементов. По своей природе оно не

отличается от светового: 1) не отклоняется в электрических и магнитных полях и, следовательно, не имеет электрического заряда; 2) действует на фотографические эмульсии; 3) способно возбуждать люминесценцию; 4) может отражаться, преломляться, обладает поляризацией; 5) может интерферировать и образует дифракционные максимумы и минимумы.

Рентгеновские лучи, так же как и оптические, обнаруживают корпускулярно-волновые свойства. В некоторых явлениях (распространения дифракции, интерференции и т. п.) рентгеновские лучи ведут себя как электромагнитное излучение с очень короткими длинами волн, примерно от 0,06 до $120 \cdot 10^{-10}$ м; в процессах же взаимодействия с частицами вещества рентгеновские лучи ведут себя как поток фотонов с энергиями $h\nu = 1{,}7 \cdot 10^{-16} — 3{,}3 \cdot 10^{-14}$ Дж. Со стороны длинных волн рентгеновские лучи сливаются с коротковолновыми ультрафиолетовыми лучами, а со стороны коротких волн — с длинноволновыми гамма-лучами радиоактивных веществ.

ПОЛУЧЕНИЕ РЕНТГЕНОВСКИХ ЛУЧЕЙ

Рентгеновские лучи могут быть получены при бомбардировке металлических мишеней: 1) электронами, 2) протонами, альфа-частицами и другими массивными частицами, 3) под действием гамма-лучей радиоактивных веществ (длинноволновые рентгеновские лучи могут быть получены также при облучении мишеней коротковолновыми рентгеновскими лучами).

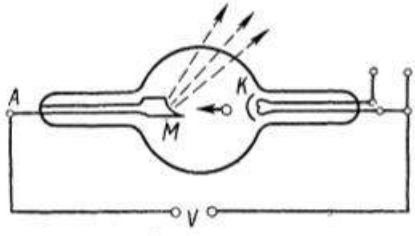


Рис. IV.82

Рассмотрим сначала первый (наиболее распространенный) способ. В технике и медицине используются *рентгеновские трубки*, схема которых показана на рис. IV.82. В этой схеме: K — катод (источник электронов), M — бомбардируемая металлическая мишень (иногда называется антикатодом), A — анод (в маломощных и портативных трубках мишенью может служить анод). Для сообщения электронам больших энергий между катодом и анодом прикладывается высокое напряжение V порядка десятков и сотен тысяч вольт. Так как при бомбардировке электронами мишень сильно нагревается, то ее приходится охлаждать каким-либо способом.

В зависимости от способа эмиссии электронов из катода рентгеновские трубки делятся на *газонаполненные* (или ионные) и *вакуумные* (с нагреваемым катодом). Первые содержат воздух при давлении 10^{-2}—10^{-4} мм рт. ст. Приложенное к трубке электрическое напряжение вызывает в газе разряд и сообщает электронам, движущимся к аноду, и положительным ионам, движущимся к катоду, большие скорости. Положительные ионы при ударе о катод выбивают из него добавочные электроны. Рентгеновские лучи возбуждаются из мишени (антикатода) при достаточно большой энергии бомбардирующих электронов.

В вакуумных рентгеновских трубках источником электронов является подогреваемый (электрическим током) катод, например вольфрамовая нить накаливания. Давление воздуха в трубке не превышает 10^{-6}—10^{-7} мм рт. ст., поэтому электроны, испаряющиеся из катода, почти свободно ускоряются электрическим полем, приобретая скорости до сотен тысяч километров в секунду (в зависимости от приложенного напряжения V).

При ударе о мишень электроны резко тормозятся и в процессе уменьшения своей кинетической энергии излучают электромагнитные волны. Интенсивность этого излучения определяется мощностью потока электронов (от числа электронов, ежесекундно падающих на мишень, и от их энергии).

Экспериментальное изучение интенсивности рентгеновских лучей показало (рис. IV.83):

1) существование минимальной длины волны $\lambda_{мин}$, не зависящей от вещества мишени и определяемой только приложенным к трубке напряжением V;

2) существование максимума интенсивности при некоторой длине волны λ_0; с повышением приложенного напряжений этот максимум смещается в сторону коротких волн.

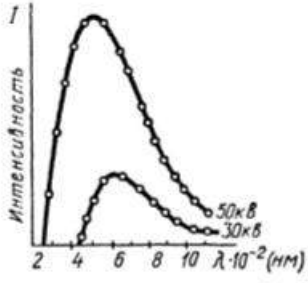


Рис. IV.83

Согласно классическим представлениям, тормозное электромагнитное излучение электронов должно теоретически содержать волны всех частот, поэтому резкой границы в спектре рентгеновского излучения не должно быть. Согласно данным измерений (рис. IV.83), в области длинных волн такая граница отсутствует, но коротковолновая часть спектра оказывается резко ограниченной. Квантовая же теория позволяет объяснить и непрерывный характер тормозного рентгеновского излучения, и существование резкой коротковолновой границы. Согласно закону сохранения энергии, при ударе электрона о мишень часть его кинетической энергии пойдет на создание рентгеновского фотона с энергией $h\nu$, а остальная часть A перейдет в энергию теплового движения атомов мишени, а также останется у электрона в виде его кинетической энергии внутри мишени:

$$\frac{mv^2}{2} = h\nu + A. \qquad (2.73)$$

При случайном характере тепловых столкновений энергия A может иметь всевозможные значения от нуля до $mv^2/2$; поэтому наиболее длинные волны будут соответствовать $A = mv^2/2$, $\nu = 0$, $\lambda = \infty$; наиболее короткие волны будут вызваны при $A = 0$ электронами, имеющими наибольшую кинетическую энергию:

$$\left(\frac{mv^2}{2}\right)_{макс} = h\nu_{макс} = \frac{hc}{\lambda_{мин}}.$$

Так как энергия свободного электрона внутри мишени, а также начальная кинетическая энергия электрона при вылете из катода очень

малы по сравнению с энергией, приобретаемой им в электрическом поле трубки, то для расчетов можно полагать:

$$\left(\frac{mv^2}{2}\right)_{\text{макс}} = eV; \quad \nu_{\text{макс}} = \frac{eV}{h}; \quad \lambda_{\text{мин}} = \frac{hc}{eV}. \tag{2.74}$$

Измерения с большой точностью подтверждают этот расчет. Заметим, что, пользуясь этой формулой, можно по измеренным значениям $\lambda_{\text{мин}}$, e и V вычислить постоянную Планка; она оказалась очень близкой к значениям, полученным другими способами.

Однако в потоке электронов, движущихся к мишени, вследствие тепловых столкновений должен существовать некоторый разброс по скоростям (энергиям). Большинство электронов будут иметь скорости, близкие к средней скорости потока $v_{\text{пот}}$; некоторые электроны будут иметь $v < v_{\text{пот}}$, другие $v > v_{\text{пот}}$ (но не больше максимальной скорости, определяемой приложенным напряжением: $v_{\text{макс}} = \sqrt{2\frac{eV}{m}}$). Поэтому наиболее вероятным будет появление фотонов с энергией

$$h\nu_{\text{в}} = mv_{\text{пот}}^2/2.$$

Таким образом, функция распределения фотонов по энергиям в тормозном и рентгеновском излучении должна иметь максимум при $\nu = \nu_{\text{в}}$. Так как средняя скорость потока электронов в трубке определяется приложенным напряжением V, то с увеличением V этот максимум должен перемещаться в сторону коротких волн (рис. IV.83).

Эмпирически установлено, что суммарная интенсивность рентгеновского излучения трубки пропорциональна силе I проходящего через трубку тока, квадрату приложенного напряжения V и порядковому номеру бомбардируемого элемента Z:

$$P = kIV^2Z.$$

Сопоставляя энергию, расходуемую трубкой, с энергией полученного рентгеновского излучения, оценивают коэффициент полезного действия трубки; оказалось, что он невелик (несколько процентов), зависит от вещества мишени и пропорционален приложенному напряжению.

Иногда (исходя из фотонной теории) полагают, что испускание электроном фотона при торможении есть процесс, обратный фотоэффекту: в одном из них фотон «рождается», в другом — «исчезает» (поглощается). В обоих процессах участвуют две элементарные частицы и происходит передача энергии от одного из них к другому (предполагается, что атомы вещества, внутри которого происходят эти процессы, создают лишь условия, необходимые для реализации этих процессов; тогда их участием в обмене энергиями можно пренебречь). Принципиальное значение этих процессов заключается в следующем. Если фотоны есть частицы поля, то в этих процессах происходит обмен энергией между веществом и полем; если же отнести фотоны к частицам вещества (находящимся в «особом» состоянии), то роль электромагнитного поля будет сведена к активному участию в этих процессах как среды, но не как объекта, обладающего энергией.

Если скорость электронов, возбуждающих рентгеновское излучение, очень велика, то на фоне непрерывного спектра тормозного излучения появляются отдельные линии.

Изучение этих рентгеновских спектральных линий показало, что их расположение в пределах непрерывного спектра зависит от вещества мишени, ввиду чего набор этих линий называют *характеристическим рентгеновским спектром*. Для выяснения происхождения такого спектра важное значение имело следующее обстоятельство: при бомбардировке мишени протонами или альфа-частицами непрерывное излучение не появляется; возникают только характеристические спектры. Если же мишень облучается коротковолновыми рентгеновскими лучами, то от нее исходит излучение, состоящее из смеси рассеянных первичных лучей и характеристического спектра вещества мишени. Эти факты показывают, что рентгеновский линейчатый спектр, по-видимому, связан с возбуждением атомов вещества и обратным переходом в нормальные состояния. Так как частоты этих линий в десятки тысяч раз превосходят частоты видимого света, то, следовательно, испускание рентгеновских фотонов должно сопровождаться большими изменениями в энергетическом состоянии атома.

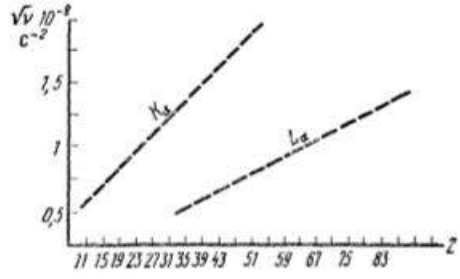


Рис. IV.84

Характеристические рентгеновские спектры различных элементов имеют однотипную структуру и состоят из нескольких серий, обозначаемых буквами K, L, M, N и т. д. Каждая из этих серий содержит некоторый набор отдельных спектральных линий, обозначаемых индексами (например, K_α, K_β, ... ; L_α, L_β, ...). Согласно закону, установленному Мозли, *частоты характеристического спектра элементов возрастают с увеличением порядкового номера*, т. е. при переходе к более тяжелым атомам весь спектр, сохраняя свою структуру, смещается в сторону коротких волн. Если выбрать одну определенную линию спектра, например самую коротковолновую K_α (или любую другую), то, согласно закону Мозли, частота этой линии определяется порядковым номером элемента Z:

$$\nu = A\,(Z-b)^2; \quad \sqrt{\nu} = aZ - c \tag{2.75}$$

(где A, a, b, c — постоянные величины), т. е. *квадратный корень из частот характеристического рентгеновского излучения есть линейная функция от порядкового номера элемента*.

На рис. IV.84 показаны прямые, на которых укладываются частоты линий K_α и L_α различных элементов.

Руководствуясь представлениями Н. Бора о происхождении оптических спектров, Мозли установил, что

$$A = R\left(\frac{1}{n^2} - \frac{1}{k^2}\right),$$

где R — постоянная Ридберга, n и k — целые числа. По измерениям Мозли константа b в формуле (2.75) оказалась в пределах каждой серии спектра одинаковой для всех элементов: для K-серии $b = 1$, для L-серии $b = 7{,}5$ и т. д. Таким образом, для характеристического рентгеновского спектра всех элементов можно написать одну формулу

$$\nu = R(Z-b)^2\left(\frac{1}{n^2} - \frac{1}{k^2}\right), \qquad (2.76)$$

где величина b выделяет ту или иную серию в этом спектре. Эта формула отличается от сериальной формулы атома водорода (см. § 9) тем, что содержит порядковый номер элемента и величину b, имеющую различное значение для каждой из серий спектра.

Излучение характеристического рентгеновского спектра обусловлено перемещениями электронов внутри атомов, переходами от дальних оболочек (энергетических уровней) к ближним. Допустим, что при помощи внешнего воздействия из слоя K (см. § 74) выбит один электрон, который затем покидает атом. Вследствие этого в структуре атома появляется энергетическая вакансия и атом оказывается в возбужденном состоянии. Для перехода в нормальное состояние потребуется заполнение этой вакансии; оно может осуществиться различным образом. Вакантное место может быть занято, например, электроном из соседнего (более удаленного от ядра) слоя L. Тогда появится вакансия в этом слое L и она также должна быть заполнена электроном из еще более удаленных слоев. Однако возможно, что первоначальная вакансия в K-слое будет заполнена электроном не из L-слоя, а из M-, N- слоев и т. д. Таким образом, если имеется некоторое множество возбужденных атомов, имеющих вакансии в K-слое, то возможны целые серии переходов: $L \to K$, $M \to K$ и т. д.; $M \to L$, $N \to L$ и т. д. Переходы электрона в вакантное место K-слоя из дальних слоев сопровождается излучением спектральных линий: K_α $(L \to K)$; K_β $(M \to K)$ и т. д. Аналогичное объяснение получают линии L_α $(M \to L)$; L_β $(N \to L)$ и т. д. Частоты испускаемых спектральных линий (или фотонов) характеристического рентгеновского спектра определяются правилом Бора:

$$h\nu = E_k - E_n,$$

где E_k и E_n — энергии атома в начальном и конечном состояниях электрона.

При расчете энергии возбужденных состояний сложных атомов необходимо (в отличие от атома водорода, содержащего только протон и электрон) учесть взаимодействие не только электронов с ядром, но и электронов между собой. Наличие электронов ослабляет поле положительных зарядов ядра, поэтому следует ожидать, что потенциальная энергия, которой обладает один электрон в поле заряда

ядра Ze в присутствии других электронов, будет меньше, чем в их отсутствии. Это уменьшение энергии можно оценить некоторой величиной, вычитаемой из Z, т. е. можно представить себе, что поле ядра соответствует не заряду Ze, а несколько меньшему заряду $(Z - b)e$. Этими рассуждениями выясняется происхождение величины b в формулах (2.75) и (2.76); ее называют *постоянной экранирования* (электроны в заполненных внутренних оболочках являются своеобразным «электрическим экраном» положительного заряда для тех электронов, которые находятся в более удаленных оболочках).

В тяжелых атомах энергия, которая выделяется при переходе электрона из удаленных орбит в K-слой, несколько больше энергии, необходимой для ионизации самого атома, т. е. для удаления одного из электронов из валентных орбит за пределы атома. Поэтому возможен и наблюдается процесс, при котором переход электрона в K-слой сопровождается не излучением рентгеновского фотона, а выбросом периферийного электрона за пределы атома (эффект Оже). Однако вероятность реализации такого процесса меньше вероятности излучения фотона.

ДИФРАКЦИЯ И ИНТЕРФЕРЕНЦИЯ РЕНТГЕНОВСКИХ ЛУЧЕЙ

Волновые свойства рентгеновских лучей наблюдаются в явлениях дифракции и интерференции. Так как длина волны рентгеновских лучей очень мала, то дифракционные решетки, употребляющиеся для видимого света, могут быть использованы для рентгеновского излучения только при больших углах падения, близких к 90° (см. § 5). Этот способ наклонного падения рентгеновских лучей на плоские отражающие решетки позволяет определить длину волны этих лучей.

Другой способ наблюдения дифракции и интерференции рентгеновских волн основан на предложении Лауэ (1913) — воспользоваться кристаллическими решетками, в которых атомы и молекулы, упорядоченно расположенные (с определенными интервалами) вдоль линий и на плоскостях, могли бы играть роль периодических неоднородностей. Схема опыта, подтвердившего возможность получения дифракционной картины от кристаллов, показана на рис. IV.85, *а*, где R — рентгеновская трубка; S — отверстие, через которое узкий пучок рентгеновского излучения проникает в свинцовый ящик P и падает на кристаллическую пластинку K, поставленную строго перпендикулярно лучам; F — фотографическая пластинка. На этой пластинке получилась картина, изображенная на рис. IV.85, *б*. Центральное темное пятно образуется неотклоненными лучами. Каждое маленькое пятнышко вокруг центрального есть результат интерференции (взаимного усиления) рентгеновских лучей, отклоненных атомами решетки; расположение этих пятен на экране оказалось различным у различных веществ.

При рассмотрении дифракции и интерференции рентгеновских лучей на кристаллах следует иметь в виду, что кристаллы представляют собой трехмерные дифракционные решетки. В обыкновенных

дифракционных решетках для видимого света отклоняющими объектами являются чередующиеся «штрихи» — прозрачные и непрозрачные узкие полоски, имеющие л и н е й н у ю протяженность. В кристаллической же решетке таких сплошных линий нет. Каждый атом или молекула кристалла должны рассматриваться как «непрозрачная»

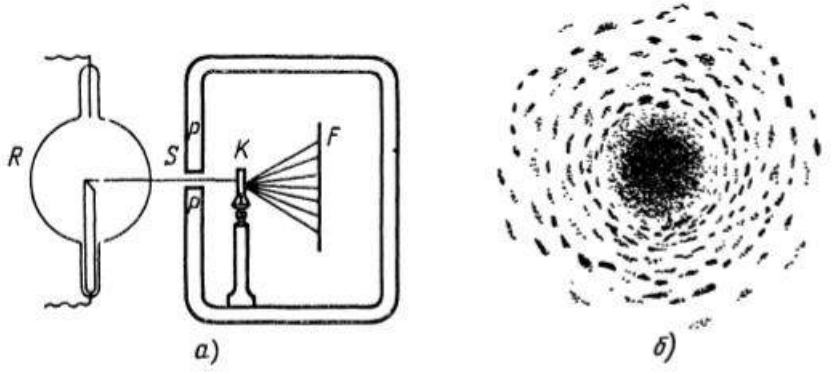


Рис. IV.85

для рентгеновских волн частица, причем они расположены в определенном порядке по объему кристалла на расстояниях, соизмеримых с длинами волн (для кристалла, например, поваренной соли расстояние между атомами натрия и хлора равно $\approx 2{,}8 \cdot 10^{-10}$ м). Отклонение рентгеновских волн от прямолинейного направления распространения может быть обусловлено не только отражением (рассеянием), но и *дифракционным огибанием* волной этих непрозрачных для нее тел. Вследствие этого в каждую точку экрана попадут волны, отклоненные от отдельных, пространственно расположенных частиц. Расчет показывает, что в этом случае в результате интерференции дифрагированных волн на экране должна получиться картина, состоящая из отдельных точек, в которых волны усиливают друг друга. Кроме того, из этих расчетов следует, что при данных периодах в пространственном расположении атомов максимумы могут быть получены только для волн определенной длины; другие волны будут равномерно рассеиваться.

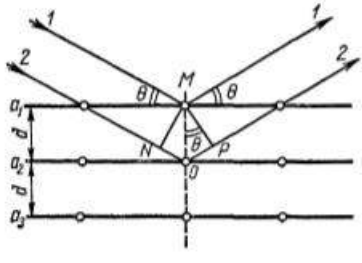


Рис. IV.86

Рассмотрим отражение плоской рентгеновской волны (т. е. пучка параллельных рентгеновских лучей) от ряда параллельных плоскостей кристаллической решетки (рис. IV.86). С одной стороны, должен соблюдаться закон отражения, поэтому «углы скольжения» θ (дополнительные к углам падения и отражения) должны быть равны. С другой стороны, для того чтобы лучи, отраженные от плоскостей a_1, a_2, a_3, ... ,

взаимно усиливались, необходимо определенное соотношение между фазами, а именно

$$2d \sin \theta = n\lambda, \tag{2.77}$$

где d — расстояние между отражающими слоями, λ — длина волны, $n = 1, 2, 3, \ldots$ Эта формула (В у л ь ф а—Б р э г г а) показывает, что данная кристаллическая решетка при выбранном угле падения будет «правильно» отражать только волны определенной длины; остальные волны, идущие на кристалл под тем же углом, будут рассеиваться по различным направлениям. Этой формулой можно воспользоваться также и для определения расстояния между плоскостями отражающего кристалла, если известна длина волны падающего рентгеновского излучения и измерены углы θ скольжения падающих и отраженных лучей.

По отношению к рентгеновскому излучению определены также показатели преломления некоторых веществ (они оказались близкими к единице), наблюдалось полное внутреннее отражение, обнаружены дисперсия (зависимость показателя преломления вещества от длины рентгеновских волн) и другие волновые свойства.

Измерение *интенсивности* рентгеновских лучей основано на ионизации вещества, вызываемой этими лучами. Степень ионизации измеряется величиной ионизационного (несамостоятельного) тока. Иногда об интенсивности рентгеновского излучения судят по степени потемнения фотопластинок под действием этого излучения.

Измерение интенсивности в абсолютных единицах (например, в эрг·см$^{-2}$·с$^{-1}$) представляет значительные трудности. Однако относительную интенсивность найти значительно легче. Для этого достаточно сравнить величины ионизационного тока (или степень потемнения фотопластинок) для двух пучков рентгеновских лучей. Отношение величин ионизационных токов равно отношению интенсивностей пучков лучей.

Как и для видимого излучения (см. § 3, формула (1.13)), закон ослабления интенсивности монохроматического пучка рентгеновских лучей имеет вид:

$$\begin{aligned} -\mathrm{d}I &= \mu I \,\mathrm{d}x, \\ I &= I_0 \mathrm{e}^{-\mu x}. \end{aligned} \tag{2.78}$$

Линейный коэффициент ослабления μ зависит от длины волны рентгеновских лучей и природы вещества, в котором происходит поглощение, и на основании формулы (2.78) выражает относительное ослабление рентгеновского излучения на единице длины слоя поглотителя; μ не зависит от физических и химических условий, в которых находятся атомы поглощающего вещества. Например, свинец, обладая большим коэффициентом μ, входя в состав свинцовых белил или стекла, делает их значительно ослабляющими рентгеновские лучи. Вместо коэффициента μ часто употребляют величину $d_{1/2}$, равную толщине слоя, ослабляющего интенсивность пучка в два раза (его называют слоем половинного ослабления). Связь между μ и $d_{1/2}$ легко устанав-

ливается из уравнения (2.78); полагая $I = 1/2I_0$, а $x = d_{1/2}$, получим

$$d_{1/2} = \frac{\ln 2}{\mu} = \frac{0{,}69}{\mu}.$$

ДОЗИМЕТРИЯ

В ряде случаев, например для медицинских целей, бывает важно знать, какое количество рентгеновского излучения проникает в какое-либо вещество. Область физики, занимающаяся методами дозировки излучения, называется *дозиметрией*.

Основными величинами дозиметрии являются:

доза D излучения, определяемая величиной энергии излучения, поглощенной в единице массы вещества,

$$D = \Delta W/\Delta m, \tag{2.79}$$

где ΔW — энергия рентгеновского излучения, поглощенная массой Δm вещества;

экспозиционная доза, — характеризующая излучение по произведенной им ионизации.

Единицей экспозиционной дозы излучения является рентген (Р): рентген — *это такая экспозиционная доза, при которой в* 0,001293 г *воздуха* (1 см^3 при нормальных условиях) *образуются ионы, несущие одну электростатическую единицу количества электричества каждого знака.*

Для определения энергетического эквивалента рентгена следует подсчитать, сколько пар ионов должно быть образовано в 1 см^3 воздуха, чтобы заряд ионов одного знака составил одну единицу СГС количества электричества. Так как заряд одновалентного иона равен $4{,}8 \cdot 10^{-10}$ СГС, то для образования единицы СГС количества электричества надо иметь

$$\frac{1}{4{,}8 \cdot 10^{-10}} = 2{,}08 \cdot 10^9$$

пар ионов в 1 см^3 воздуха.

Измерения показывают, что энергия образования одной пары ионов в воздухе при нормальных условиях равна 33 эВ. Учитывая это, получим $1\text{Р} = 33 \cdot 2{,}08 \cdot 10^9$ эВ/см^3 (воздуха) или $1\text{Р} = 0{,}11$ эрг/см^3 (воздуха). Следовательно, рентген можно определить как дозу излучения, при которой в 1 см^3 воздуха при нормальных условиях образуется $2{,}08 \cdot 10^9$ пар ионов, или как дозу излучения, которой соответствует поглощенная энергия в 0,11 эрг в 1 см^3 воздуха.

Мощностью дозы излучения называется величина, равная отношению дозы излучения ΔD ко времени излучения Δt:

$$P = \Delta D/\Delta t. \tag{2.80}$$

Для ионизирующего излучения единицами мощности дозы служат: Р/с; Р/мин; Р/ч и др.

Приборы, служащие для практических измерений дозы рентгеновских лучей, называются *дозиметрами* или *рентгенометрами.*

Рентгеновские лучи находят широкое применение в науке и технике. При их помощи изучается строение аморфных тел, жидкостей, кристаллов, атомов и молекул (рентгеноструктурный анализ). Рентгеновское просвечивание получило большое распространение как в медицине для обнаружения переломов костей, нахождения инородных тел (иголок, пуль и т. п.), обнаружения заболевания тканей, так и в технике для просвечивания металлических и других изделий. Последнее особенно важно для контроля качества изделий и сварных швов; при этом выявляется наличие раковин, посторонних включений, внутренних разрывов и трещин (при обработке давлением). Большое распространение получил также рентгеновский спектральный анализ, который в ряде случаев оказывается более простым и удобным способом определения химического состава тела, чем обычный химический анализ. Рентгеновский спектральный анализ применяется: 1) для химического анализа тонких слоев вещества (несколько микрон), 2) для выявления элементов, содержащихся в очень малых количествах, 3) для экспресс-анализа и т. д.

§ 17. ЛЮМИНЕСЦЕНТНОЕ ИЗЛУЧЕНИЕ

Некоторые вещества при их облучении (например, видимым, ультрафиолетовым или рентгеновским излучением) испускают собственное свечение, спектр которого отличается от спектра поглощенного излучения и определяется химическим составом и внутренней молекулярной структурой этих веществ. Такого рода свечение называется *люминесцентным излучением* или просто *люминесценцией*, а испускающие их вещества — *люминофорами*.

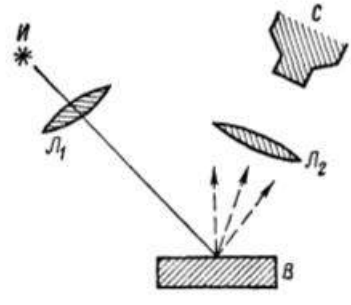


Рис. IV.87

На рис. IV.87 показана схема установки для изучения люминесценции различных веществ: *И* — источник возбуждающего излучения (для детальных исследований он должен давать монохроматические участки спектра различных длин волн); $Л_1$ — линза, фокусирующая это излучение на исследуемое вещество *В*; $Л_2$ — линза, собирающая люминесцентное излучение на спектральный аппарат *С*, с помощью которого определяются состав излучения и распределение энергии по спектру.

Люминофорами являются: 1) пары и газы некоторых элементов (натрий, сера, иод, кислород); 2) соли некоторых веществ и их растворы (большинство неорганических солей не люминесцирует; исключение составляют соединения лантаноидов: цезия, празеодима, неодима и др.); 3) ряд органических веществ (бензол и его производные, нафталин, антрацен и др.; красители: флуоресцеин, родамины и др.); 4) кристаллические неорганические вещества, содержащие в своей структуре (в небольшом количестве) примеси ионов тяжелых металлов (например,

сернистый цинк, в кристаллическую структуру которого введены ионы меди или марганца). Металлы ни в твердом, ни в жидком состоянии не люминесцируют. Некоторые вещества обнаруживают люминесцентные свойства только при определенных условиях (низкие температуры, выбор специального растворителя, необходимая концентрация в растворе и т. п.).

Возбуждение люминесценции может быть произведено различными способами, в связи с чем употребляются следующие названия:

1) *фотолюминесценция,* вызванная облучением, содержащим оптический участок спектра — видимый и ультрафиолетовый (свечение некоторых красок, керосина, флуоресцеина и других веществ под действием света или ультрафиолетовых лучей);

2) *рентгенолюминесценция,* вызванная рентгеновскими лучами;

3) *катодолюминесценция,* вызванная бомбардировкой вещества потоком быстрых электронов (свечение экрана телевизора, осциллографа, свечение минералов, облучаемых потоком электронов);

4) *электролюминесценция,* вызванная электрическими полями, которые освобождают слабо связанные в структуре кристалла электроны и сообщают им энергию, достаточную для возбуждения вещества;

5) *хемилюминесценция,* возникающая при химических реакциях; возбуждение вещества происходит под действием выделяющейся при реакции энергии (свечение фосфора, гниющего дерева, светлячков, моллюсков, вызываемое окислением, и др.).

ПРИРОДА ЛЮМИНЕСЦЕНЦИИ

Люминесцентное излучение отличается от других видов излучения следующими характерными особенностями:

1. Каждое вещество (люминофор) имеет определенный спектр люминесценции (по составу и распределению энергии), не зависящий от спектра поглощенного излучения, вызвавшего люминесценцию. Если возбуждение вещества и последующая люминесценция вызваны бомбардировкой быстрыми частицами (электронами, альфа-частицами), то в исходном веществе возможны химические и структурные изменения и поэтому спектр люминесценции будет определяться не только исходным веществом, но и составом и структурой образующихся новых веществ (впрочем, это обстоятельство может иметь место и при облучении ультрафиолетовыми, рентгеновскими и гамма-лучами). Спектр люминесцирующего вещества не зависит от способа возбуждения.

2. Люминесцентное излучение при одной и той же температуре имеет бо́льшую спектральную плотность энергетической светимости по сравнению с тепловым излучением (сравнение обычно производится с излучением абсолютно черного тела). Тепловое излучение всегда добавляется к люминесцентному, но имеет значительно меньшую мощность. Многие люминофоры излучают видимый свет при нормальных и низких температурах, тогда как при таких температурах в тепловом излучении абсолютно черного тела и самих люминофоров эти участки спектра имеют исчезающе малую мощность. Ввиду этого одно

только наличие видимых участков спектра при невысоких температурах («холодный свет») достаточно для того, чтобы отнести это излучение к люминесценции.

3. В спектре люминесценции имеется коротковолновая граница $\lambda_{гр}$, различная для различных веществ. Фотолюминесценцию можно вызвать только при условии, если в падающем (и поглощенном этим веществом) излучении содержатся длины волн, меньшие $\lambda_{пред}$, т. е. *поглощенное излучение должно быть более коротковолновым, чем люминесцентное*. Это правило Стокса имеет некоторые отступления, рассмотренные ниже.

4. Люминесцирующие тела не подчиняются закону Кирхгофа: отношение спектральных плотностей энергетической светимости люминесцентного излучения r_λ к соответствующим коэффициентам поглощения α_λ не одинаково для всех люминесцирующих веществ. Кроме того, согласно указанному выше правилу Стокса, люминесцентное излучение всегда более длинноволновое, чем излучение, возбуждающее люминесценцию, поэтому излучение, исходящее от данного люминофора, не может быть использовано для возбуждения таких же люминофоров. Ввиду этого люминесцентное излучение называют неравновесным.

5. Люминесцентное свечение продолжается некоторое время после прекращения облучения. Это время у различных веществ не одинаково и имеет значения от 10^{-10} с (например, у паров и газов) до многих часов (у некоторых твердых люминофоров).

Различная длительность люминесценции после прекращения внешнего возбуждающего действия послужила основой для разделения люминесценции на:

а) *флуоресценцию,* исчезающую одновременно с прекращением облучения (таким свойством обладает, например, минерал флуорит);

б) *фосфоресценцию,* продолжающуюся некоторое время после устранения возбуждающего воздействия.

Однако исследования показали, что длительность люминесценции можно изменять в широких пределах; увеличивать у флуоресцирующих веществ и уменьшать у фосфоресцирующих веществ. Ввиду этого деление люминесценции по продолжительности свечения потеряло однозначный смысл. В настоящее время флуоресценцией называют спонтанное (самостоятельное) излучение возбужденных молекул при переходе их в нормальное состояние под действием внутренних электрических и магнитных полей; фосфоресценцией же называют излучение, которое возникает при рекомбинации частиц, разделенных при облучении (например, электрона и иона или двух частей диссоциированной молекулы).

6. Способность данного вещества к люминесценции может быть сильно ослаблена (и даже сведена к нулю) добавлением специальных примесей — «тушителей люминесценции» (анилин, гидрохинон, иодистый калий, ионы железа и др.).

Заметим, что тепловое излучение не может быть ослаблено или устранено какими-либо примесями, так как оно не зависит от вещества, а определяется только температурой. Поэтому зависи-

мость интенсивности излучения от наличия примесей является весьма характерной отличительной особенностью люминесценции.

7. Люминесцентное излучение, как и тепловое, является некогерентным, что свидетельствует об отсутствии связи («согласованности») между частицами — источниками этого излучения.

8. Энергия, излучаемая люминесцирующим телом, не превышает энергии, поглощенной при облучении.

Перечисленные выше особенности люминесцентного излучения получают удовлетворительное объяснение только на основе квантовых представлений. Существование продолжительного свечения после прекращения внешнего воздействия свидетельствует о том, что частицы вещества, поглотившие излучение, переходят не на обычные уровни энергии (с которых возможно спонтанное излучение), а на особые *метастабильные* уровни, среднее время пребывания на которых велико. Свободные атомы таких уровней не имеют, поэтому среднее время пребывания электрона на каком-нибудь возбужденном уровне оказывается порядка 10^{-10} с. Метастабильные уровни образуются только в сложных системах: молекулах, ячейках кристаллических решеток и т. п.

Появление метастабильных уровней в люминофорах (а также и в других веществах) объясняется следующим образом. В § 13 было указано, что в сложных молекулах газов и жидкостей, а также в кристаллической решетке твердых тел энергия одного атома уже не определяется только взаимодействием между его составными частями — ядром и электронами, а зависит также от взаимодействия с другими соседними атомами. Это взаимодействие вносит изменения в значения энергии уровней. Определенный уровень *свободного* атома превращается в сложный уровень *связанного* атома, состоящий из очень большого числа близко расположенных подуровней. Можно также допустить, что каждый уровень энергии имеет уже не «естественную ширину», определяемую соотношением неопределенностей Гейзенберга, а значительно большую ширину ΔE, соизмеримую с разностью энергий соседних уровней.

В кристаллических решетках ввиду малых расстояний между соседними частицами (атомами, ионами, молекулами) и симметричного их расположения уровни со значениями энергии E, $E + \Delta E$ уже не локализуются возле частиц, а соединяются в «энергетические зоны», простирающиеся по объему кристалла. Электрон, «попавший в эту зону» (т. е. имеющий соответствующую энергию), может без затраты своей энергии свободно перемещаться по объему кристалла, разумеется, не выходя за пределы этой зоны. Таких зон может быть несколько; одни из них (соответствующие малым энергиям) будут полностью заполнены электронами (с соблюдением принципа Паули), другие могут оказаться незаполненными. Существование электронов вне этих зон запрещено. Если зона заполнена, то перемещения электронов в пределах этой зоны возможны только путем «обмена местами» между двумя соседними электронами; такие перемещения не вызывают заметных изменений в состоянии этой системы. Если же в объеме кристалла имеется незаполненная зона, то перемещения электронов в пре-

делах объема кристалла возможны и могут быть вызваны небольшими внешними полями, что означает способность вещества проводить электрический ток. Поэтому незаполненные энергетические зоны называются *зонами проводимости*.

Однако наличие таких (незаполненных) зон и возможность возбуждения вещества путем переброски в эти зоны электронов из низших (заполненных) зон еще не означает, что это вещество может быть люминофором. Установлено, что кристаллы с правильной внутренней структурой не обладают люминесцентными свойствами. Такие свойства обнаруживаются у кристаллов с дефектами; введение весьма малого количества примесей (иногда сотые доли процента) деформирует решетку кристалла и создает в его узлах «активные центры», являющиеся источниками люминесцентного излучения. Такими примесями (активаторами) являются, в частности, ионы тяжелых металлов — серебра, меди, висмута, марганца и др. Введенные в структуру кристалла какого-нибудь вещества (например, сернистого цинка), ионы тяжелых металлов создают вокруг себя дополнительные уровни энергии, лежащие несколько ниже зоны проводимости. Так как расстояние между ионами сравнительно велико, то эти уровни оказываются *локализованными* в ближайшей окрестности ионов и не соединяются в зоны, простирающиеся по объему кристалла. Ввиду этого электроны, попавшие в эти локальные уровни (если эти уровни оказались свободными), уже не имеют возможности свободно перемещаться по объему вещества и оказываются «связанными» в данном месте. Таким образом, эти дополнительные уровни энергии вблизи примесных ионов оказываются потенциальными ямами для электронов, в которых они будут вынуждены пребывать некоторое время. Освобождение электронов из этих метастабильных уровней может быть произведено либо под воздействием тепловых столкновений, либо при внешнем воздействии на вещество. В первом случае время пребывания электрона на метастабильном уровне будет зависеть от «глубины ямы» (т. е. от разности энергии между данным метастабильным уровнем и зоной проводимости) и от температуры.

Для объяснения всех особенностей люминесцентного излучения должны быть известны не только значения энергий уровней и зон, но и вероятности переходов из верхних уровней на нижние. В квантовой физике установлено, что вероятности различных процессов определяются не только значениями энергии начального и конечного состояний, но и значениями квантовых чисел, определяющих эти состояния (правила отбора, см. § 14). Некоторые процессы имеют столь малую вероятность реализации, что их можно полагать «запрещенными». Поэтому, рассматривая систему уровней и зон в люминесцирующем веществе, необходимо учесть, какие переходы являются «разрешенными» или «запрещенными».

Люминесцентное излучение в зависимости от характера вызывающих его процессов делят, по предложению С. И. Вавилова, на следующие виды:

1) *спонтанное* (самостоятельное) излучение; оно происходит при переходе почти свободной частицы (атома, молекулы) из возбужден-

ного состояния в нормальное с излучением фотона определенной энергии:

$$h\nu = E_{\text{возб}} - E_{\text{норм}}.$$

Характерная особенность такого свечения: вероятность перехода определяется внутренними полями, но не внешним воздействием на возбужденную частицу. На рис. IV.88, *а* стрелка *1* указывает на процесс возбуждения, стрелка *2* — на разрешенный переход из возбужденного состояния в нормальное. Продолжительность люминесценции, т. е. среднее время пребывания частицы в возбужденном состоянии, составляет 10^{-9}—10^{-10} с;

2) *вынужденное* излучение, требующее наличия внешнего воздействия на излучающий электрон. Для такого излучения характерно наличие третьего (метастабильного) уровня $E_{\text{мет}}$, расположенного между $E_{\text{норм}}$ и $E_{\text{возб}}$ (рис. IV.88, *б*). При возбуждении люминофора электрон оказывается на уровне $E_{\text{мет}}$, однако переход с этого уровня на нормальный запрещен. Люминесцентное излучение возможно только после переброски электрона с метастабильного уровня $E_{\text{мет}}$ на верхний уровень $E_{\text{возб}}$, откуда переход на нормальный уровень разрешен. Длительность люминесценции (после прекращения возбуждающего действия на люминофор) зависит от времени, которое затрачивается на переброску электрона на верхний уровень, и колеблется в широких пределах — от 10^{-5} с до десятков секунд;

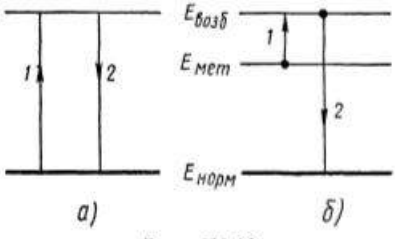


Рис. IV.88

3) *рекомбинационное* излучение, возникающее следующим образом: под действием облучения происходит разъединение связанных частиц (отделение электрона от иона, диссоциация молекул на атомы и т. п.). При обратном воссоединении (рекомбинации) израсходованная энергия выделяется и вызывает возбуждение люминесцирующей частицы вещества. Последующее излучение происходит либо по схеме спонтанного, либо по схеме вынужденного излучения. Длительность свечения колеблется от 10^{-7} с до минут и часов. Наличие очень большой длительности люминесценции, не свойственной спонтанному и вынужденному излучению, является достаточным признаком наличия рекомбинационных процессов.

В некоторых люминофорах процессы поглощения и излучения происходят (локализованы) в одном и том же месте их структуры, т. е. одна и та же частица и поглощает поступивший извне фотон и испускает люминесцентный фотон; такими частицами могут быть отдельные атомы, ионы, молекулы или даже сложные комплексы, находящиеся в определенных местах структуры люминофора. По предложению В. Л. Лёвшина, такой вид люминесценции называется *свечением дискретных центров*; он характерен для многих жидких люминофоров. В других веществах элементарные процессы возбуждения и излучения

происходят, как правило, в различных местах структуры люминофора. Например, у кристаллов с деформированной решеткой, в которую вкраплены ионы тяжелого металла, падающее излучение поглощается основным веществом, а излучающими центрами являются ионы примеси. В таких люминофорах необходим перенос энергии от мест поглощения к местам излучения, что осуществляется главным образом электронами. Такой вид люминесценции также называется рекомбинационным.

На рис. IV.89 показаны схемы электронных переходов в кристаллических люминофорах при отсутствии (*а*) и при наличии (*б*) локальных

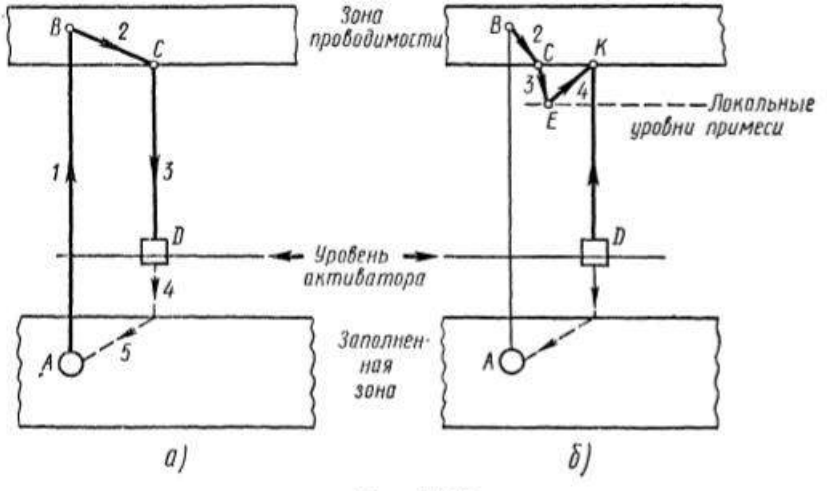


Рис. IV.89

уровней, появившихся вследствие деформации решетки ионами примесей. Электрон, получивший энергию возбуждения в точке A, перебрасывается в точку B зоны проводимости. Постепенно теряя энергию вследствие тепловых столкновений, электрон приближается к нижнему краю зоны проводимости (точка C), откуда он имеет разрешенный переход к уровню активатора. Однако для этого перехода на уровне активатора должна иметься вакансия; она образуется вследствие перемещения связанных электронов (по пунктирным линиям $D \to 4 \to 5 \to A$) в «дырку», образовавшуюся в заполненной зоне после ухода электрона из A (иногда предпочитают отмечать перемещение из A в D самой «дырки», рассматриваемой как условная положительно заряженная частица). Таким образом, люминесцентный фотон испускается при переходе $C \to D$. Время, прошедшее от момента получения энергии возбуждения до момента излучения, оказывается порядка 10^{-6} с.

При наличии локальных уровней, образованных деформацией решетки и расположенных ниже зоны проводимости, существует вероятность безызлучательного перехода электронов на эти уровни (из точки C в точку E). Однако переход $E \to D$ запрещен, поэтому электрон будет

вынужден находиться на локальном (метастабильном) уровне до тех пор, пока какое-нибудь внешнее воздействие не вернет его в зону проводимости, откуда он может использовать разрешенный канал перехода $K \to D$ и рекомбинировать с «дыркой» на уровне активатора. В зависимости от глубины потенциальной ямы (разности энергии между нижним краем зоны проводимости и локальным уровнем) время пребывания электрона на локальном уровне может колебаться в широких пределах и это отразится на длительности люминесценции после прекращения возбуждения.

Интересным примером является *сенсибилизированная люминесценция* в смеси двух газов, где молекулы одного газа поглощают излучение, но не люминесцируют, так как уровень возбуждения является метастабильным; но молекулы второго газа, не поглощающие падающего излучения, возбуждаются при столкновениях с молекулами первого газа и люминесцируют, имея разрешенный переход в нормальное состояние.

ХАРАКТЕРИСТИКИ ЛЮМИНЕСЦИРУЮЩИХ ВЕЩЕСТВ

Люминесцентные вещества характеризуются:

1) спектром излучения (составом спектра и функцией распределения мощности излучения по длинам волн);

2) спектром поглощения (зависимостью спектрального коэффициента поглощения от длины волны);

3) выходом люминесценции (отношением энергии, затраченной на возбуждение, к энергии люминесцентного излучения);

4) кинетикой процесса люминесценции (нарастание и затухание люминесценции в зависимости от внешнего воздействия).

Спектр люминесценции определяется: 1) составом и внутренней структурой вещества и 2) воздействием внешней среды. У одноатомных газов, например у паров ртути, натрия, спектр излучения линейчатый; у паров бензина — полосатый; твердые и жидкие вещества имеют спектр, состоящий из более или менее широких непрерывных полос. Структура вещества определяет энергетический спектр, расположение уровней и зон, их ширину, вероятности переходов. Внешняя среда может произвести некоторые изменения в структуре вещества и энергетических уровней, в значения вероятностей различных переходов (уменьшить вероятность одних, увеличить вероятность других процессов, в частности устранить запреты); тепловое движение в среде может способствовать попаданию электрона в потенциальные ямы или извлечению их из метастабильных уровней и т. д.

Спектр поглощения вместе с спектром излучения является важной характеристикой состава и строения люминофоров. Коэффициент поглощения определяется следующим образом: допустим, что на толщине dl слоя люминофора поток монохроматического излучения $\Phi_e(\lambda)$ ослабевает на $d\Phi_e(\lambda)$. Полагая, что поглощение прямо пропорционально $\Phi_e(\lambda)$ и dl, получим:

$$d\Phi_e(\lambda) = -k_\lambda \Phi_e(\lambda)\, dl; \quad \Phi_e(\lambda) = \Phi_{e0}(\lambda)\, e^{-k_\lambda l}; \quad k_\lambda = \frac{1}{l} \ln \frac{\Phi_{e0}(\lambda)}{\Phi_e(\lambda)},$$

где l — толщина слоя, в котором мощность проходящего через люминофор внешнего излучения ослабевает от $\Phi_{e0}(\lambda)$ до $\Phi_e(\lambda)$, k_λ — спектральный коэффициент поглощения для излучения с длиной волны λ.

Однако не всякое излучение, поглощенное данным люминофором, вызывает люминесценцию. Поэтому из всего спектра поглощения выделяются те его участки, которые могут вызвать необходимое возбуждение люминофора; эти участки называются *спектром возбуждения*.

Связь между спектрами излучения и поглощения при фотолюминесценции выражается несколькими правилами:

1) правило Стокса: *излучение, вызывающее люминесценцию, должно иметь, как правило, меньшую длину волны, чем коротковолновая граница спектра люминесценции.*

Это означает, что при поглощении фотонов малой частоты нельзя получить фотоны больших частот. Если вещество поглощает фотоны с энергией $h\nu_0$, то при наличии различных потерь в веществе люминофора энергия излученного фотона $h\nu$ будет меньше $h\nu_0$ и в идеальном случае равна ей. Однако возможно, что молекулы, поглощающие фотон, уже находятся в возбужденных состояниях, имея небольшой избыток энергии ΔE по сравнению с нормальным состоянием. В этом случае излучаемый фотон может унести энергию $h\nu = h\nu_0 + \Delta E$. Такое объяснение должно быть дано многочисленным отступлениям от правила Стокса. Область излучения с частотами ν, бóльшими, чем частота ν_0 возбуждающего фотона, называется *антистоксовой*.

Более общим является другая закономерность:

2) правило Стокса — Ломмеля: *спектр люминесцентного излучения в целом и его максимум находятся в области более длинных волн (меньших частот) по отношению к спектру поглощения и его максимуму.*

Это правило относится уже не к отдельному акту взаимодействия фотона и молекулы, а к совокупности всех процессов поглощения и излучения в веществе, т. е. имеет статистический (вероятностный) смысл: более вероятными являются акты, при которых фотоны больших энергий вызывают появление фотонов меньших энергий и менее вероятными — присоединение избыточных энергий молекул ΔE к энергии $h\nu_0$ падающего фотона и выход фотона с большей энергией.

В. Л. Лёвшиным было установлено еще одно правило, относящееся к большой группе органических люминофоров:

3) *функции распределения энергии в спектрах поглощения и излучения оказываются зеркально симметричными.*

На рис. IV.90 показаны спектры поглощения и излучения одного из органических красителей (родамина 6*G* в ацетоне), иллюстрирующие все перечисленные выше правила. По оси абсцисс отложены частоты колебаний, по оси ординат — квантовая интенсивность $I_{кв} = \psi(\nu)/h\nu$, где $\psi(\nu)$ — функция распределения энергии по спектру. Кривая люминесцентного излучения показана сплошной линией, кривая поглощения — пунктирной. Оба спектра оказываются симметричными относительно частоты ν_0, соответствующей точке пересечения кривых. На этом рисунке показана также антистоксовская область в том частном

случае, когда на люминофор действует внешнее излучение частоты ν_0: заштрихованная часть спектра излучения имеет $\nu > \nu_0$.

Выход люминесценции. Люминесценцию можно рассматривать как процесс преобразования энергии, подаваемой для возбуждения, в излучение определенного спектрального состава, характерного для данного люминофора. Этот процесс может иметь некоторый коэффициент полезного действия, называемый в данном случае *выходом*. Различают *энергетический выход* — соотношение энергии, поглощенной люминофором при возбуждении $W_{погл}$, и полной энергии излучения (до его прекращения) $W_{люм}$:

$$B_э = W_{люм}/W_{погл}, \qquad (2.81)$$

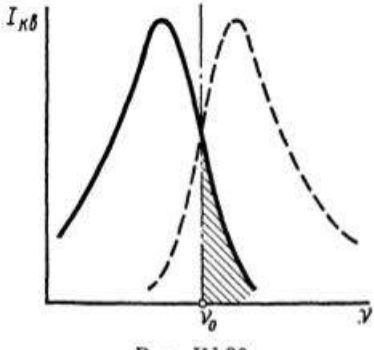


Рис. IV.90

и *квантовый выход* — отношение числа фотонов люминесцентного излучения $N_{люм}$ к числу поглощенных фотонов $N_{погл}$ (если люминесценция вызвана облучением):

$$B_{кв} = N_{люм}/N_{погл}. \qquad (2.82)$$

Отличие величин $B_э$ и $B_{кв}$ от единицы обусловлено тем, что часть энергии, поглощенной люминофором, превращается в теплоту (или в другие виды энергии). Заметим, что даже в том идеальном случае, когда каждый поглощенный фотон порождает один люминесцентный фотон (т. е. когда квантовый выход равен единице), энергетический выход будет меньше единицы, так как, согласно правилу Стокса, возбуждающие фотоны должны иметь меньшую длину волны (т. е. большую энергию), чем люминесцентные; обусловленное этим уменьшение энергетического выхода называется *стоксовскими потерями*.

Измерения показали, что в некоторых случаях величина выхода люминесценции может быть весьма большой; например, для флуоресцеина он равен 0,76. Однако в большинстве случаев выход люминесценции оказывается значительно меньше единицы. Одно и то же вещество в разных условиях может дать разные выходы люминесценции; в частности, при некоторых условиях (например, при нагревании) почти вся энергия возбуждения может растратиться на процессы тушения, т. е. на процессы, не вызывающие люминесценций.

С. И. Вавилов установил следующий закон:

величина энергетического выхода люминесценции сначала растет пропорционально длине волны возбуждающего света, а затем быстро падает до нуля.

На рис. IV.91 представлена полученная Вавиловым зависимость энергетического выхода люминесценции $B_э$ от длины волны возбуждающего излучения для одной из солей флуоресцеина, подтверждающая сформулированный выше закон. Из этого рисунка видно, что вначале

от $\lambda \approx 254$ до $\lambda \approx 500$ нм выход люминесценции нарастает пропорционально увеличению длины волны. Затем в области более длинных волн происходит быстрое падение величины $B_э$ до нуля.

Пропорциональное увеличение энергетического выхода с увеличением длины волны возбуждающего света означает, что в этом интервале длин волн квантовый выход люминесценции остается постоянным (не зависит от длины волны; рис. IV.91). Это утверждение является иной формулировкой закона С. И. Вавилова, который является следствием квантового характера фотолюминесценции.

Действительно, допустим, что каждый фотон возбуждающего света создает фотон люминесценции, т. е. квантовый выход равен единице: $B_{кв} = 1$. В этом случае энергетический выход, очевидно, равен отношению энергии фотона люминесценции $h\nu_{люм}$ к энергии возбуждающего фотона $h\nu_{возб}$: $B_э = \frac{h\nu_{люм}}{h\nu_{возб}}$, или $B_э = \frac{h\nu_{люм}}{hc}\lambda_{возб}$, т. е. энергетический выход пропорционален длине волны возбуждающего света. Однако как только энергия падающего фотона окажется недостаточной для возбуждения атома, люминесценция прекращается; квантовый и энергетический выходы становятся равными нулю, о чем свидетельствует быстрое падение кривых на рис. IV.91.

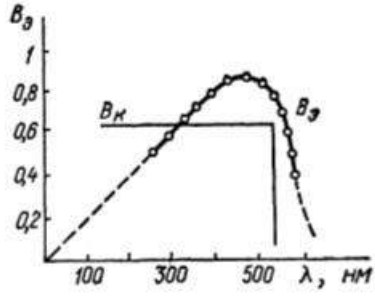


Рис. IV.91

Кинетика процесса люминесценции зависит от характера этого процесса. Рассмотрим простейший — спонтанный — процесс и допустим, что переход электрона с возбужденного уровня на нормальный имеет вероятность w и поэтому за время dt из общего числа n электронов такие переходы совершат dn электронов. Имеем:

$$dn = -wn\,dt; \quad n = n_0 e^{-wt},$$

где n_0 — число электронов на возбужденном уровне в начальный момент. Если каждый переход сопровождается излучением фотона (т. е. если нет потерь энергии возбуждения), то интенсивность люминесцентного свечения будет пропорциональна dn/dt и поэтому

$$I = I_0 e^{-wt}.$$

При таком характере высвечивания *длительностью люминесценции* τ называют время, в течение которого интенсивность уменьшается в e раз, т. е. $w\tau = 1$, $\tau = 1/w$.

Для рекомбинационного излучения dn должно быть пропорционально как числу свободных электронов, так и числу ионов, образовавшихся при ионизации (или числу каждой компоненты при диссоциации), т. е. должно быть пропорционально n^2. Если p — вероятность рекомбинации, то

$$dn = -pn^2\,dt; \quad n = \frac{n_0}{pn_0 t + 1}; \quad I = I_0 \frac{1}{(at+1)^2},$$

где $a = pn_0$. Однако этот закон изменения интенсивности люминесценции (с течением времени) не соблюдается, так как начало рекомбинации несколько отстает от начала возбуждения; в частности, некоторые электроны, находящиеся в зоне проводимости, могут попасть в потенциальные ямы (на метастабильные уровни) и продолжительное время не будут участвовать в рекомбинационных процессах.

ГАШЕНИЕ ЛЮМИНЕСЦЕНЦИИ

В структуре люминофоров могут быть созданы такие условия (или вызваны такие процессы), при которых энергия возбуждения переходит не в энергию люминесцентного излучения, а, например, в тепловую энергию. Осуществление безызлучательных переходов из возбужденных состояний в нормальные называется *тушением* (или *гашением*) люминесценции. Тушение может быть вызвано: повышением температуры; посторонними примесями (соли галоидных кислот, ионы серебра, железа и др.); помещением в среду, обладающую гасящим свойством (газообразный кислород, электролиты); переводом в другое агрегатное состояние (из твердого состояния в парообразное или в раствор); изменение концентрации раствора и др. По предложению В. Л. Лёвшина, различают:

1) *внешнее тушение,* при котором энергия возбужденных молекул, способных к люминесценции, передается молекулам, не обладающим этим свойством. Такой способ тушения не сопровождается изменениями в свойствах люминесцирующего вещества, их спектров поглощения и излучения;

2) *внутреннее тушение,* обусловленное процессами, происходящими внутри люминесцирующей молекулы, например превращение поглощенной энергии возбуждения сначала в энергию колебаний частей молекулы, а затем в тепловую энергию; переход молекулы в другую модификацию с потерей люминесцентной способности и т. п.

По классификации С. И. Вавилова, тушением *первого рода* называется уменьшение выхода фотолюминесценции в результате воздействия на невозбужденные молекулы, а *второго рода* — на возбужденные молекулы (в частности, внешнее тушение относится к второму роду).

Если гасящее действие обусловлено взаимодействием двух молекул, то оно будет сильно зависеть от времени существования возбужденного состояния одной из них; чем больше это время, тем больше вероятность реализации гасящего взаимодействия.

При длительном возбуждении люминофоров возможно установление *равновесия* между излучением и поглощением, при котором за некоторый промежуток времени Δt вещество будет получать извне такую же энергию, какую расходует на люминесцентное излучение и на безызлучательные потери. Возможно также и явление «насыщения» люминесценции: при очень мощном внешнем воздействии все люминесцирующие центры могут оказаться возбужденными и поэтому дальнейшее увеличение интенсивности внешнего воздействия не сможет вызвать увеличения яркости люминесцентного свечения.

Перечислим важнейшие применения люминесценций:

1. *Люминесцентная лампа* представляет собой стеклянную трубку (рис. IV.92), наполненную парами ртути и аргоном. Стенки лампы покрыты изнутри тонким слоем люминесцирующего состава — люминофором. В концы трубки впаяны электроды в виде вольфрамовых спиралей. При работе лампы температура ее стенок близка к 50 °С и давление паров ртути при этом составляет около 10^{-2} мм рт. ст., а давление аргона — около 4 мм рт. ст.

При включении лампы в электрическую цепь с достаточно высоким напряжением между электродами происходит разряд в парах ртути (аргон только улучшает условия возбуждения паров ртути и замедляет разрушение электродов). Ртутные пары дают кроме видимого излучения еще и ультрафиолетовые лучи с длинами волн 0,254 и 0,185 мкм; эти лучи используются для возбуждения дополнительного видимого свечения люминофора, нанесенного на стенках лампы. Таким образом, в люминесцентных лампах часть электрической энергии сначала превращается в энергию ультрафиолетового излучения, а затем благодаря люминесценции в энергию видимого излучения.

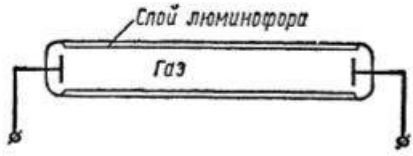


Рис. IV.92

Изменяя состав люминофоров, можно подобрать спектральный состав излучения люминесцентных ламп в соответствии с требованиями эксплуатации.

Основные преимущества люминесцентных ламп:

а) большая светоотдача (порядка 40—50 лм/Вт, что в 3—4 раза превышает светоотдачу ламп накаливания);

б) небольшая поверхностная яркость (около одного стильба);

в) большой срок службы (от 3000 до 10 000 ч, что в 3—10 раз превышает срок службы ламп накаливания).

К недостаткам люминесцентных ламп следует отнести сложность их включения и мелькание при питании переменным током.

2. *Применение люминесценции для создания слабых освещенностей* (аварийное и маскировочное освещение).

Такие освещенности необходимы для получения местного освещения, незаметного при наблюдении издали, или запасного освещения на случай аварий. Сюда же относится изготовление светящихся циферблатов и стрелок измерительных приборов и т. п.

3. *Люминесцентный анализ и дефектоскопия*. Так как люминесцентное излучение имеет спектр, характерный для каждого вещества, то можно обнаружить и исследовать различные объекты с помощью люминесценции.

Так, в биологии и микробиологии с помощью люминесценции можно наблюдать свечение отдельных бактерий. В археологии исследование старинных рукописей при ультрафиолетовом облучении позволяет читать на них стертые и попорченные места. В палеонтологии это дает возможность рассмотреть многие дополнительные детали отпечатков доисторических растений и животных, включенных в осадочные породы.

Цвет люминесценции позволяет отличать одни продукты от других, выявлять фальсификацию и примеси, обнаруживать подделки документов, характер пятен (судебная экспертиза).

Люминесцентный анализ по сравнению с химическим и спектральным обладает некоторыми преимуществами:

1) не требуется такого воздействия на исследуемое вещество, при котором могли бы измениться его состав, структура или агрегатное состояние (например, нет необходимости оказывать химические воздействия или переводить исследуемое вещество в газообразное состояние и т. п.);

2) очень высокая чувствительность, позволяющая обнаружить миллиардные доли процента люминесцирующих примесей в различных средах, растворах или смесях (например, следы нефти в породе и т. д.);

3) при люминесцентном анализе определяются не химические элементы, из которых состоит исследуемое вещество, а непосредственно наличие того или иного интересующего нас вещества: органического соединения, краски и т. п.;

4) простота методики анализа и дешевизна необходимой аппаратуры.

Важное значение имеет в настоящее время люминесцентная дефектоскопия, т. е. обнаружение невидимых глазу трещин в металлических изделиях с помощью люминесцентных веществ. Для этой цели поверхность исследуемой детали смазывается люминесцентным раствором, который заполняет и трещины. Далее люминесцирующий раствор смывается с поверхности образца, но в трещинах он остается. При облучении образца ультрафиолетовым светом трещина обнаруживается ярким свечением заключенного в ней раствора.

Широкое применение получил люминесцентный анализ при сортировке оптических стекол.

§ 18. ИНДУЦИРОВАННОЕ ИЗЛУЧЕНИЕ. ЛАЗЕРЫ

СПОНТАННЫЕ И ИНДУЦИРОВАННЫЕ ПЕРЕХОДЫ

Переход возбужденной системы (атома, молекулы) с верхних энергетических уровней на нижние может происходить либо спонтанно, либо индуцированно.

Спонтанным называется самопроизвольный (самостоятельный) переход, обусловленный только факторами, действующими внутри системы и свойственными ей. Эти факторы определяют среднее время пребывания системы в возбужденном состоянии; согласно соотношению Гейзенберга (см. § 11),

$$\Delta E\,\Delta t \geqslant \frac{h}{2\pi};\quad \Delta t \geqslant \frac{1}{\Delta E}\,\frac{h}{2\pi}.$$

Теоретически это время может иметь различные значения в пределах:

$$\Delta t_{\text{мин}} = \frac{1}{\Delta E_{\text{макс}}}\,\frac{h}{2\pi};\quad \Delta t_{\text{макс}} = \frac{1}{\Delta E_{\text{мин}}}\,\frac{h}{2\pi},$$

т. е. зависит от свойств системы — разброса ΔE значений энергии возбужденного состояния (за характеристику системы обычно принимается среднее значение времени пребывания в возбужденных состояниях в зависимости от среднего значения ΔE). Следует учесть также воздействие на систему окружающего пространства («физического вакуума»), в котором даже в отсутствие электромагнитных волн существует, согласно квантовой теории, флуктуирующее поле («вакуумные флуктуации»); это поле может стимулировать переход возбужденной системы к низшим уровням и должно быть включено в число неустранимых факторов, вызывающих спонтанные переходы.

Индуцированным называется вынужденный (стимулированный) переход в энергетически низшее состояние, вызванное каким-нибудь внешним воздействием на возбужденную систему: тепловыми столкновениями, взаимодействием с соседними частицами или проходящей через систему электромагнитной волной. Однако в литературе установилось более узкое определение: индуцированным называется переход, вызванный только электромагнитной волной, причем той же частоты, которая излучается системой при этом переходе (поля других частот не будут резонировать с собственными колебаниями системы,

поэтому их стимулирующее действие будет слабым). Так как «носителем» электромагнитного поля является фотон, то из этого определения следует, что при индуцированном излучении внешний фотон стимулирует рождение нового фотона такой же частоты (энергии).

Рассмотрим важнейшие особенности спонтанного и индуцированного переходов на одном простом идеализированном примере. Допустим, что в объеме V с зеркальными стенками имеется N одинаковых систем (атомов, молекул), из которых в начальный фиксированный момент времени $t_0 = 0$ некоторая часть ($N_{возб}$) переведена в возбужденное состояние с энергией ε; суммарная избыточная энергия в этом объеме будет равна $\varepsilon N_{возб} = W$. Для спонтанных переходов характерно следующее:

1) процесс перехода возбужденных систем в нормальные состояния (т. е. излучение избыточной энергии W) растянут во времени. Одни системы пребывают в возбужденном состоянии малое время τ, для других это время больше. Поэтому поток (мощность) излучения будет с течением времени изменяться, достигнет максимума в некоторый момент и затем будет асимптотически убывать до нуля. Среднее значение потока излучения будет равно

$$\Phi_e = \frac{W}{\tau_{ср}} = \frac{N_{возб}}{\tau_{ср}} h\nu;$$

2) момент времени, когда начинается излучение одной системы, и местонахождение этой системы совершенно не связаны с моментом излучения и местонахождением другой, т. е. между излучающими системами нет «согласованности» (корреляции) ни в пространстве, ни во времени. Спонтанные переходы являются совершенно случайными процессами, разбросанными во времени, по объему среды и по всевозможным направлениям; плоскости поляризации и фазы электромагнитных излучений от различных систем имеют вероятностный разброс, поэтому сами излучатели не являются источниками когерентных волн.

Для характеристики индуцированных переходов допустим, что в рассматриваемый объем V в момент времени $t_0 = 0$ вводится один фотон с энергией, в точности равной $h\nu = \varepsilon$. Имеется некоторая вероятность того, что этот фотон при одном из столкновений с невозбужденной системой поглотится ею; эта вероятность будет учтена ниже в более общем случае (когда в объеме V происходит взаимодействие рассматриваемых систем с фотонным газом). Будем полагать, что фотон не поглощается, многократно отражается от стенок сосуда и при столкновениях с возбужденными системами стимулирует излучение таких же фотонов, т. е. вызывает индуцированные переходы. Однако каждый появившийся при этих переходах новый фотон будет также возбуждать индуцированные переходы. Так как скорости фотонов велики, а размеры объема V малы, то понадобится очень малое время для того, чтобы все имеющиеся в начальный момент времени возбужденные системы были вынуждены перейти в нормальное состояние. Следовательно, для индуцированных переходов характерно следующее:

1) время τ, необходимое для излучения избыточной энергии W, может быть регулируемо и сделано очень малым, поэтому поток излучения $\Phi_e = W/\tau$ может быть очень большим;

2) кроме того, фотон, вызвавший переход, и фотон такой же энергии (частоты), появившийся при этом переходе, находятся в одинаковой фазе, имеют одинаковые поляризацию и направление движения. Следовательно, электромагнитные волны, образующиеся при индуцированном излучении, когерентны.

Однако не каждое столкновение фотона с возбужденной системой приводит к ее переходу в нормальное состояние, т. е. вероятность индуцированного перехода в каждом «акте взаимодействия» фотона с системой не равна единице. Обозначим эту вероятность через $w^0_{инд}$. Допустим, что в данный момент времени в объеме V имеется n фотонов и каждый из них в среднем может иметь z столкновений в единицу времени. Тогда число индуцированных переходов в единицу времени $\Delta N_{инд}$, следовательно, и число появившихся фотонов в объеме V будет равно

$$\Delta N_{инд} = w^0_{инд} n z.$$

Обозначим число возбужденных систем в объеме V через $N_{возб}$. Число столкновений z фотонов с возбужденными системами будет пропорционально концентрации таких систем, т. е. $z \sim N_{возб}/V$. Тогда $\Delta N_{инд}$ может быть выражено в зависимости от $N_{возб}$:

$$\Delta N_{инд} = w_{инд} n N_{возб}, \tag{2.83}$$

где $w_{инд}$ учитывает все другие факторы, кроме числа фотонов (n) и числа возбужденных систем ($N_{возб}$).

Увеличение числа фотонов в объеме V будет происходить также и вследствие спонтанного излучения. Вероятность спонтанного перехода $w_{сп}$ есть обратная величина среднего времени пребывания в возбужденном состоянии $\tau_{ср}$. Следовательно, число фотонов, появляющихся в единицу времени вследствие спонтанных переходов, будет равно

$$\Delta N_{сп} = \frac{N_{возб}}{\tau_{ср}} = w_{сп} N_{возб}. \tag{2.84}$$

Уменьшение числа фотонов в объеме V будет происходить в результате их поглощения невозбужденными системами (при этом будет увеличиваться число возбужденных систем). Так как не каждый «акт взаимодействия» фотона с системой сопровождается поглощением, то следует ввести *вероятность реализации поглощения* $w_{погл}$. Число столкновений в единицу времени одного фотона с невозбужденными системами будет пропорционально числу таких систем ($N_{невозб}$), поэтому по аналогии с (2.83) можно для убыли фотонов написать:

$$\Delta N_{погл} = w_{погл} n N_{невозб} = w_{погл} n (N - N_{возб}). \tag{2.85}$$

Найдем разность между интенсивностями процессов излучения и поглощения фотонов, т. е. процессов перехода систем из высших уровней на низшие и обратно:

$$\Delta N = \Delta N_{погл} - \Delta N_{инд} - \Delta N_{сп}. \tag{2.86}$$

В зависимости от значения ΔN в рассматриваемом объеме могут происходить следующие изменения:

1) если $\Delta N > 0$, то в этом объеме будет происходить постепенное уменьшение плотности фотонного газа, т. е. поглощение лучистой энергии. Необходимым условием для этого является малая концентрация возбужденных систем: $N_{\text{возб}} \ll N$;

2) если $\Delta N = 0$, то в системе установится равновесное состояние при некоторой определенной концентрации возбужденных систем и плотности лучистой энергии;

3) если $\Delta N < 0$ (что возможно при больших значениях $N_{\text{возб}}/N$), то в рассматриваемом объеме будет происходить увеличение плотности фотонного газа (лучистой энергии).

Очевидно, что уменьшение (п.1) или увеличение (п.3) энергии излучения будет иметь место не только в изолированном объеме с отражающими стенками, но и в том случае, когда поток монохроматической лучистой энергии (поток фотонов с частотой ν) распространяется в среде, содержащей возбужденные частицы с избыточной энергией $\varepsilon = h\nu$.

Найдем относительное изменение числа фотонов, приходящееся на один фотон и на одну систему; воспользовавшись (2.86), (2.83), (2.84) и (2.85), получим

$$\frac{\Delta N}{nN} = w_{\text{погл}} - \frac{N_{\text{возб}}}{N}\left(w_{\text{погл}} + w_{\text{инд}} + \frac{1}{n}\, w_{\text{сп}}\right). \tag{2.87}$$

Заметим, что в равновесном состоянии (которое возможно только при положительной температуре T), согласно формуле (2.42), приведенной в § 12, отношение $N_{\text{возб}}/N$ равно

$$\frac{N_{\text{возб}}}{N} = \frac{e^{-h\nu/kT}}{\sum e^{-h\nu/kT}} = \frac{e^{-h\nu/kT}}{e^{-\varepsilon_0/kT} + e^{-(\varepsilon_0 + h\nu)/kT}}. \tag{2.88}$$

Статистическая сумма в знаменателе в данном случае состоит только из двух слагаемых, соответствующих: 1) системам в нормальных состояниях с энергией ε_0 и 2) возбужденным системам с энергией $\varepsilon_0 + h\nu$. Из этой формулы следует, что при бесконечно большой положительной температуре $N_{\text{возб}}/N = 1/2$. Это означает, что *путем повышения температуры невозможно достигнуть состояния, при котором число возбужденных систем было бы больше числа невозбужденных.*

Вернемся к соотношению (2.87) и применим его к состоянию с большой положительной температурой, когда число фотонов n в объеме V очень велико; для этого состояния третьим членом в скобках можно пренебречь. Кроме того, существование положительной температуры означает равновесность состояния, т. е. $\Delta N = 0$. Если дополнительно выбрать T очень большим, то, согласно (2.88), $N_{\text{возб}} \approx N/2$. Тогда из (2.87) следует, что $w_{\text{инд}}$ должно быть равно $w_{\text{погл}}$. Эти величины не зависят от температуры, следовательно, полученное равенство $w_{\text{инд}} = w_{\text{погл}} = w$ выражает свойство системы, а не характеризует ее состояние. Тогда соотношение (2.87) можно записать в виде

$$\frac{\Delta N}{nN} \approx w\left(1 - 2\frac{N_{\text{возб}}}{N}\right). \tag{2.89}$$

Таким образом, для того чтобы ΔN было отрицательной величиной и можно было получить усиление лучистого потока, проходящего через данную среду, необходимо, чтобы $N_{возб}$ было больше, чем $N_{невозб}$, т. е. $N_{возб} > N/2$ (это необходимо, чтобы число фотонов, появляющихся при переходах на низшие уровни, было больше числа фотонов, поглощаемых за то же время). Выше было указано, что такое состояние не может быть достигнуто повышением температуры. Поэтому для получения среды, способной усиливать проходящий через нее лучистый поток, необходимо использовать другие (не температурные) способы возбуждения атомов и молекул.

Можно показать, что $N_{возб}$ может быть больше $N/2$ (т. е. $N_{возб} > N_{невозб}$) только при отрицательной температуре, т. е. при неравновесном состоянии рассматриваемой среды. Если, кроме того, это неравновесное состояние является метастабильным (см. ч. II, § 3), то можно при помощи подходящего внешнего воздействия вызвать скачкообразный переход к равновесному состоянию с освобождением избыточной энергии W за очень короткое время. Эта идея и лежит в основе работы лазеров.

Состояние среды, при котором верхние энергетические уровни имеют большие коэффициенты заполнения по сравнению с низшими, называется *инверсионным*. Так как в этом состоянии среда не ослабляет, как обычно, а усиливает проходящее через нее излучение, то в формуле для изменения интенсивности лучистого потока в среде

$$I = I_0 e^{-kx}$$

коэффициент k будет отрицательной величиной (следовательно, показатель степени — положительной величиной). Ввиду этого среду в инверсионном состоянии называют *средой с отрицательным показателем поглощения*. Возможность получения таких сред, их свойства и использование для усиления оптического излучения были установлены и разработаны В. А. Фабрикантом и его сотрудниками (1939—1951).

ЛАЗЕРЫ

Лазеры, разработанные Н. Г. Басовым, А. И. Прохоровым, Ч. Таунсом и др. (1954), представляют собой устройства, в которых определенными приемами создается метастабильное инверсионное состояние специально подобранной «активной среды». Освобождение избыточной энергии этой среды осуществляется либо слабым внешним воздействием — электромагнитным полем, имеющим необходимую частоту колебаний («входной сигнал»), либо производится фотонами, появившимися в этой же среде вследствие спонтанного излучения. Таким образом, лазер может работать либо как у с и л и т е л ь входного сигнала, либо как г е н е р а т о р когерентного монохроматического излучения большой мощности.

Используется несколько способов приготовления инверсионных состояний:

1) метод увеличения концентрации возбужденных атомов вещества путем с е п а р а ц и и (удаления из среды невозбужденных атомов);

2) метод «оптической накачки» — возбуждения атомов среды внешним излучением;

3) путем бомбардировки электронами атомов среды (например, при газовом разряде в активной среде);

4) передачей энергии возбуждения от неизлучающих атомов излучающим при тепловых столкновениях.

Рассмотрим первый способ на примере аммиачного *газового лазера*, схематическое устройство которого показано на рис. IV.93.

В возбужденной молекуле аммиака атом азота совершает колебания с частотой 23 870 МГц, вследствие чего излучается электромагнитная волна длиной 1,26 см. В камере K газовый разряд создает смесь возбужденных и невозбужденных атомов, однако $N_{\text{возб}} < N_{\text{невозб}}$. Выходящий из отверстия пучок атомов проходит через сепаратор C, в котором имеется неоднородное электрическое поле. Невозбужденные молекулы аммиака, имеющие постоянный электрический момент, частично отклоняются в сторону, вследствие чего в резонатор P попадает смесь, в которой концентрация возбужденных атомов достаточно велика. В резонатор (весьма точно настроенный на длину волны 1,26 см) подается слабое электромагнитное поле, колеблющееся с частотой 23 870 МГц («входной сигнал», условно обозначенный на рисунке стрелкой *а*). Вызванные им индуцированные переходы возбужденных атомов аммиака образуют мощный «выходной сигнал», отводимый из резонатора через волновод (стрелка *б*).

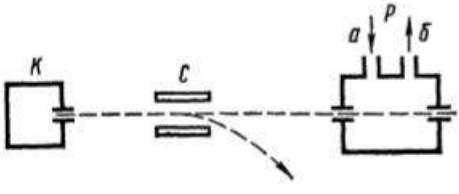


Рис. IV.93

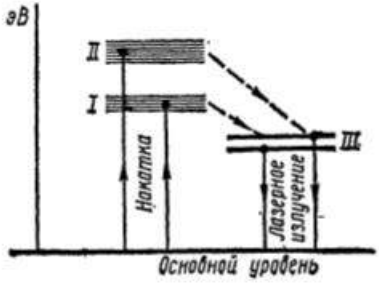


Рис. IV.94

Метод оптической накачки используется в *рубиновых лазерах*. На рис. IV.94 показана схема энергетических уровней в рубине (кристалл Al_2O_3, в котором часть ионов алюминия заменена ионами хрома: 0,05% — в розовом и 0,5% — в красном рубине). На рис. IV.95 показано устройство самого лазера. Рубиновый стержень P помещен внутри мощной газонаполненной (неоном, криптоном, ксеноном) лампы $Л$, возбуждаемой импульсным разрядом конденсатора (время вспышки около 0,0005 с). Излучение этой лампы ($\lambda \approx 5600 \cdot 10^{-10}$ м), поглощенное ионами хрома, переводит их на широкие (и поэтому имеющие малое время существования, порядка 10^{-7} с) уровни *I* и *II* (этот процесс называется *оптической накачкой*). Из уровней *I* и *II* происходит переход на узкие метастабильные уровни *III* с временем существования порядка $5 \cdot 10^{-3}$ с; эти переходы не сопровождаются излучением. Таким образом, со временем ионы хлора оказываются в состояниях с энерги-

ями, соответствующими уровню *III* (этот уровень состоит из двух очень близких подуровней).

Для того чтобы вызвать мощное индуцированное лазерное излучение, торцы рубинного стержня (строго параллельные) тщательно шлифуются и покрываются серебром с таким расчетом, чтобы от одного торца излучение полностью отражалось, а второе было полупрозрачно. При спонтанных переходах в объеме рубина появятся фотоны, летящие в различных направлениях. Фотоны, летящие под большими углами к оси стержня, быстро покинут объем стержня (вместе с фотонами, появление которых было ими стимулировано). Фотоны, летящие в направлениях, параллельных оси стержня, будут многократно отражаться от зеркальных торцов стержня и вследствие этого вызовут большое число индуцированных переходов. Фотоны, появившиеся при переходах, движутся в том же направлении, что и первичные, и поэтому поток фотонов, параллельных оси стержня, будет лавинообразно

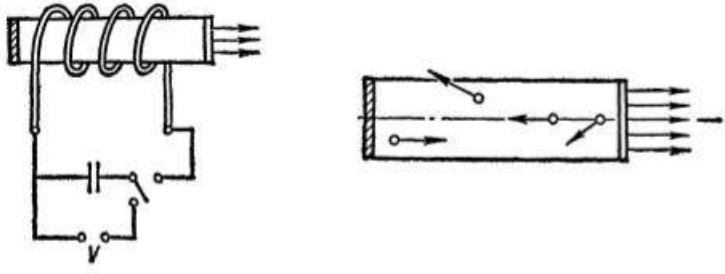
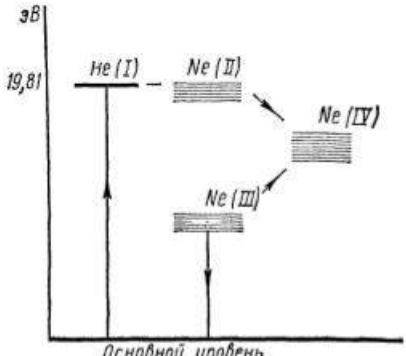


Рис. IV.95

нарастать. Ввиду большой скорости фотонов и малых размеров стержня понадобится весьма малое время для перевода возбужденных ионов хрома в основное (нормальное) состояние. Образовавшийся мощный поток фотонов выходит из стержня через полупрозрачный торец. Излучение рубинового лазера состоит из двух спектральных линий с длинами волн 6927 и $6943 \cdot 10^{-10}$ м, соответствующих переходам от каждого подуровня *III* на основной уровень.

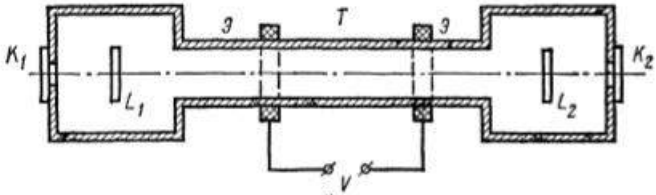
Третий и четвертый способы приготовления инверсионных сред можно показать на примере гелий-неонового газового лазера непрерывного действия. На рис. IV.96, *а* приведена схема энергетических уровней гелия и неона, используемых в этом лазере, а на рис. IV.96, *б*— схема его устройства. Трубка *T* содержит гелий при парциальном давлении около 1 мм рт. ст. и неон при давлении около 0,1 мм рт. ст. Возбуждение атомов осуществляется газовым разрядом, вызванным напряжением *U*, приложенным к электродам *Э*. Электроны, получившие в трубке большие скорости, при столкновениях с атомами гелия переводят их на метастабильный уровень *1*. Затем при столкновениях возбужденных атомов гелия с невозбужденными атомами неона происходит передача энергии: атом гелия возвращается в нормальное состояние (без излучения), а атом неона переходит из нормального состояния на свой уровень *II*, который является метастабильным (передача энергии облегчена тем, что энергия уровня *1* гелия весьма близка к энергии

уровня *II* неона; разность энергий этих уровней 0,04 эВ переходит в кинетическую энергию атомов; для облегчения этих процессов и подобраны оптимальные значения парциальных давлений гелия и неона). Лазерное излучение соответствует переходу атомов неона с уровня *II* на уровень *IV* (когда на уровне *II* будет достигнута инверсионная по отношению к *IV* концентрация возбужденных атомов неона); излучаются инфракрасные волны длиной $\lambda = 11523 \cdot 10^{-10}$ м, выходящие из установки через полупрозрачные зеркала L_1 и L_2 и кварцевые окна K_1 и K_2. В дальнейшем происходят каскадные переходы: с уровня *IV* на уровень *III* с излучением характерной для неона спектральной линии $\lambda \approx 6000 \cdot 10^{-10}$ м и затем с уровня *III* на основной без излучения, с передачей энергии через столкновения другим атомам и стенкам сосуда.



а)

Отражающий слой на зеркалах *L* (строго параллельных между собой) состоит из большого числа (около 10) чередующихся тонких слоев ($d \approx \lambda/4$) сернистого цинка и фтористого магния, напыленных на стеклянную или кварцевую пластинку; такой слой имеет хороший коэффициент отражения для инфракрасного излучения $\lambda = 11\,523 \cdot 10^{-10}$ м. «Слоистые» зеркала весьма полезны при подборе необходимого коэффициента отражения для тех или иных узких участков спектра или спектральных линий.



б)

Рис. IV.96

Несколько иной принцип работы имеют *полупроводниковые лазеры*; на рис. IV.97 показаны расположение энергетических уровней и схематическое устройство такого лазера. В качестве активной среды (кристалл *K*) подбираются вещества, у которых вероятность перехода электронов из зоны проводимости в валентную зону с испусканием фотона велика. Таким веществом оказался арсенид галлия (GaAs), у которого путем ввода (в качестве примесей) доноров — атомов Te, Se, Sn и других и акцепторов — атомов Zn, Cd и других создаются зоны *p*- и *n*-проводимостей. Если к граням кристалла, параллель-

ным плоскости p-n-перехода, при помощи электродов $Э$ приложить напряжение, то электроны из n- зоны будут переходить в p-зону и рекомбинировать с «дырками», движущимися в обратном направлении. При этом возникает излучение, распространяющееся вдоль плоскости p-n-перехода перпендикулярно зеркальным поверхностям L_1 и L_2 (которые предназначены для получения лазерного эффекта). Возможность образования инверсионного состояния кристалла достигается введением примесей и образованием дополнительных энергетических уровней между валентной зоной и зоной проводимости кристалла. Эти уровни локализуются вблизи атомов примесей и играют роль потенциальных ям для электронов. Кроме того, оказалось необходимым содержать кристалл при низких температурах.

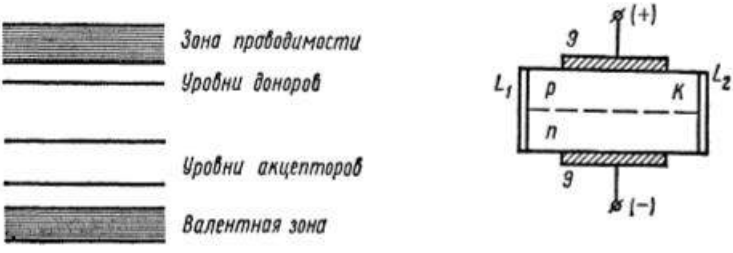


Рис. IV.97

Возбуждение полупроводника (создание неравновесных концентраций носителей тока — электронов и «дырок») может производиться различными способами: 1) при помощи сильного импульсного электрического поля. Величина поля зависит от разности энергий между валентной зоной и зоной проводимости, т. е. от энергетической ширины запретной зоны (и если эта ширина велика, то возникают большие тепловые потери); 2) при помощи «оптической накачки» — либо обычными источниками света (что оказалось нецелесообразным, так как длинноволновые участки спектра не приводят к возбуждению, а коротковолновые — приводят к энергетическим потерям), либо излучением от какого-нибудь (например, рубинового) лазера, что означает преобразование коротковолнового излучения возбуждающего лазера в длинноволновое излучение возбуждаемого; 3) бомбардировкой электронным пучком, что позволяет получить лазерное излучение от полупроводников в широком диапазоне длин волн. Важные исследования полупроводниковых лазеров проведены Н. Г. Басовым, Б. М. Вулом, Ю. М. Поповым, О. Н. Крохиным и др.

ПРЕИМУЩЕСТВА ЛАЗЕРОВ

Лазеры как источники излучения обладают следующими преимуществами:

1) лазеры дают весьма *монохроматическое когерентное* излучение. Напомним, что монохроматичность и когерентность излучения взаимосвязаны. Если в данную точку приходят два цуга волн одинаковой

частоты и плоскости поляризации, то в течение короткого времени τ, пока в этой точке существуют обе волны, их можно полагать когерентными, так как в течение этого времени результат их интерференции изменяться не будет. Некогерентность таких ограниченных цугов волн проявляется в том, что в следующие промежутки времени через ту же точку проходят уже другие цуги волн с иной разностью фаз и, следовательно, с иным результатом интерференции. Поэтому чем длиннее цуги волн, тем в большей степени (в течение большего времени) они будут когерентны. Совершенно когерентными будут бесконечно длинные, т. е. гармонические, волны. С другой стороны, короткие цуги волн не являются монохроматическими; если их разложить на гармонические составляющие, то получается широкая спектральная линия, содержащая непрерывный интервал частот $\Delta\nu$, тем больший, чем короче цуг волн. Длинные цуги волн имеют малые $\Delta\nu$ и поэтому являются более монохроматическими. Таким образом, чем больше степень когерентности излучения (т. е. длиннее цуги испускаемых волн), тем более монохроматическим является это излучение; бесконечная гармоническая волна является идеально когерентной и одновременно идеально монохроматической. Следовательно, всякое устройство, которое обеспечивает (повышает) когерентность излучения, способствует также монохроматизации этого излучения (уменьшению разброса частот $\Delta\nu$ в пределах спектральных линий). Заметим кстати, что при импульсном режиме работы лазера, когда электромагнитная волна прерывается, излучение получается менее когерентным, чем при непрерывном режиме;

2) *мощность излучения* лазеров весьма велика. Например, если рубиновый стержень получил от лампы при импульсной накачке энергию $W = 50$ Дж и высветился за $t = 0{,}5 \cdot 10^{-3}$ с, то поток излучения будет равен

$$\Phi_e = \frac{50}{0{,}5 \cdot 10^{-3}} \text{ Дж/с} = 10^5 \text{ Вт} = 100 \text{ кВт}.$$

При помощи линз можно сфокусировать это излучение на небольшой площадке и получить на ней колоссальные концентрации мощности; например, если диаметр фокального пятна равен 1,1 мм, то $S = \pi d^2/4 \approx 1$ мм^2 и

$$\frac{\Phi_e}{S} = \frac{100}{10^{-6}} \text{ кВт/м}^2 = 10^8 \text{ кВт/м}^2.$$

При такой концентрации мощности (энергетической освещенности) можно произвести почти мгновенные процессы в различных телах. Однако не следует забывать, что энергия, содержащаяся в излучении лазера, все же невелика, поэтому процессы, требующие большой затраты энергии, не могут быть осуществлены излучением, не содержащим этой энергии, как бы ни были велики поток излучения или концентрации мощности на площади или в объеме. Очевидно, необходимую энергию можно получить, увеличивая размеры лазера, повторяя импульсные выбросы лучистой энергии с большой частотой либо используя лазеры непрерывного действия. Таким образом, лазер работает

как прибор, концентрирующий поступающую извне и разбросанную во времени энергию;

3) *напряженности электрического и магнитного полей* в лазерной волне весьма велики. Если, например, в солнечном свете $E = 7$ В/см, то в излучении лазера средней мощности напряженность E достигает десятков и сотен киловольт на сантиметр. Расчет производится следующим образом: допустим, что из рубинового стержня за время $t = 5 \cdot 10^{-4}$ с выходит поток излучения с сечением $S = 1$ см$^2 = 10^{-4}$ м^2 и мощностью $\Phi_e = 10^6$ Вт. В объеме $V = Sct$ (c — скорость света) будет содержаться энергия $W = \Phi_e t$. Так как плотность энергии в электромагнитной волне

$$w = \frac{\varepsilon_0 E^2}{2} = \frac{W}{V} = \frac{\Phi_e t}{Sct} = \frac{\Phi_e}{Sc},$$

то

$$E = \sqrt{\frac{2w}{\varepsilon_0}} = \sqrt{\frac{2\Phi_e}{Sc\varepsilon_0}} = \sqrt{\frac{2 \cdot 10^6}{10^{-4} \cdot 3 \cdot 10^8 \cdot 8{,}85 \cdot 10^{-12}}} \approx 1{,}15 \cdot 10^7 \text{ В/м}.$$

Следовательно, лазерный луч может осуществить процессы, требующие больших напряженностей электрического и магнитного полей;

4) *направленность* излучения лазеров может быть сделана очень точной (угловое расхождение в пучке — очень малым).

Если зеркала лазера строго параллельны, то лазерный луч будет представлять собою плоскую волну. Дифракция этой волны на круглом отверстии (торец рубинового стержня) диаметра D дает угловое расхождение для пятна Эйри (см. § 5)

$$\varphi \approx 1{,}22\lambda/D.$$

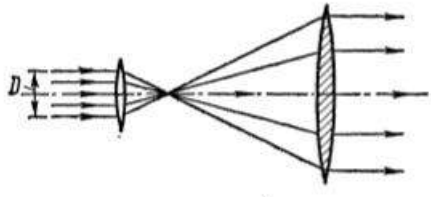


Рис. IV.98

При $\lambda = 7000 \cdot 10^{-10}$ м $D = 3{,}5$ мм $= 0{,}0035$ м, $\varphi \approx 1{,}22 \times (7 \cdot 10^{-7})/(3{,}5 \cdot 10^{-3}) \approx 2{,}44 \cdot 10^{-4}$ рад, или около 0,75 мин (45″). Если воспользоваться двумя линзами (рис. IV.98), то можно увеличить диаметр D лазерного пучка до 10 раз и получить угол расхождения порядка нескольких секунд. При такой направленности лазерного излучения пятно, образованное на Луне ($L \approx 3{,}8 \cdot 10^8$ м), имело бы радиус

$$r = L\varphi = 3{,}8 \cdot 10^8 \cdot 2{,}44 \cdot 10^{-5} \approx 9{,}3 \text{ км}.$$

Заметим, что при использовании коротковолновых радиоволн, например $\lambda = 1$ см, для получения пятна таких размеров понадобилось бы фокусирующее зеркало диаметром около 500 м;

5) *спектральная плотность энергетической светимости* в излучении лазеров значительно превосходит соответствующую плотность спонтанного излучения, а также теплового излучения при той же температуре. Например, в рубиновом лазере атомы хрома излучают волну $\lambda = 6943 \cdot 10^{-10}$ м с разбросом, равным при спонтанных переходах $\Delta\lambda = 5{,}4 \cdot 10^{-10}$ м, в лазерном режиме $\Delta\lambda = 0{,}02 \cdot 10^{-10}$ м. Следова-

тельно, одна и та же мощность спектральной линии в лазерном режиме приходится на интервал длин волн, в 270 раз меньший, чем при спонтанном излучении; поэтому яркость лазерной спектральной линии $r_\lambda = \Delta\Phi_{e\lambda}/\Delta\lambda$ будет во столько же раз больше.

Если лазерное устройство предназначено для получения когерентного излучения, то его качество может быть оценено по содержанию в этом излучении некогерентных волн. Кроме вынужденных переходов в активной среде лазера происходят также и спонтанные переходы. Та часть излучения, которая приходится на долю спонтанных переходов, является некогерентной и должна рассматриваться как п о м е х а по отношению к основному когерентному излучению; эту часть излучения называют *шумом* (термин, принятый в радиотехнике). Характеристикой когерентного источника света является отношение интенсивности шума к интенсивности «упорядоченного» когерентного излучения. С этой точки зрения тепловые источники света являются генераторами шумов в широкой области спектра. Радиотехнические генераторы электромагнитных волн имеют малый уровень шумов, благодаря чему возможна передача обширной информации при помощи модуляции и демодуляции этих волн;

6) *коэффициент полезного действия* лазеров, использующих различные «активные» вещества, колеблется в широких пределах: от долей процента до значений, приближающихся к 100%.

К. п. д. газового лазера на гелии и неоне очень мал (около 1%). Более высокий к. п. д. (до 10%) получен у газового лазера, содержащего смесь углекислого газа с азотом (иногда с небольшой примесью гелия). В газе вызывается электрический разряд, при котором энергия электронов оказывается достаточной для перевода молекул азота на низший колебательный уровень (E_1). Однако для атомов азота переход с этого уровня на основной (E_0) с испусканием фотона запрещен (имеет очень малую вероятность реализации), но возможна безызлучательная передача энергии молекулам CO_2 при тепловых столкновениях; эта передача облегчена тем, что энергия $E_1—E_0$ мало отличается от энергии, необходимой для возбуждения молекул CO_2 на «лазерный уровень» $E_л$. У молекул CO_2 имеются еще и другие, более низкие уровни энергии; существенно, что они освобождаются значительно быстрее, чем уровень $E_л$ (вследствие столкновений с невозбужденными молекулами и переходом энергии возбуждения в теплоту). Таким образом, у молекул CO_2 верхние уровни (и в том числе главный уровень $E_л$) оказываются более заполненными по сравнению с низшими уровнями, что обеспечивает возможность появления лазерного эффекта. Большим преимуществом лазеров на углекислом газе является (кроме высокого к. п. д.) возможность получения больших мощностей.

Полупроводниковые лазеры имеют к. п. д., близкий к 100%. Это особенно ценно в приборах, где используются *миниатюрные лазеры* (полупроводниковые лазеры имеют активный слой толщиной в несколько микрон, а сам кристалл может иметь размеры порядка 2—3 мм). Недостатком полупроводниковых лазеров по сравнению с газовыми является большая «размытость» спектральных линий. Даже если (для устранения «шумов») поместить лазер на арсениде галлия в тер-

мостат с температурой жидкого азота (—196° С), то излучаемая им линия $\lambda = 8430 \cdot 10^{-10}$ м имеет ширину $0{,}1 \cdot 10^{-10}$ м, т. е. размытость линии $\sim 1 : 10^5$; в газовых лазерах эта величина значительно меньше ($1 : 10^{10}$ и менее).

ПРИМЕНЕНИЕ ЛАЗЕРОВ

Перечисленные свойства лазерного излучения обеспечили широкое применение лазеров: в научных исследованиях (в физике, химии, биологии, медицине и т. д.), в технических процессах (сварка и резка материалов на участках порядка микронов; пробивание отверстий в сверхтвердых и хрупких телах; плавление и испарение, требующие высоких концентраций энергии и т. д.), в измерительной технике (создание сверхточных часов; измерение расстояний космических масштабов; управление искусственными спутниками и т. п.), в медицинской практике (лечение опухолей, глазных и кожных заболеваний и т. п.).

Лазеры, генерирующие излучение в оптическом диапазоне, называются также *оптическими квантовыми генераторами*. В области радиотехнических волн СВЧ (сверхвысоких частот) аналогичные устройства получили название *мазеров*.

В заключение приведем два интересных исследования. В приведенных выше рассуждениях всегда предполагалось, что поглощение веществом фотонов состоит из элементарных актов, в которых один атом поглощает один фотон; вероятность одновременного поглощения двух фотонов полагалась исчезающе малой и в рассуждениях игнорировалась. Однако благодаря той колоссальной плотности фотонов, которая имеется в лазерном излучении, удалось обнаружить «двухфотонное поглощение». Опыт состоял в следующем: ионы европия в кристалле фтористого кальция возбуждаются голубым светом с $\lambda = 4250 \cdot 10^{-10}$ м. Рубиновый лазер, излучающий $\lambda \approx 7000 \cdot 10^{-10}$ м, может вызвать это возбуждение только в том случае, если ион европия поглотит два фотона этого излучения. В обнаружении такого поглощения важное значение имело то обстоятельство, что вероятность однофотонного поглощения пропорциональна первой степени, а вероятность двухфотонного поглощения — квадрату интенсивности падающего излучения. Измерения показали, что интенсивность голубого излучения кристалла увеличивается пропорционально квадрату интенсивности красного света рубинового лазера.

Другое явление — самофокусировка и самоканализация интенсивных пучков света (Г. Аскарян, 1962 г.) — заключается в следующем. Допустим, что идеальный пучок лазерного излучения проходит через среду с показателем преломления n_0. Однако в объеме пучка при большой концентрации электромагнитной энергии (больших значениях напряженности поля E) может возникнуть столь сильная поляризация молекул среды, что показатель преломления в пределах этого объема окажется больше, чем за пределами пучка (при нелинейной зависимости между поляризуемостью среды и вектором внешнего электрического поля значение n в объеме пучка будет выражаться формулой $n \approx n_0 + aE^2$, где a — некоторый коэффициент). Тогда лазерный пучок

окажется внутри прозрачного шнура («оптического волновода»), на цилиндрической поверхности которого возможно полное внутреннее отражение. Если, например, в в а к у у м е луч *1* (рис. IV.99) мог бы выйти за пределы лазерного пучка, то в с р е д е этот луч будет испытывать полное внутреннее отражение от границы шнура и останется

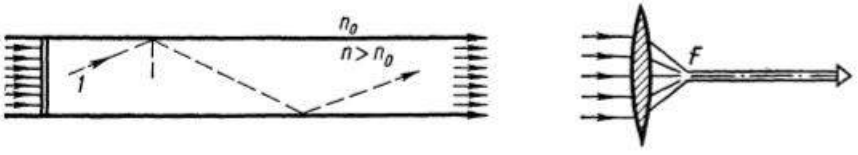


Рис. IV.99

в пределах пучка. Можно показать, что кроме такой «самоканализации» пучка будет иметь место и некоторое уменьшение сечения пучка («самофокусировка»). В частности, лучи, собранные линзой в фокусе F, в дальнейшем не расходятся, а образуют тонкий световой шнур очень малого диаметра (порядка нескольких длин волн).

Глава 4
АТОМНЫЕ ЯДРА И ЭЛЕМЕНТАРНЫЕ ЧАСТИЦЫ

§ 19. ОСНОВНЫЕ СВОЙСТВА АТОМНЫХ ЯДЕР

СОСТАВ И СТРУКТУРА ЯДЕР

Свойства атомных ядер определяются их составом и структурой
1) числами протонов Z и нейтронов N в ядре;
2) свойствами этих частиц; характером взаимодействия между ними;
3) относительным расположением и движением протонов и нейтронов внутри ядра.

Протоны и нейтроны как составные части ядер объединены общим названием — *нуклоны*; $Z + N = A$ есть число нуклонов в ядре.

Ядра принято обозначать химическим символом элемента, снабженным числами, указывающими содержание ядра, по схеме

$$^{A}_{Z}X_{N}.$$

Например, ядро натрия, содержащее 11 протонов и 12 нейтронов, записывается так: $^{23}_{11}Na_{12}$.

Известно и изучено около 1300 ядер, из которых 267 являются стабильными. Большинство стабильных ядер имеют четные значения Z и N. Из них 159 являются четно-четными, т. е. имеют одновременно четные и Z, и N; 54 ядра являются четно (Z)-нечетными (N), 50 — нечетно-четными и лишь 5 стабильных ядер являются нечетно-нечетными:

$$^{2}_{1}H_{1};\quad ^{6}_{3}Li_{3};\quad ^{10}_{5}B_{5};\quad ^{14}_{7}N_{7};\quad ^{50}_{23}V_{27}$$

(впрочем, ванадий имеет очень большой период полураспада — около 10^{15} лет). Эти сведения показывают, что при комплектовании внутренней структуры стабильных ядер четность чисел протонов и нейтронов имеет важное значение.

Ядра, имеющие:

1) одинаковые Z, но разные N, называются *изотопами*;
2) одинаковые N, но разные Z, называются *изотонами*;
3) разные Z и N, но одинаковые $A = Z + N$, называются *изобарами*.

Кроме того, могут существовать ядра, имеющие одинаковый состав (Z и N), но отличающиеся некоторыми свойствами, в частности периодами полураспада. Такие ядра называются *изомерами*.

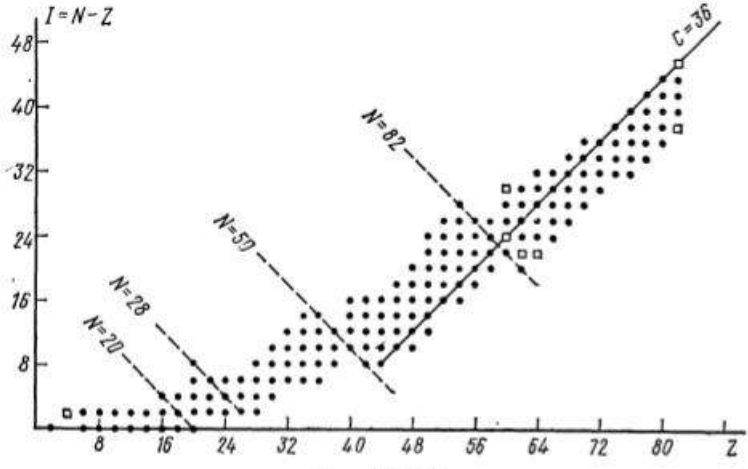


Рис. IV.100

В исследовательской аппаратуре ядра, особенно тяжелые, изучаются вместе с окружающей их электронной оболочкой. Ввиду этого введен новый термин *нуклид* — ядро вместе с электронной оболочкой, нейтрализующей его заряд. Атом же есть химическое понятие; в большинстве случаев химический элемент представляет собой смесь изотопов (следовательно, нейтральные атомы — смесь нуклидов с одинаковыми числами протонов).

Число $I = N - Z$ показывает избыток нейтронов в ядре. Имеется только одно ядро с отрицательным значением I (гелий-3 имеет $I = -1$); у остальных ядер избыток нейтронов является положительным числом и возрастает с увеличением атомного номера. На рис. IV.100 показано распределение стабильных четно-четных ядер по значениям Z и I. Квадратиками обозначены радиоактивные ядра, имеющие очень большие периоды полураспада (10^4 лет и более). Ядра, имеющие одинаковое число нейтронов (изотопы), лежат на прямых, проведенных под углом 45° к осям координат. На рисунке проведено только несколько таких линий, соответствующих $N = 20$, 28, 50 и 82. Ядра, расположенные на прямых, перпендикулярных линиям $N = \text{const}$, имеют одинаковое значение величины $C = 2Z - N$ (C — число нейтронов, которое не-

обходимо прибавить к ядру, чтобы число нейтронов в нем сделалось вдвое больше числа протонов). Рис. IV.100 показывает, что семейства ядер с $C = \text{const}$ более многочисленны, чем семейства $Z = \text{const}$ (изотопы), $N = \text{const}$ (изотоны) и $I = \text{const}$.

Два ядра, у которых число протонов одного равно числу нейтронов другого

$$Z_1 = N_2; \quad N_1 = Z_2; \quad A_1 = A_2,$$

называются *зеркальными*. К ним относятся, например,

$$^{3}_{1}H_2 \text{ и } ^{3}_{2}He_1; \quad ^{11}_{5}B_6 \text{ и } ^{11}_{6}C_5 \text{ и др.}$$

Имеется 19 пар зеркальных ядер; характерно, что в 16 из этих пар одно ядро является стабильным, второе — радиоактивным (в остальных трех парах оба ядра являются радиоактивными).

Ядра, у которых числа протонов или нейтронов равны 2, 8, 20, 28, 50, 82 (а для нейтронов еще и 126), отличаются большой распространенностью и резко выделяются по основным свойствам среди соседних ядер. Эти ядра получили название *магических*. Ядра, у которых магическими являются оба числа Z и N, называются *дважды магическими*; к ним относятся, например,

$$^{4}_{2}He_2; \quad ^{16}_{8}O_8; \quad ^{208}_{82}Pb_{126}$$

Для иллюстрации приведем относительное содержание (c, %) изотопов в естественном элементе — хроме ($Z = 24$) при различных значениях N:

N	26	28	29	30
c	4,31%	83,76%	9,55%	2,8%

Магическое значение $N = 28$ резко выделяется среди остальных. У трех соседних элементов — кадмия ($Z = 48$), олова ($Z = 50$) и теллура ($Z = 52$), имеющих одинаковое значение $N = 68$, наибольшее содержание имеет изотоп олова с магическим значением $Z = 50$. Для $N = 68$:

Z	48	50	52
c	7,58%	24,01%	0,089%

Приведем распространенность в Земной коре ряда соседних элементов (по А. Е. Ферсману, 1959 г.; распространенность водорода принята за единицу):

Индий ($Z=49$)	$1 \cdot 10^{-5}$	Таллий ($Z=81$)	$1 \cdot 10^{-5}$
Олово ($Z=50$)	0,003	Свинец ($Z=82$)	0,0016
Сурьма ($Z=51$)	$5 \cdot 10^{-5}$	Висмут ($Z=83$)	$1 \cdot 10^{-5}$

Элементы с магическими значениями чисел протонов имеют большую распространенность.

Кроме того, по сравнению с соседними ядрами:

1) магические ядра очень слабо поглощают нейтроны;

2) у радиоактивных магических ядер энергия выбрасываемых альфа-частиц очень мала;

3) для перевода магического ядра из основного состояния в ближайшее возбужденное состояние требуется значительно бо́льшая энергия, чем у соседних ядер.

Радиоактивность ядер есть самопроизвольное (обусловленное только внутренними факторами, действующими в ядре) превращение неустойчивых ядер одного элемента в ядра другого элемента. Радиоактивность элементов, существующих в природе, называется *естественной*; к ним относятся элементы, расположенные в конце периодической системы Менделеева, а также калий. Искусственные радиоактивные ядра могут быть получены при бомбардировке ядер протонами, нейтронами, альфа-частицами, а также ионами элементов (этим частицам сообщается энергия, достаточная для проникновения в ядро-мишень).

Изучение радиоактивности показало, что распад ядра следует полагать совершенно случайным событием; невозможно предвидеть момент распада данного ядра. Одни ядра могут существовать в неустойчивом состоянии большое время, другие — малое. Из наблюдений за радиоактивностью большого числа одинаковых ядер определяют либо среднее «время жизни» ядра, либо вероятность распада.

Допустим, что в начальный момент времени $t_0 = 0$ имеется N_0 радиоактивных ядер, а в момент t их стало N. Число dN ядер, испытавших распад за время dt, очевидно, будет пропорционально N и dt.

Вероятностью λ *распада* (или *постоянной распада*) называется относительное уменьшение числа радиоактивных ядер в единицу времени, следовательно,

$$\lambda = -\frac{dN}{N\,dt}; \quad dN = -\lambda N\,dt; \quad N = N_0 e^{-\lambda t}. \tag{4.1}$$

Время T, в течение которого число радиоактивных ядер уменьшается вдвое, называется *периодом полураспада*. Из (4.1) следует, что

$$\frac{1}{2}N_0 = N_0 e^{-\lambda T}; \quad T = \frac{\ln 2}{\lambda} = \frac{0{,}693}{\lambda}. \tag{4.2}$$

Число ядер, распадающихся в единицу времени, называется *активностью* данного радиоактивного элемента:

$$\left|\frac{dN}{dt}\right| = \lambda N_0 e^{-\lambda t} = \lambda N = \frac{N \ln 2}{T}. \tag{4.3}$$

Следовательно, активность радиоактивного вещества в данный момент времени прямо пропорциональна имеющемуся числу радиоактивных ядер и обратно пропорциональна их периоду полураспада, поэтому она убывает со временем по экспоненциальному закону.

Единицей для измерения радиоактивности выбрана *кюри*—активность, соответствующая $3{,}7 \cdot 10^{10}$ расп./с. Заметим, что единица активности определяется не по количеству радиоактивных ядер, а по интенсивности распада. Например, кюри соответствует активности 1 г радия или 3 т урана-238. Чем меньше период полураспада, тем меньшее количество данного вещества будет иметь единичную активность. На практике используется также *милликюри* $= 10^{-3}$ Ки $= 3{,}7 \cdot 10^7$ расп./с.

Внутри ядра действуют:

1) *электрические* силы отталкивания между протонами и

2) *ядерные* силы между нуклонами (отталкивания — на малых и притяжения — на больших расстояниях).

Установлено, что ядерные силы одинаковы для нуклонов обоих сортов. Ядерное притяжение между протонами значительно превосходит электрическое отталкивание, вследствие чего протон прочно удерживается в составе ядра.

Ядро окружено потенциальным барьером, обусловленным ядерными силами. Выход из ядра нуклона и системы нуклонов (например, альфа-частиц) возможен либо путем «туннельного эффекта», либо при получении энергии извне. В первом случае происходит спонтанный радиоактивный распад ядра, во втором — вынужденная ядерная реакция. Оба процесса позволяют вынести некоторые суждения о размерах ядра. Ценные сведения о протяженности потенциального барьера вокруг ядер получены при изучении рассеяния ядрами различных бомбардирующих частиц — электронов, протонов, нейтронов и др.

Исследования показали, что ядерные силы притяжения между нуклонами очень быстро убывают с увеличением расстояния между ними. Средний радиус действия ядерных сил, который можно трактовать так же, как некоторый условный («эффективный») размер ядра, на основании экспериментальных данных выражается оценочной формулой

$$r = (1{,}2 \div 1{,}5) \cdot 10^{-15} \sqrt[3]{A},\ \text{м}.$$

Если полагать, что ядра с большим числом нуклонов $A = Z + N$ состоят из сердцевины, где частицы равномерно распределены по объему, и сферической оболочки, в которой плотность частиц убывает к границам ядра до нуля, то в этом случае

$$r = 1{,}2 \cdot 10^{-15} \left(0{,}7 + \sqrt[3]{A}\right),\ \text{м}.$$

Эти формулы показывают, что «эффективный» объем ядра $(4/3)\pi r^3$ прямо пропорционален числу нуклонов A, поэтому нуклоны во всех ядрах упакованы в среднем с почти одинаковой плотностью.

Плотность ядер весьма велика; например, у ядра с $A = 81$ масса $M \approx 81 \cdot 1{,}67 \cdot 10^{-27}$ кг, радиус

$$r \approx 1{,}2 \cdot 10^{-15} \left(0{,}7 + \sqrt[3]{81}\right) \approx 4{,}4 \cdot 10^{-15}\ \text{м},$$

плотность

$$\rho = \frac{M}{(4/3)\,\pi r^3} = \frac{81 \cdot 1{,}67 \cdot 10^{-27}}{(4/3) \cdot 3{,}14 \cdot (4{,}4 \cdot 10^{-15})^3} \approx 3{,}6 \cdot 10^{17}\ \text{кг/м}^3.$$

Состояние нуклона в различных местах внутри ядра можно характеризовать величиной энергии E, которую нужно затратить, чтобы извлечь этот нуклон из ядра. Она называется *энергией связи* данного нуклона в ядре. В общем случае эта энергия различна для протонов и нейтронов и может зависеть от того, в каком месте объема ядра находится данный нуклон.

Взаимодействие нуклонов в ядре можно сопоставить с аналогичным взаимодействием атомов в кристаллических решетках металлов, где

существенную роль играют электроны как «передатчики взаимодействия».

Различие заключается в том, что в ядрах «передатчиками взаимодействия» между нуклонами являются более тяжелые частицы — пи-мезоны (или пионы), масса которых в 273 раза больше массы электрона. Полагают, что нуклоны непрерывно порождают и поглощают пи-мезоны по схеме

$$p \rightleftarrows n + \pi^+; \quad n \rightleftarrows p + \pi^-,$$

так что каждый нуклон окружен облаком виртуальных пи-мезонов. Внутри ядра, где частицы находятся на относительно малых расстояниях друг от друга, пи-мезонное облако активно участвует в ядерных процессах, обусловливая взаимодействие и взаимные превращения нуклонов.

ОСНОВНЫЕ ХАРАКТЕРИСТИКИ ЯДЕР

Важнейшими характеристиками атомных ядер являются:

1) заряд (Ze),
2) масса (M),
3) дефект массы (ΔM) или энергия связи ($E = c^2\Delta M$),
4) механический момент (спин S),
5) магнитный момент (μ),
6) электрический квадрупольный момент (Q),
7) спектр возбужденных состояний (т. е. значения энергии связи для всех состояний, в которых может находиться ядро данного состава).

Масса ядра M меньше суммарной массы протонов m_p и нейтронов m_n, измеренных в свободном состоянии, вне ядра. Разность

$$Zm_p + Nm_n - M = \Delta M$$

называется *дефектом массы ядра*. Более удобно определение ΔM по массе нуклида $M_{\text{нук}}$.

$$\Delta M = ZM_{\text{H}} + Nm_n - M_{\text{нук}}, \tag{4.4}$$

где M_{H} — масса атома водорода. Существование дефекта масс ядер является иллюстрацией общего положения, согласно которому масса физической системы всегда меньше суммарной массы ее составных частей (измеренных вне системы). Ввиду этого масса нуклида также несколько меньше суммы масс ядра и Z электронов. Дефект массы данной системы (ΔM), умноженный на квадрат скорости света (c^2), равен той энергии, которая выделилась при образовании этой системы или необходима для ее разрушения на составные элементы:

$$\Delta Mc^2 = E.$$

У атомных ядер эта «энергия связи», как правило, увеличивается с ростом числа нуклонов, но есть и отступление от этого правила, например: изотоп титана $A = 49$ имеет $E = 426{,}844$ МэВ, а изотоп марганца $A = 50$ имеет несколько меньшую энергию связи $E = 426{,}659$ МэВ.

Прочность ядра принято характеризовать удельной (средней) энергией связи, приходящейся на один нуклон:

$$E_1 = \frac{E}{Z+N} = c^2 \frac{\Delta M}{A} = c^2 \Delta M_1.$$

$\Delta M/A = \Delta M_1$ — есть среднее уменьшение массы одного нуклона при формировании данного ядра из свободных протонов и нейтронов. При увеличении числа нуклонов в ядре E_1 изменяется, но не монотонно. Рассмотрим таблицу стабильных изотопов кальция ($Z = 20$; $N = 20 \div 28$) и ряд ядер, имеющих одинаковое число нейтронов $N = 20$. В этой таблице указаны:

1) содержание данного изотопа в естественном элементе c (в процентах),

2) полная энергия связи ядра E (в МэВ),

3) удельная энергия связи E_1,

4) изменение энергии связи ΔE при добавлении к предыдущему ядру одного нейтрона или одного протона.

I. Ядра с $Z = 20$ (присоединяются нейтроны)

$N = 20$	$c = 96{,}97\%$	$E = 342{,}056$	$E_1 = 8{,}551$	$\Delta E = 8{,}364$
21	—	350,420	8,547	11,471
22	0,64	361,891	8,617	7,928
23	0,145	369,819	8,60	11,135
24	2,06	280,954	8,658	7,420
25	—	388,374	8,63	10,401
26	0,0033	398,775	8,669	7,281
27	—	406,056	8,6395	9,940
28	0,185	415,996	8,667	—

II. Ядра с $N = 20$ (присоединяются протоны)

$Z = 16$	$c = 0{,}017\%$	$E = 308{,}707$	$E_1 = 8{,}575$	$\Delta E = 8{,}399$
17	24,6	317,106	8,570	10,243
18	9,063	327,349	8,614	6,374
19	93,08	333,723	8,557	8,333
20	96,97	342,056	8,551	—

Эти таблицы показывают, что при присоединении нуклонов к четно-четным ядрам изменение энергии связи заметно меньше по сравнению с другими ядрами. Заметим также, что в семействе ядер с $Z = 20$ основной изотоп кальция с $N = 20$ обладает меньшей удельной энергией связи E_1, чем другие, менее распространенные изотопы; более того, у этого стабильного ядра удельная энергия связи меньше, чем у радиоактивных изотопов с $N = 25$ и $N = 27$. В семействе ядер с $N = 20$ изотопы кальция $Z = 20$ и калия $Z = 19$, выделяющиеся своей распространенностью, также имеют меньшую удельную энергию связи по сравнению с менее распространенными ядрами серы $Z = 16$ и аргона $Z = 18$. То же самое можно отметить и для других значений Z и N. Поэтому удельная энергия связи, по-видимому, не определяет ни устойчивости ядра, ни содержания данного изотопа в естественном элементе.

На рис. IV.101 показаны значения энергии связи одного нуклона для ядер с различным числом нуклонов.

Из таблицы для изотопов кальция $Z = 20$ видно, что присоединение первого нейтрона к изотопу с $N = 20$ дает увеличение энергии связи на $\Delta E_1 = 8{,}364$ МэВ, а присоединение второго нейтрона вызывает несколько большее изменение энергии: $\Delta E_2 = 11{,}47$ МэВ. Разность этих значений

$$E_n = \Delta E_2 - \Delta E_1$$

можно рассматривать как энергию, которая дополнительно выделяется при соединении обоих нейтронов внутри ядра в особую систему — «пару нейтронов». Предполагается, что такие «пары» в структуре ядра представляют собой отдельную составную часть. Энергия образования «пары нейтронов» внутри данного ядра рассматривается как новая энергетическая характеристика этого ядра. Предполагается, что протоны внутри ядра также соединяются в «пары». Энергию, выделяющуюся при образовании внутри ядра «пары протонов», можно вычислить из энергии связи соседних ядер с $N = \text{const}$. Например, для $N = 20$ и $Z = 16$, 17 и 20 можно получить энергию образования «пары протонов» $\Delta E_p = 1{,}844$ МэВ. Однако величины ΔE_n и ΔE_p, рассчитанные, например, для четно-четных ядер, не обнаруживают определенной зависимости от Z, N или A и, по-видимому, представляют собой такие же усредненные характеристики ядра, какими являются удельная энергия E_1 и энергия присоединения нуклона ΔE.

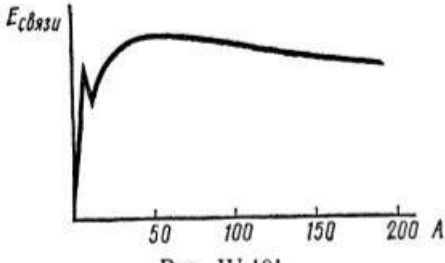


Рис. IV.101

МЕХАНИЧЕСКИЙ МОМЕНТ ЯДРА

Механический момент (момент импульса или спин) атомных ядер определяется их составом и внутренней структурой и выражается через постоянную Планка. В квантовой физике показано, что механический момент элементарных частиц и образованных ими систем, так же как и энергия этих систем, имеет дискретный спектр устойчивых («собственных») значений. Формула для энергии атома (см. ч. IV, § 14) имеет вид

$$E = -\frac{Z^2 me^4}{n^2 \cdot 8\varepsilon_0^2 h^2},$$

где $n = 1, 2, 3, \ldots$ — квантовое число, определяющее спектр дискретных значений энергии. Для модуля механического момента получена формула

$$|L| = \sqrt{l(l+1)}\,\frac{h}{2\pi}, \qquad (4.5)$$

где $l = 0, 1, 2, \ldots$ — квантовое число, определяющее спектр дискретных значений L. Эта формула применима как для орбитального,

так и для собственного (с п и н о в о г о) механического момента частицы; спиновое квантовое число обозначается через s, поэтому

$$|\mathbf{S}| = \sqrt{s(s+1)}\,\frac{h}{2\pi}. \qquad (4.6)$$

Если заданы l или s, то число возможных значений механического момента будет равно $2l + 1$ (или $2s + 1$).

С механическим моментом частиц (орбитальным или спиновым) неразрывно связано существование и м а г н и т н о г о момента (см. ч. III, § 20). Ниже будет приведена связь между спиновыми механическим и магнитным моментами частиц. Здесь же мы отметим важное обстоятельство, вытекающее из этой связи.

Допустим, что частица, имеющая магнитный момент μ, введена во внешнее магнитное поле B. При этом частица приобретает дополнительную кинетическую энергию $W = \mu B \cos\alpha$ (см. ч. III, § 20) и ее ось (которую ориентируем по направлению вектора магнитного момента) будет прецессировать с некоторой частотой вокруг направления поля. Энергия W зависит от угла α между направлениями векторов μ и B, однако, ввиду того что эта энергия может иметь только дискретный набор значений, угол α не может быть любым, а также должен иметь определенные значения («пространственное квантование»). Следовательно, проекция вектора μ на направление внешнего магнитного поля B может иметь только определенные значения. Так как вектор μ неразрывно связан с вектором орбитального $\mathbf{L}_{\text{ор}}$ или спинового $\mathbf{S}$ моментов, то и проекции механического момента на направление поля также должны иметь дискретный набор значений, включая максимальные значения при их параллельности и антипараллельности. Спектр значений вектора механического момента определяется квантовым числом l или s, поэтому число допустимых проекций этого вектора на направление магнитного поля также будет определяться квантовыми числами l и s.

Были поставлены опыты для определения числа возможных ориентаций μ относительно B. В опытах Штерна и Герлаха пучок атомов серебра (у которых суммарный орбитальный момент равен нулю и поэтому магнитный момент определяется только о д н и м электроном) пропускался через неоднородное магнитное поле, в котором частицы с различными ориентациями μ относительно B отклоняются в различных направлениях. Опыт показал расщепление пучка только на две части, т. е. магнитный момент электрона может иметь только две допустимые ориентации относительно B. Приравнивая $2s + 1 = 2$, получим $s = 1/2$, т. е. квантовое число для спина электрона равно 1/2. Само же значение механического момента (спина) электрона будет равно

$$|S| = \sqrt{\frac{1}{2}\left(\frac{1}{2}+1\right)}\,\frac{h}{2\pi} = \frac{\sqrt{3}}{2}\,\frac{h}{2\pi}.$$

Согласно установившейся терминологии, под спином частицы или атомного ядра понимается не модуль вектора механического момента S, а только квантовое число s; зная это число, можно не только вычислить

действительное значение механического момента частицы и ядра по формуле (4.6), но и число возможных проекций этого момента на направление внешнего магнитного поля, равное $2s + 1$.

Спиновые квантовые числа (или просто спины) протонов и нейтронов также оказались равными половине. Спины ядер составляются из орбитальных моментов вращающихся нуклонов и их собственных моментов, причем взаимная ориентация этих векторов (от которой зависит энергия взаимодействия между нуклонами) не может быть произвольной. Для спина ядер важное значение имеет четность или нечетность чисел Z и N:

1) у четно-четных ядер спины в основных состояниях равны нулю, а в возбужденных — имеют целочисленные значения: 0, 1, 2, 3, ...;

2) у ядер с нечетным значением общего числа нуклонов $A = Z + N$ спины в основных и возбужденных состояниях имеют значения 1/2, 3/2, 5/2 и т. д.

3) у нечетно-нечетных ядер спины имеют целочисленные значения: 0, 1, 2, ..., 10.

Для некоторых ядер указываются два разных значения спина, например:

Изотоп кальция	($Z=20$, $N=25$)	$S=5/2$ и $7/2$
Изотоп марганца	($Z=25$, $N=29$)	$S=3$ и 2

У некоторых ядер присоединение одного протона или нейтрона изменяет спин на 1/2, например:

Изотопы кремния	$Z=14$,	$N=14$	имеют	$S=0$
		$N=15$	»	$S=1/2$
		$N=16$	»	$S=0$
Фосфор	$Z=15$,	$N=16$	»	$S=1/2$

У других ядер присоединение нуклона сопровождается сильным изменением спина, например:

Изотопы титана	$Z=22$,	$N=26$	имеют	$S=0$
		$N=27$	»	$S=7/2$
		$N=28$	»	$S=0$
Ванадий	$Z=23$,	$N=27$	»	$S=6$

МАГНИТНЫЕ МОМЕНТЫ ЯДЕР

Магнитные моменты атомных ядер имеют различные значения, которые, однако, не связаны простыми соотношениями с механическим моментом. Теоретически показано, что частица, обладающая электрическим зарядом e и массой m, должна иметь собственный магнитный момент μ_0, равный

$$\mu_0 = \frac{1}{2}\frac{e}{m}\frac{h}{2\pi}. \tag{4.7}$$

В этой формуле величина $\frac{1}{2}\frac{h}{2\pi} = s_0$ определяется спином частицы S, поэтому

$$\boldsymbol{\mu}_0 = \frac{e}{m}\mathbf{S}_0. \tag{4.8}$$

Если заряд частицы положительный, то векторы магнитного и механического моментов имеют одинаковые направления: в этом случае μ_0 считают положительным. При отрицательном заряде частицы направления векторов μ_0 и S_0 противоположны и μ_0 считают отрицательным.

Если в формулу (4.7) подставить заряд и массу покоя электрона, то получим величину

$$\mu_B = 9{,}274 \cdot 10^{-24}\ \text{А} \cdot \text{м}^2,$$

называемую *магнетоном Бора*. Однако действительное (измеренное) значение магнитного момента электрона оказывается равным

$$\mu_e = 9{,}284 \cdot 10^{-24}\ \text{А} \cdot \text{м}^2.$$

Еще большее расхождение между формулой (4.7) и результатами измерений обнаруживается у протона. Если в формулу (4.7) подставить заряд и массу протона, то получается величина

$$\mu_N = 5{,}05 \cdot 10^{-27}\ \text{А} \cdot \text{м}^2,$$

называемая *ядерным магнетоном*. Измеренное значение магнитного момента протона оказалось равным

$$\mu_p = 1{,}41 \cdot 10^{-26}\ \text{А} \cdot \text{м}^2,$$

т. е. в 2,793 раза больше.

Нейтральные частицы не должны иметь магнитного момента, тогда как измерения показали, что нейтрон имеет магнитный момент

$$\mu_n = -1{,}9128\,\mu_N.$$

Знак минус показывает, что вектор магнитного момента направлен против вектора механического момента, что характерно для отрицательно заряженных частиц, поэтому полагают, что нейтрон, будучи в целом электрически нейтрален, имеет в своей структуре заряды противоположных знаков, расположенные в его объеме таким образом, что магнитные моменты, связанные с отрицательными зарядами, оказываются несколько больше магнитного момента, связанного с положительными зарядами.

Аномальные значения магнитных моментов протона и нейтрона объясняются существованием указанного выше пи-мезонного облака, окружающего их. Пи-мезоны имеют магнитный момент $\approx 6{,}6\mu_N$, поэтому если определение магнитного момента протона или нейтрона в измерительной аппаратуре длится некоторое время Δt, то появляющиеся и исчезающие пи-мезоны могут внести некоторый вклад в результаты измерений; эти результаты будут статистически средними за время Δt.

Таким образом, магнитные моменты ядер не могут быть представлены в виде векторной суммы собственных (спиновых) магнитных моментов составляющих их нуклонов (которые вместе с механическими моментами должны быть или параллельны, или антипараллельны между собой); некоторый вклад могут внести и орбитальные магнитные моменты вращающихся зарядов (протонов).

Приведем магнитные моменты некоторых ядер:

Гелий	$A=3$	$\mu=+2{,}1274\ \mu_p$
Азот	14	$+0{,}40$
Азот	15	$-0{,}28$
Фтор	19	$+2{,}63$
Магний	25	$-0{,}85$

Обнаружена только одна закономерность: магнитные моменты всех четно-четных ядер равны н у л ю. Так как у таких ядер равны нулю и спины, то можно утверждать, что в структуре четно-четных ядер протоны соединены в «пары» с противоположными ориентациями механического и, следовательно, магнитного моментов; такие же «пары» с нулевыми значениями обоих моментов образуют и нейтроны.

Рассмотрим интересный пример: у ядра «тяжелого водорода» — дейтрона, состоящего из протона и нейтрона, спин равен единице. Это означает, что в его структуре протон и нейтрон должны иметь одинаковые ориентации векторов механического момента (спины складываются: $1/2 + 1/2 = 1$) и, следовательно, противоположные ориентации магнитных моментов (у протона механический и магнитный моменты имеют одинаковые направления, у нейтрона — противоположные). Тогда полный магнитный момент дейтрона должен был бы равняться

$$\mu_d = \mu_p + \mu_n = (2{,}7928 - 1{,}9128)\,\mu_p = 0{,}880\,\mu_p;$$

измерения же показали, что $\mu_d = 0{,}857\mu_p$, т. е. меньше разности магнитных моментов протона и нейтрона. Это означает, что в ядре кроме собственных магнитных моментов нуклонов имеются дополнительные магнитные моменты (орбитальные и, возможно, другие), которые в данном случае направлены противоположно моменту протона.

Магнитные моменты ядер измеряются р е з о н а н с н ы м м е т о д о м (см. ч. III, § 25). Идея этого метода заключается в следующем: допустим, что некоторая частица (или система частиц — атом, атомное ядро) обладает магнитным моментом μ. Во внешнем магнитном поле B она приобретает дополнительную кинетическую энергию, равную μB, и прецессируют вокруг направления поля с некоторой частотой ν, пропорциональной μ. В таком состоянии частица может поглотить излучение с частотой ν и изменить ориентацию вектора μ относительно B на противоположную. Поэтому, если пропустить через вещество, содержащее множество таких прецессирующих частиц, излучение, частоту которого можно плавно изменять, то при $\nu_{\text{внеш}} = \nu_{\text{прец}}$ обнаруживается резкий (резонансный) максимум поглощения. (Частота ν оказывается в радиотехническом диапазоне.) Затем по значениям ν и B можно вычислить μ.

ЭЛЕКТРИЧЕСКИЕ МОМЕНТЫ ЯДЕР

Для атомных ядер можно найти не только «центр масс» (см. ч. I, § 14), но и «центр электрических зарядов». Если x_i — координата i-го протона в данной системе координат, то соответствующая координата

X «центра зарядов» будет равна: $X = \frac{\sum ex_i}{\sum e} = \frac{\sum x_i}{Z}$; такие же формулы дадут значения координат Y и Z.

В основном состоянии атомных ядер центр масс нуклонов совпадает с «центром зарядов», т. е. атомные ядра имеют общий «центр симметрии». При совпадении центров масс и зарядов электрический дипольный момент ядра равен нулю. Атомных ядер с отличным от нуля электрическим дипольным моментом не обнаружено.

При помощи весьма тонких измерений удается выяснить, является ли распределение электрических зарядов (протонов) внутри ядра сферически симметричным или иным. Такие измерения показали, что у некоторых ядер распределение протонов в структуре ядра не является сферически симметричным, но имеет ось симметрии и перпендикулярную к ней плоскость симметрии. Такие ядра имеют *электрический квадрупольный момент*. Если допустить, что ядро представляет собой эллипсоид вращения с постоянной по объему плотностью заряда, то квадрупольный момент будет равен

$$Q = \frac{2}{5} Ze(c^2 - a^2), \tag{4.9}$$

где Ze — заряд ядра; c и a — полуоси эллипсоида вдоль и перпендикулярно к оси симметрии. Электрический квадрупольный момент является положительным, если эллипсоид вытянут вдоль оси симметрии ($c > a$), и отрицательным, если $c < a$.

Электрические квадрупольные моменты обычно выражаются в единицах Q/e (см^2). Замечено, у всех магических ядер этот момент равен нулю, а у соседних с ними ядер отличен от нуля и имеет различные знаки. Большие значения электрического квадрупольного момента встречаются только у тяжелых ядер.

СПЕКТР ВОЗБУЖДЕННЫХ СОСТОЯНИЙ

Важной характеристикой ядер является *спектр возбужденных состояний*. Ядро данного состава Z, N может находиться, во-первых, в основном состоянии с наименьшим возможным значением энергии связи E_0 и, во-вторых, в некотором наборе других состояний с более высокими значениями энергии связи E. Совокупность всех возможных состояний данного ядра с энергиями связи $E > E_0$ есть спектр возбужденных состояний или *спектр ядерных уровней*.

Разность между энергиями возбужденного и основного состояний называется *энергией возбуждения*:

$$\Delta E_i = E_i - E_0,$$

где индекс i обозначает номер уровня.

Перевод ядра из основного состояния в возбужденное происходит:

1) при поглощении ядром гамма-фотонов;

2) при бомбардировке ядра различными частицами;

3) при электромагнитном взаимодействии между ядром и пролетающей мимо него заряженной частицей (протоном или другим ядром),

обладающей достаточно большой кинетической энергией («кулоновское возбуждение ядра»);

4) в некоторых процессах ядро данного состава появляется не в основном, а в возбужденном состоянии.

Например, при проникновении протонов с энергией 7,26 МэВ в ядра бериллия-9 образуются ядра лития-6 в трех состояниях: в основном и в двух возбужденных состояниях с энергиями возбуждения 2,184 и 3,560 МэВ.

Возбужденные состояния ядер являются неустойчивыми и ядро переходит в основное состояние (с минимально возможным значением энергии связи). Переход из возбужденного состояния в основное может осуществляться различными способами, в зависимости от энергии возбуждения и структуры данного ядра:

1) испусканием гамма-фотона;

2) испусканием частиц из ядра (протона, нейтрона, альфа-частицы). В этом случае энергия возбуждения должна быть больше энергии связи испускаемой частицы в ядре;

3) передачей энергии возбуждения одному из электронов, входящих в состав электронной оболочки атома (этот процесс называется «внутренней конверсией»).

Измерение энергии испускаемых частиц и фотонов позволяет определить спектр уровней данного ядра.

Время τ существования возбужденного состояния с энергией E связано с «энергетической шириной» этого уровня $\Delta E_i^?$ (заметим, что физические системы никогда не имеют строго определенного значения энергии; любое состояние системы можно характеризовать значением энергии, лежащей в узких пределах: $E_i \pm \frac{1}{2}\Delta E_i^?$). Согласно принципу Гейзенберга:

$$\tau \Delta E_i^? = h/2\pi.$$

Величины $\Delta E_i^?$ и τ для каждого уровня определяются составом ядра, характером взаимодействия между нуклонами и структурой ядра. Если $\Delta E_i^?$ равно или больше разности энергий между соседними уровнями, то спектр возбужденных состояний будет непрерывным.

Среднее время пребывания ядер в возбужденных состояниях, т. е. время, в течение которого первоначальное число одинаково возбужденных ядер уменьшается в $e = 2{,}718$ раз:

1) у легких ядер лежит в пределах $10^{-11} \div 10^{-14}$ с, т. е. значительно меньше времени существования возбужденных состояний в электронных оболочках атомов ($\sim 10^{-8}$с);

2) у тяжелых ядер — значительно больше и для состояний, близких к основному, достигает нескольких секунд, минут и даже часов.

Изучение спектров возбужденных состояний ядер приводит к следующим представлениям. Каждый нуклон внутри ядра может находиться только в некотором дискретном наборе состояний. Спектр возможных состояний для протонов может не совпадать с аналогичным спектром для нейтронов. Каждое из возможных состояний нуклона в ядре характеризуется определенным значением энергии связи;

поэтому можно утверждать, что в ядре нуклоны располагаются по определенным «уровням энергии». Эти уровни нумеруются по возрастающим значениям энергии. В основном состоянии ядра нуклоны занимают все низшие уровни энергии без пропусков. Если один из нуклонов оказывается на более высоком уровне, то ядро будет иметь в спектре низших уровней одно «вакантное место». Согласно принципу Паули, на каждом уровне для протонов находится только два протона с противоположно направленными спинами. На нейтронных уровнях могут находиться также два нейтрона с противоположными ориентировками спинов. Если уровни протонов и нейтронов совпадают, то на каждом из них могут находиться только четыре нуклона.

Полный спектр всевозможных состояний данного ядра можно условно разделить на следующие части:

1) область низших уровней. В основном состоянии нуклоны данного ядра полностью занимают все эти уровни;

2) область возбужденных уровней, у которых энергия возбуждения меньше энергии, необходимой для удаления нуклона из ядра. В этом случае переход ядра из возбужденного состояния в основное может осуществляться путем излучения гамма-фотона: такие уровни называются также «стационарными»;

3) область возбужденных уровней, у которых энергия возбуждения больше энергии связи одного нуклона в ядре. Эти уровни обнаруживаются по резонансным явлениям в ядерных реакциях, происходящих при различных внешних воздействиях на ядро (при изменении энергии бомбардирующих частиц и т. д.).

Установлены следующие закономерности:

1) легкие ядра имеют небольшое число уровней возбужденных состояний, а разность энергий между ними велика и вблизи основного состояния достигает нескольких МэВ. С увеличением числа нуклонов в ядре общее число уровней увеличивается, а разности энергий между ними уменьшаются;

2) при больших значениях энергии возбуждения ядерные уровни расположены ближе друг к другу, т. е. их «плотность» («густота» или число уровней, приходящихся на единичный интервал значений энергии) возрастает;

3) в окрестности определенных значений энергии возбуждения плотность ядерных уровней у четно-нечетных ядер больше, чем у четно-четных;

4) разности энергий уровней у ядер с магическими числами Z и N больше, чем у соседних ядер (в одной и той же области спектра возбуждений);

5) спектры уровней энергии у пары «зеркальных» ядер похожи. Например, энергии возбуждения у ядер $^{11}_{5}B_6$ и $^{11}_{6}C_5$ имеют следующие значения, МэВ:

$^{11}_{5}B_6$ . . .	2,1	4,46	5,30	6,80	7,30
$^{11}_{6}C_5$ . . .	1,95	4,30	4,85	6,48	6,83

При общей характеристике возбужденных состояний ядер, кроме энергии связи E (или энергии возбуждения) и времени существования,

необходимо указать также спин ядра в каждом из этих состояний. Изменение спина при переходе из основного состояния в возбужденное или из одного возбужденного состояния в другое является некоторой характеристикой внутренних изменений в ядрах при этих переходах.

Возбуждение ядер осуществляется сообщением им энергии извне. Если эта энергия достаточно велика, то можно перевести нуклон из его уровня в структуре ядра на уровень, лежащий за пределами объема ядра. Так, например, при поглощении ядром гамма-фотонов большой энергии наблюдается выбрасывание из ядра протона, нейтрона или альфа-частицы. Это явление называется *фоторасщеплением ядра* или «ядерным фотоэффектом».

Такими процессами являются, например, фоторасщепление дейтрона

$$^{2}_{1}H_1 + \gamma \rightarrow {}^{1}_{1}H_0 + n; \quad {}^{2}_{1}H_1 (\gamma, n) {}^{1}_{1}H_0$$

(справа дана условная запись этой реакции; в скобках записаны символы падающей и излучаемой частиц) или реакции

$$^{71}_{31}Ga_{40} (\gamma, n) {}^{70}_{31}Ga_{39}; \quad {}^{6}_{3}Li_3 (\gamma, p) {}^{3}_{2}He_3.$$

При помощи гамма-фотонов можно вызвать и разложение ядра, например

$$^{16}_{8}O_8 (\gamma, \alpha) {}^{12}_{6}C_6.$$

В этой реакции ядро кислорода-16 разлагается на ядро углерода-12 и ядро гелия-4.

§ 20. МОДЕЛЬНЫЕ ПРЕДСТАВЛЕНИЯ О СТРУКТУРЕ ЯДЕР

ОСНОВНЫЕ ПРЕДПОЛОЖЕНИЯ

Для объяснения свойств ядер и ядерных процессов используются различные «модельные» представления о структуре ядер; к ним относятся: «капельная», «оболочечная», «коллективная», «оптическая» и другие модели. Каждая из предлагаемых моделей ядра содержит ряд упрощающих предположений, благодаря которым удается рассчитать наблюдаемые свойства ядер, объяснить течение известных ядерных реакций и предсказать новые. Перечислим важнейшие из таких предположений:

1) внутри ядра протоны и нейтроны сохраняют свои отличительные («индивидуальные») свойства. Ядро представляется как система взаимосвязанных нуклонов, в которой возможны различные изменения в их относительном расположении и характере движения;

2) взаимные превращения между нуклонами:

$$n \rightarrow p + e^- + \bar{\nu}, \quad p \rightarrow n + e^+ + \nu,$$

при которых из ядра вылетают электрон или позитрон и нейтрино, происходят только в особых условиях, возникающих в радиоактивных ядрах. Однако взаимные превращения другого характера, которые сопровождаются обменом π-мезонами, происходят в ядре постоянно и

обусловливают ядерное взаимодействие между нуклонами («мезонная теория ядерных сил»);

3) каждый нуклон движется внутри ядра в некотором общем (усредненном) ядерном поле, которое создано остальными нуклонами. Ввиду этого все нуклоны в ядре будут находиться в одинаковых условиях. Однако эти условия на поверхности ядра несколько отличны от условий, существующих внутри объема ядра;

4) движение отдельного нуклона в ядре не зависит от движения остальных. В частности, если данный нуклон может перемещаться в ядре по каким-то «дозволенным» орбитам, то переход этого нуклона из одной орбиты на другую может осуществляться независимо от того, в каких состояниях (на каких орбитах) находятся остальные нуклоны. В некоторых случаях такая зависимость допускается и оценивается;

5) ядерное взаимодействие между протоном и нейтроном, протоном и протоном, нейтроном и нейтроном одинаковое. Такой универсальный характер ядерного взаимодействия (т. е. независимость взаимодействия от наличия или отсутствия у нуклона электрического заряда) рассматривается, как необходимое следствие из некоторых опытных данных;

6) полная энергия взаимодействия между нуклонами может быть представлена, как сумма энергий взаимодействий между каждой парой нуклонов. Это предположение сохраняется и в тех моделях, в которых ядерное взаимодействие происходит только между соседними («соприкасающимися») нуклонами;

7) энергия ядра не зависит от формы его поверхности. В некоторых моделях такая зависимость допускается и дается ее оценка. Энергия нуклонов, расположенных на поверхности ядра, может несколько отличаться от энергии нуклонов, находящихся внутри;

8) ядро имеет форму шара, в пределах которого нуклоны расположены равномерно (с одинаковой по всему объему ядра плотностью). В некоторых моделях предполагается, что ядра имеют другие формы: эллипсоида вращения, трехосного эллипсоида и др.; при этом допускаются вращение ядра и колебания его формы;

9) сложные ядра представляют собой систему из более простых ядер. В частности, четно-четные ядра рассматриваются как системы, состоящие из альфа-частиц («альфа-частичная модель»), которые существуют и движутся в ядре как самостоятельные частицы. Допускается, что эти альфа-частицы внутри ядра могут распадаться на протоны и нейтроны и вновь собираться из них; важно, чтобы время пребывания в собранном состоянии было значительно больше, чем в разложенном.

Это представление может быть расширено (для тяжелых ядер), если в качестве структурных частей рассматривать не только альфа-частицы, т. е. ядра гелия, но и ядра других легких элементов — лития, бериллия, бора, углерода и т. д.;

10) протоны расположены в пределах объема ядра равномерно. Это предположение применимо для ядер, содержащих большое число протонов и используется для расчета потенциальной энергии их электростатического взаимодействия. Однако имеются экспериментальные данные, которые рассматриваются как доказательство неравно-

мерного распределения электрического заряда по объему ядра (плотность заряда убывает от центра ядра к поверхности);

11) ядерное поле описывается при помощи определенной функции, выражающей зависимость потенциала ядерных сил от расстояния до центра ядра (если предполагается, что поле сферически симметричное) или до оси ядра (если ядро имеет, например, форму эллипсоида вращения). Выбирая конкретную форму этой зависимости, можно попытаться объяснить свойства ядер и их взаимодействия с бомбардирующими частицами. Одно из таких предположений позволило разработать теорию ядра, в которой нуклоны вращаются по особым устойчивым орбитам, причем оказалось, что состояние нуклона при параллельной ориентации собственного спина и орбитального момента отличается от состояния с антипараллельной ориентацией. Это приводит к важному следствию: каждый нуклонный уровень энергии в ядре расщепляется на два уровня с различными значениями энергии («спин — орбитальное расщепление уровней»).

Кроме перечисленных выше основных предположений, которые используются в той или иной модели, пользуются и другими предположениями, позволяющими объяснить те или иные экспериментальные данные с необходимой точностью. Изложение этих деталей приводится в специальных монографиях. Ниже дадим краткое изложение некоторых сравнительно простых ядерных моделей.

КАПЕЛЬНАЯ МОДЕЛЬ ЯДРА

В «капельной модели», предложенной Н. Бором, ядро рассматривается как система нуклонов, упакованных под действием ядерных сил притяжения с предельно возможной плотностью, наподобие капли жидкости. Каждая частица взаимодействует только с ближайшими соседями, с которыми она «соприкасается». При большом числе нуклонов ядро (в равновесном или основном состоянии) должно иметь форму шара с равномерным распределением нуклонов по объему. При предельно плотной упаковке объем ядра будет пропорционален числу нуклонов, что находит удовлетворительное подтверждение в пропорциональности между радиусом ядра и кубичным корнем из числа нуклонов.

Капельная модель ядра позволила получить полуэмпирическую формулу для вычисления энергии связи, предложенную Вейцзеккером. Потенциальная энергия взаимодействия нуклонов представляется в виде суммы энергий:

$$E = (E_{об} - E_{пов}) + E_{кул} + E_{изб} + E_{пар}.$$

Так как внутри объема ядра каждый отдельный нуклон взаимодействует с одинаковым числом соседей, то можно полагать, что эти нуклоны вносят одинаковые вклады в полную энергию ядра. Следовательно, какая-то часть энергии ядра должна быть пропорциональна числу нуклонов, а следовательно, и объему ядра; эта часть и обозначена выше через $E_{об}$, поэтому

$$E_{об} = c_1 A,$$

где c_1 — вклад одного нуклона. Однако нуклоны, расположенные на поверхности ядра, имеют соседей только с одной стороны. Поэтому из $E_{об}$ следует вычесть некоторую энергию $E_{пов}$, пропорциональную числу нуклонов на поверхности ядра. Можно полагать, что число нуклонов, находящихся на периферии ядра, пропорционально поверхности ядра $S = 4\pi r^2$. Так как радиус ядра $r = r_0 A^{1/3}$, то $S = 4\pi r_0^2 A^{2/3}$. Таким образом, энергия $E_{пов}$ будет пропорциональна $A^{2/3}$:

$$E_{пов} = c_2 A^{2/3}.$$

Вследствие кулоновского отталкивания положительно заряженных протонов ядро будет иметь энергию электростатического отталкивания $E_{кул}$. Эта энергия будет зависеть от числа протонов в ядре и их взаимного расположения. Полагая, что электрический заряд ядра $q = Ze$ распределен равномерно по объему шара радиуса r, можно написать:

$$E_{кул} = \frac{3}{5}\frac{q^2}{r} = \frac{3}{5}\frac{Z^2e^2}{r_0 A^{1/3}} = c_3 Z^2 A^{-1/3}.$$

(Напомним, что потенциальные энергии взаимодействия при отталкивании и притяжении имеют противоположные знаки, поэтому энергия кулоновского отталкивания имеет знак, противоположный энергии ядерного притяжения.)

Последние два члена выражения для энергии связи означают: энергию $E_{изб}$, зависящую от избытка нейтронов над протонами $N - Z$, и энергию $E_{пар}$, обусловленную особыми силами «спаривания», благодаря которым в структуре ядра образуются «пары протонов» и «пары нейтронов». Обе эти энергии также имеют знаки, противоположные знаку энергии ядерного притяжения.

Из условия: E равно минимуму при постоянном A, можно найти то значение Z, которое соответствует наиболее прочному ядру среди других ядер с таким же числом нуклонов.

Расчет ($dE/dZ = 0$ при $A = \text{const}$) приводит к формуле

$$Z \approx \frac{A}{2 + 0{,}014\, A^{2/3}}. \qquad (4.10)$$

В частности для $A = 10$ получается $Z = 5$ (ядро бора $^{10}_{5}B_5$)
$A = 100$ » $Z = 44$ (ядро рутения $^{100}_{44}Ru_{56}$)
$A = 200$ » $Z = 80$ (ядро ртути $^{200}_{80}Hg_{120}$)

и т. д.

Капельная модель допускает возможность колебания формы ядра (при постоянном объеме). Эти колебания могут привести к делению ядра-капли на части. Условие, при котором такое ядро сохраняет свою устойчивость, связывает между собой число протонов и общее число нуклонов в ядре (оно было получено Я. И. Френкелем)

$$Z^2 < 49\, A.$$

Из этого соотношения следует, что при данном числе нуклонов в ядре число протонов не может превышать некоторой «нормы».

Упомянем еще одно следствие, к которому приводит «капельная модель». При попадании в ядро бомбардирующей частицы (протона, нейтрона и т. д.) ее кинетическая энергия в течение очень короткого времени должна перераспределяться между всеми нуклонами ядра. Если новое ядро испытывает распад, то его называют «промежуточным», а испускание частицы при распаде рассматривают как процесс, эквивалентный испарению молекулы с поверхности жидкой капли. Очевидно, что этот распад не должен зависеть от того, каким образом было получено промежуточное ядро; в частности, направление скорости испускаемых частиц не должно быть связано с направлением скорости ранее влетевшей частицы. Эта гипотеза о независимости распада промежуточного ядра от способа его образования, обоснованная Н. Бором, получила ряд экспериментальных подтверждений, однако наблюдались и некоторые отклонения.

ОБОЛОЧЕЧНАЯ МОДЕЛЬ

В «оболочечной модели» ядра предполагается, что нуклоны создают общее ядерное поле, которое имеет центр и поэтому действует на каждый нуклон с силой, зависящей от расстояния до этого центра. Далее предполагается, что нуклоны движутся в общем ядерном поле по особым устойчивым орбитам, почти независимо друг от друга, и на каждой орбите имеют определенные механические и магнитные моменты (собственный и орбитальный). Это предположение требует специального обоснования. Дело в том, что ядерное взаимодействие очень сильное, размеры ядра малы, а плотность частиц в единице объема n велика:

$$n = \frac{A}{(4/3)\,\pi r^3} = \frac{3}{4\pi r_0^3} \approx 0{,}24 \cdot 10^{39} \text{ см}^{-3}.$$

Ввиду этого длина свободного пробега нуклона в объеме ядра, равная

$$l = \frac{1}{\pi\sigma^2 n} = \frac{4}{3}\,\frac{r_0^3}{\sigma^2}$$

(где σ — эффективный диаметр нуклона), оказывается соизмеримой с размерами ядра. При энергии около 20 МэВ эксперименты дают для σ значение $\approx 4{,}5 \cdot 10^{-13}$ см, поэтому

$$l \approx \frac{2}{9} \cdot 10^{-13} \text{ см}.$$

При этих условиях вероятность столкновений между нуклонами (сопровождающихся обменом энергиями) будет велика и понятие нуклонной орбиты теряет смысл.

Однако следует учесть два обстоятельства:

1) уровни энергии в ядре (соответствующие определенным нуклонным орбитам) заполняются последовательно, начиная с наинизших. Поэтому в невозбужденном состоянии ядра все A низших уровней энергии заняты A нуклонами;

2) согласно принципу Паули, на каждом уровне могут находиться только два нуклона одного сорта с противоположно направленными спинами.

Ввиду этих условий, столкновения между нуклонами, в результате которых происходил бы обмен энергиями, исключаются: нуклон, если он отдаст энергию, не сможет перейти на низший уровень, так как все они «заняты» (имеются в виду невозбужденные состояния ядер). Таким образом, каждый нуклон будет двигаться в ядре, не имея возможности изменить свою энергию. Флуктуации ядерного поля, вызванные движением нуклонов, по той же причине не могут вызвать существенного изменения энергии каждого отдельного нуклона. Траектория движения нуклонов внутри ядра будет достаточно «гладкой» и можно говорить о существовании нуклонных орбит.

Для оболочечной модели ядра, на основании определенного предположения о характере ядерного поля, получена совокупность устойчивых уровней энергии $E(n, l)$, где l и n — квантовые числа: $n = 1, 2, 3, \ldots$ означает номер уровня, начиная с наинизшего, l выражает орбитальный момент импульса нуклона. Однако вследствие упомянутого выше «спин-орбитального расщепления» каждый из этих уровней распадается на два уровня: при параллельной ориентации собственного спина нуклона ($\pm 1/2$) и орбитального момента l (когда полный момент равен $j = l + 1/2$) энергия уровня оказывается меньше, чем при антипараллельной ориентации (когда $j = l - 1/2$). Разность энергий между этими уровнями пропорциональна величине $(2l+1)\ A^{-2/3}$, т. е. возрастает с увеличением орбитального момента и убывает при переходе к тяжелым ядрам. С учетом этого расщепления удалось установить, что уровни энергии разделены на отдельные группы, названные «оболочками». (Разность энергии между уровнями одной оболочки мала, по сравнению с разностями энергии между уровнями, принадлежащими соседним оболочкам.) Выяснилось, что ядра с магическими значениями Z и N имеют заполненные оболочки: у $^{4}_{2}\mathrm{He}_{2}$ заполнена первая оболочка; у $^{16}_{8}\mathrm{O}_{8}$ — первая и вторая; у $^{40}_{20}\mathrm{Ca}_{20}$ — три оболочки и т. д.

В ядрах с заполненными оболочками механические и магнитные моменты нуклонов оказываются скомпенсированными. Если же сверх заполненной оболочки имеется один нуклон (например, у $^{17}_{8}\mathrm{O}_{9}$) или же оболочка не заполнена и имеется одна вакансия, то, очевидно, механический и магнитный моменты ядра будут отличны от нуля и обусловлены лишним или недостающимся нуклоном. Это следствие подтверждается для некоторых легких ядер.

Для объяснения известных свойств различных ядер в оболочечной модели также используются дополнительные предположения, уточняющие характер взаимодействия нуклонов в ядре, расположение уровней, формы орбит и т. д.

ОБОБЩЕННАЯ МОДЕЛЬ

Более общей, чем капельная и оболочечная, является «обобщенная модель» ядра, предложенная О. Бором. В этой модели учитывается деформация оболочечной структуры в тех ядрах, в которых оболочки не заполнены нуклонами. Эта деформация зависит от числа нуклонов в незаполненных оболочках и поэтому у магических ядер отсутствует.

В первом приближении можно полагать, что деформация оболочечной структуры ядра означает превращение сферической формы в эллипсоид вращения. Вследствие этого у ядра появится электрический квадрупольный момент; кроме того, ядро может вращаться вокруг оси, перпендикулярной к оси эллипсоида. Однако это вращательное движение деформированного ядра не аналогично вращению твердого тела, так как между нуклонами нет жесткой связи. Под вращением ядра понимается изменение его формы, которое в виде некоторой «поверхностной волны» будет перемещаться вокруг оси вращения; иногда говорят, что нуклоны совершают «коллективное вращательное движение».

Обобщенная модель используется также для расчета электрических квадрупольных и магнитных моментов ядер.

ОПТИЧЕСКАЯ МОДЕЛЬ

«Оптическая модель» ядра специально приспособлена для описания взаимодействия ядер с налетающими на них частицами. Здесь существенно, что длина дебройлевской волны бомбардирующих частиц большой энергии соизмерима с размерами ядер-мишеней. У протона при энергии $E = 20$ МэВ длина этой волны будет равна

$$\lambda = \frac{h}{mv} = \frac{h}{\sqrt{2mE}} \approx 6{,}4 \cdot 10^{-13} \text{ см},$$

что соответствует радиусу ядра, содержащего $A = 80$ нуклонов. При такой ситуации ядро можно представить в виде сферы, обладающей по отношению к падающему волновому пакету частицы определенными «оптическими» свойствами. При взаимодействии набегающей волны с таким шаром будут происходить отражение, преломление, поглощение и дифракция. Соответственно налетающая частица будет иметь определенные вероятности отражения от ядра, проникновения внутрь, поглощения или рассеяния по различным направлениям. Вследствие дифракции распределение рассеянных частиц по направлениям будет иметь максимумы и минимумы; в некоторых экспериментах эта дифракционная картина отчетливо наблюдается.

Каждая из перечисленных выше ядерных моделей позволяет удовлетворительным образом охватить лишь часть экспериментальных данных. Некоторые модели оказываются полезными для объяснения свойств только легких ядер, другие — только тяжелых. Оболочечная модель позволила понять происхождение магических значений чисел протонов и нейтронов; в капельной модели этого не сделано, но она оказалась полезной для объяснения деления тяжелых ядер и т. д. Одной модели, которая дала бы возможность объяснить все известные свойства ядер (легких, тяжелых, стабильных, радиоактивных), не найдено. Такое положение в ядерной физике приводит к следующим выводам: либо ядра являются очень сложными системами (имеется в виду не только «сложность» самих нуклонов, но и сложный характер ядерных взаимодействий) и поэтому трудно найти для них аналогии среди известных и изученных физических систем, либо же внутри

ядра между нуклонами действуют силы неизвестной природы, для понимания которых нужны совершенно новые, пока еще не известные идеи и представления.

УСЛОВИЕ СТАБИЛЬНОСТИ

Важным вопросом ядерной физики является *выяснение причин, по которым ядро одного состава является стабильным, а другого — радиоактивным*. Для некоторых групп ядер последовательное добавление к стабильному ядру нескольких протонов или нейтронов приводит к образованию стабильных же ядер; например, ядра, имеющие:

$$N=20 \quad \text{и} \quad Z=16,\ 17,\ 18,\ 19,\ 20$$

или

$$Z=22 \quad \text{и} \quad N=24,\ 25,\ 26,\ 27,\ 28$$

являются стабильными. В других случаях добавление только одного протона или нейтрона приводит к образованию радиоактивного ядра. Существуют также стабильные ядра, которые могут быть получены из соседних стабильных ядер добавлением только «пары» протонов или нейтронов. Располагая ядра по возрастающим значениям чисел протонов и нейтронов (т. е. в «координатной системе» Z, N), можно заметить существование *особых значений* этих чисел: при $Z = 43$ и 61 стабильные ядра не образуются ни при каких значениях N; точно так же не существуют стабильные ядра при «особых» значениях $N =$ $= 19, 21, 35, 39, 45, 61, 89, 115$ и 123.

В настоящее время не удается найти способ вычисления стабильных сочетаний Z и N, но если их считать заданными, то можно определить характер радиоактивности соседних ядер. Для большинства ядер соблюдается следующее п р а в и л о:

если стабильное ядро сохраняет свою устойчивость при добавлении к нему нейтрона, то присоединение протона приводит к образованию ядра, испытывающего β^+*-распад или электронный захват. Если же ядро сохраняет стабильность при добавлении протона, то присоединение нейтрона образует ядро, испытывающее* β^-*-распад.*

Например, стабильное ядро $^{17}_{8}O_9$ при добавлении нейтрона образует устойчивое ядро $^{18}_{8}O_{10}$, но присоединение протона дает радиоактивное ядро $^{18}_{9}F_9$, испускающее позитрон. Ядро $^{18}_{8}O_{10}$ при присоединении протона образует стабильное ядро $^{19}_{9}F_{10}$, но добавление нейтрона приводит к β^--радиоактивному ядру $^{19}_{8}O_{11}$. Среди изученных ядер имеется весьма незначительное число отступлений от этого правила (отметим, что все они имеют указанные выше «особые» значения чисел Z и N).

В большинстве ядерных моделей, в соответствии с экспериментальными данными, предполагается, что каждое атомное ядро имеет характерный для него спектр уровней энергии, на которых могут находиться нуклоны, причем уровни для протонов отличаются от уровней для нейтронов. Предполагается, что это отличие заключается только в *значениях энергии уровней*. Однако возможно существование физических различий и более глубокого характера. Допуская их,

можно дать простое и образное объяснение указанному выше правилу. Если в данном ядре все протоны занимают «протонные уровни энергии», а все нейтроны находятся соответственно на «нейтронных уровнях», то это ядро будет стабильным. Однако ядерные силы могут (для достижения минимума потенциальной энергии или же из-за отсутствия «вакансий») втянуть протон на «нейтронный» уровень; на этом уровне протон не может долго существовать и со временем вынуждается к превращению в нейтрон. Точно так же, если нейтрон был вынужден занять «протонный» уровень, то со временем он превращается в протон. Ввиду этого, если в стабильном ядре произвести замену протона нейтроном или нейтрона протоном, то получится радиоактивное ядро. В качестве примера приведем две тройки ядер с такой заменой (стрелки указывают направление радиоактивного превращения):

$$^{15}_{6}C_{9} \rightarrow {}^{15}_{7}N_{8} \leftarrow {}^{15}_{8}O_{7}; \quad {}^{67}_{29}Cu_{38} \rightarrow {}^{67}_{30}Zn_{37} \leftarrow {}^{67}_{31}Ga_{36}.$$

В следующей тройке можно заметить, что ядра, обладающие одновременно β^+ и β^--активностью, имеют соседние стабильные ядра, отличающиеся тем, что в них протон заменен на нейтрон и обратно:

$$^{102}_{44}Ru_{58} \leftarrow {}^{102}_{45}Rh_{57} \rightarrow {}^{102}_{46}Pd_{56}.$$

Можно указать также на все пары «зеркальных ядер», у которых замена нейтрона протоном в стабильном компоненте всегда дает радиоактивное ядро, испускающее позитрон или испытывающее электронный захват, например:

$$^{7}_{3}Li_{4\,(\text{стаб})} \text{ и } {}^{7}_{4}Be_{3\,(\text{э.з})}; \quad {}^{21}_{10}Ne_{11\,(\text{стаб})} \text{ и } {}^{21}_{11}Na_{10\,(+)},$$
$$^{11}_{5}B_{6\,(\text{стаб})} \text{ и } {}^{11}_{6}C_{5\,(+)}; \quad {}^{39}_{19}K_{20\,(\text{стаб})} \text{ и } {}^{39}_{20}Ca_{19(+)}.$$

§ 21. РАДИОАКТИВНЫЕ ЯДРА И ИХ ИЗЛУЧЕНИЯ

Установлено, что радиоактивность элемента не изменяется, если он находится в составе любого химического соединения. Кроме того, на активность препарата не влияют обычные внешние воздействия: механическое давление, температура, электрические и магнитные поля и т. п. Это свидетельствует о том, что радиоактивный распад есть ядерный процесс.

ВИДЫ РАДИОАКТИВНЫХ РАСПАДОВ

Известны три вида радиоактивных распадов:

1) *альфа-распад*, при котором из ядра вылетает альфа-частица — прочная система из двух протонов и двух нейтронов (ядро атома гелия). При этом числа протонов и нейтронов в ядре уменьшаются на 2, а число нуклонов — на 4. Распад может быть записан в виде (X и Y — условное обозначение ядер)

$$^{A}_{Z}X_{N} \rightarrow {}^{A-4}_{Z-2}Y_{N-2} + {}^{4}_{2}He_{2}.$$

Одним из первых процессов такого типа был распад радия:

$$^{226}_{88}Ra_{138} \rightarrow {}^{222}_{86}Rn_{136} + {}^{4}_{2}He_{2}.$$

Заметим, что при альфа-распаде изменяется как число нуклонов (масса ядра), так и состав ядра (числа Z и N);

2) *бета-распад*, при котором из ядра вылетают либо электрон (e^-), либо позитрон (e^+) (напомним, что позитрон — частица, отличающаяся от электрона только положительным знаком заряда); электроны испускаются ядрами природных радиоактивных веществ; позитроны вылетают из радиоактивных ядер, полученных искусственным путем.

При бета-распаде масса ядра почти не изменяется (так как массы электрона и позитрона очень малы по сравнению с массами нуклонов); число нуклонов в ядре $A = Z + N$ остается постоянным. Однако заряд ядра, следовательно число протонов в ядре, изменяется на единицу; увеличивается — при вылете электрона и уменьшается — при вылете позитрона. Схема бета-распадов имеет вид

$$^{A}_{Z}X_N \rightarrow {}^{A}_{Z+1}Y_{N-1} + e^-; \quad {}^{A}_{Z}X_N \rightarrow {}^{A}_{Z-1}Y_{N+1} + e^+$$

(в некоторых записях вместо e пишут $\beta^{\pm}$). Таким образом, при бета-распадах состав ядра изменяется при постоянном числе нуклонов; это указывает на существование внутри ядра процессов превращения: нейтрона в протон (при испускании электрона) и протона в нейтрон (при испускании позитрона). В качестве примеров бета-распада приведем распад естественного тория и искусственно полученного радиоактивного изотопа меди:

$$^{234}_{90}Th_{144} \rightarrow {}^{234}_{91}Pa_{143} + \beta^-; \quad {}^{64}_{29}Cu_{35} \rightarrow {}^{64}_{28}Ni_{36} + \beta^+;$$

3) *гамма-распад*, при котором ни число нуклонов в ядре, ни состав ядра (Z и N) не изменяются. Ядро испускает электромагнитное излучение очень коротких длин волн: 10^{-10} — 10^{-13} м (обладающие большой проникающей способностью). Испускание гамма-фотонов, согласно квантовой теории, следует рассматривать как процесс перехода ядер из возбужденных (метастабильных) состояний в основное. Энергия этих фотонов может выделяться только в результате некоторых изменений в относительном расположении или движении нуклонов, т. е. в их распределении по ядерным уровням энергии. Если энергии начального и конечного состояния ядра равны E_1 и E_2, то энергия гамма-фотона будет равна $h\nu = E_1 - E_2$. Гамма-излучение часто сопровождает альфа- и бета-распады.

СЕМЕЙСТВА РАДИОАКТИВНЫХ ЭЛЕМЕНТОВ

Радиоактивные элементы разделены на «семейства»:

1. *Семейство урана — радия*, у которых числа нуклонов объединены формулой $A = 4n + 2$, где n — целое число, уменьшающееся на единицу, начиная с первого элемента этого семейства — урана-238 ($^{238}_{92}U_{146}$; $n = 59$); следующий элемент, полученный из урана-238 путем бета-распада, — торий-234 ($^{234}_{90}Th_{144}$) — имеет $n = 58$ и т. д. В нижеследующей таблице дан список членов этого семейства с указанием вида

распада и периода полураспада. Семейство заканчивается устойчивым изотопом свинца $^{206}_{82}Pb_{124}$ ($n = 51$).

Семейство $^{238}_{82}U - ^{206}_{82}Pb$; $A = 4n + 2$

Элемент	Тип превращения	Период полураспада	Элемент	Тип превращения	Период полураспада
Уран-238	α	$4{,}5 \cdot 10^9$ лет	Свинец-214	β	26,8 мин
Торий-234	β	24,1 дня	Висмут-214	β и α	19,7 мин
Протактиний-234	β	1,14 мин	Полоний-214	α	$1{,}5 \cdot 10^{-4}$ с
Уран-234	α	$2{,}7 \cdot 10^5$ лет	Таллий-210	β	1,32 мин
Торий-230	α	$8{,}2 \cdot 10^4$ лет	Свинец-210	β	22,2 года
Радий-226	α	1622 года	Висмут-210	β	4,97 дня
Радон-222	α	3,8 дня	Полоний-210	α	139 дней
Полоний-218	α	3,05 мин	Свинец-206	—	Стабилен

2. *Семейство тория*, имеющее формулу $A = 4n$. Семейство начинается с тория $^{232}_{90}Th_{142}$ ($n = 58$) и заканчивается изотопом свинца $^{207}_{82}Pb_{105}$ ($n = 52$).

3. *Семейство актиния* ($A = 4n + 3$) начинается с ядра урана $^{235}_{92}U_{143}$ ($n = 58$) и заканчивается устойчивым изотопом свинца $^{207}_{82}Pb_{105}$ ($n = 51$).

В связи с получением большого числа искусственных радиоактивных ядер было обнаружено еще одно семейство, имеющее формулу $A = 4n + 1$. Первым элементом этого семейства является ядро нептуния $^{237}_{93}Np_{144}$ ($n = 59$), последним — устойчивый изотоп висмута $^{209}_{83}Bi_{126}$ ($n = 52$).

ВЗАИМОДЕЙСТВИЯ ЯДЕРНЫХ ИЗЛУЧЕНИЙ С ВЕЩЕСТВОМ

Радиоактивные излучения (альфа-частицы, электроны, позитроны, гамма-фотоны) обнаруживаются по их взаимодействию с веществом. Прежде всего это взаимодействие приводит к *ионизации* вещества, которая оценивается по числу пар ионов, возникающих на единице длины пробега данной частицы в веществе (это число называется *удельной ионизацией*).

Альфа-частицы вылетают из ядер с различными скоростями, однако лежащими в сравнительно узких пределах: от 1,4 до $2 \cdot 10^7$ м/с, что соответствует энергиям от 4 до 8 МэВ. Энергия альфа-частиц, вылетающих из одинаковых ядер, почти одинакова. В среднем альфа-частица образует в воздухе при нормальных условиях около 30 000 пар ионов на один сантиметр своего пробега. С уменьшением ее скорости удельная ионизация увеличивается сначала медленно, затем быстро и лишь перед остановкой резко уменьшается до нуля.

На рис. IV.102, *а* показана удельная ионизация, производимая альфа-частицей на протяжении ее пробега R (расстояние, проходимое частицей в веществе до ее остановки, называется *пробегом* или *проникающей способностью*). Заметное увеличение ионизации перед остановкой частицы объясняется тем, что при малой скорости движения увеличивается время взаимодействия между частицей и встречными

молекулами вещества, а следовательно, и вероятность самого процесса ионизации. В связи с этим число альфа-частиц, проходящих через среду, слабо изменяется на первых участках пробега ($r < R$) и резко уменьшается, когда $r \approx R$ (рис. IV. 102, *б*). В воздухе при нормальных условиях быстрые альфа-частицы имеют средний пробег около 8 см; в более плотных средах пробеги гораздо короче и могут быть порядка сотых долей миллиметра. Поэтому стенки стеклянных ампул, содержащих радиоактивные препараты, почти полностью поглощают их альфа-излучения.

Бета-частицы (электроны), испускаемые радиоактивными веществами, имеют различные скорости — от очень малых до скоростей, близких к скорости света (до 0,988 от этой скорости, что соответ-

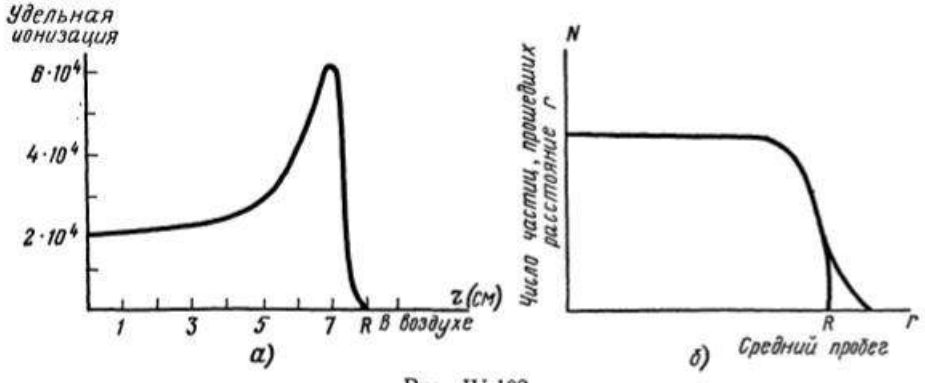


Рис. IV.102

ствует энергии около 5 МэВ). При прохождении через вещество электроны теряют энергию на:

1) ионизацию атомов вещества. Вследствие малой массы электронов и их большой скорости (и поэтому малой вероятности взаимодействия с атомом) способность электронов вызывать ионизацию слабее, чем у альфа-частиц ($\sim$ в 100 раз). Поэтому пробег электронов в веществе значительно больше, чем у альфа-частиц, и достигает в воздухе 200 м, в свинце — 3 мм и т. д.;

2) возбуждение рентгеновского излучения при торможении электронов в веществе (см. § 16). Этот процесс протекает интенсивно только в достаточно плотных средах, содержащих тяжелые атомы.

Поглощение потока электронов с одинаковыми скоростями в однородном веществе происходит приблизительно по экспоненциальному закону

$$N = N_0 e^{-\mu x}, \tag{4.11}$$

где N_0 — начальное число электронов, N — число электронов, прошедших слой вещества толщиной x, μ — коэффициент поглощения. Следует отметить, что для каждого вещества имеется некоторое значение $x = d$, при котором все электроны поглощаются, поэтому формула (4.11) применима только при $x < d$.

Измерение скоростей (энергий) электронов производится приборами (бета-спектрометрами), схема которых показана на рис. IV.103. Электроны, испускаемые препаратом A, проходят через щель $Б$ и поступают внутрь кольца K, помещенного в магнитное поле, перпендикулярное его плоскости. Это поле искривляет траекторию электронов, но только при условии, когда сила Лоренца будет равна центростремительной силе для окружности радиуса R:

$$evB = mv^2/R; \quad v = eBR/m,$$

электроны смогут попасть на приемник D. Следовательно, подбирая индукцию магнитного поля, приводящего электрон к приемнику, можно вычислить скорость электронов. Если препарат испускает электроны, имеющие различные скорости, то с помощью этого прибора можно определить полный спектр скоростей электронов.

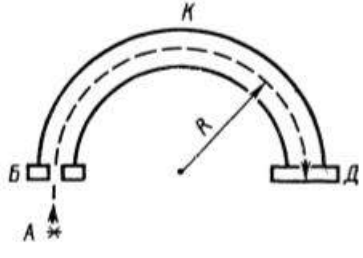


Рис. IV.103

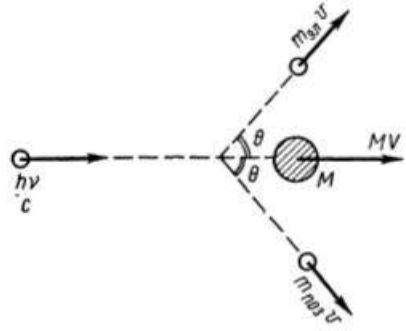


Рис. IV.104

Гамма-лучи поглощаются в веществе вследствие двух процессов:

1) фотоэлектрического эффекта; гамма-фотоны, поглощаясь атомами вещества, выбивают из них электроны. Если процесс происходит в газе, то эти электроны, обладая достаточной энергией, вызывают ионизацию;

2) образования «пары частиц»; установлено, что гамма-фотон, обладающий энергией не менее 1,02 МэВ, при столкновении с достаточно массивным ядром M исчезает и вместо него появляются две частицы: электрон и позитрон. Схема процесса показана на рис. IV.104. В этом процессе соблюдаются закон сохранения энергии (при этом необходимо учесть ту энергию, которая связана с массами появившихся частиц согласно соотношению Эйнштейна $\Delta E = \Delta mc^2$)

$$h\nu = m_{эл}c^2 + m_{поз}c^2 + MV^2/2 \qquad (4.12)$$

и закон сохранения импульса

$$h\nu/c = MV + m_{эл}v\cos\Theta + m_{поз}v\cos\Theta. \qquad (4.13)$$

Энергия, приобретаемая в этом процессе массивным ядром M, незначительна, поэтому энергия гамма-фотона расходуется в основном на «рождение» электрона и позитрона. Так как масса покоя этих частиц

$m_{эл} = m_{поз} = 9{,}11 \cdot 10^{-31}$ кг, то энергия, необходимая для их появления, $\Delta E = m_{эл}c^2 + m_{поз}c^2 = 2 \cdot 9{,}11 \cdot 10^{-31} \cdot (3 \cdot 10^8)^2$ Дж $= 16{,}4 \cdot 10^{-16}$ Дж $\sim$ $\sim$ 1,02 МэВ. К этой энергии необходимо добавить кинетическую энергию электрона и позитрона, разлетающихся с места столкновения, поэтому энергия фотона $h\nu$ должна быть несколько больше 1,02 МэВ.

Энергия гамма-лучей может передаваться среде также и при упругом столкновении со слабо связанными электронами; рассеянные вследствие комптон-эффекта фотоны имеют меньшую энергию, чем падающие.

Гамма-лучи являются одними из самых проникающих излучений. Если полагать, что их поглощение в веществе происходит по экспоненциальному закону $I = I_0 e^{-\mu x}$, то коэффициент μ для многих веществ оказывается очень малой величиной. Кроме того, μ зависит от энергии фотонов; с увеличением этой энергии уменьшаются потери на фотоэффект и комптон-эффект и начиная с $h\nu > 1{,}2$ МэВ резко возрастают потери на образование пар. Толщина слоя вещества, уменьшающая интенсивность гамма-излучения вдвое («толщина половинного поглощения»), равна для свинца 1,6 см, железа — 2,4 см, алюминия — 12 см, земли — около 15 см и более в зависимости от плотности.

БЕТА-РАСПАД ЯДЕР

Радиоактивные излучения являются основной информацией о структуре ядра и процессах, происходящих внутри них. В этом отношении важные результаты были получены при изучении бета-распадов.

Прежде всего необходимо было решить вопрос о происхождении вылетающих из ядра электронов. Предположение, что электроны вылетают не из ядра, а из электронной оболочки, было отвергнуто, так как при этом должно было бы наблюдаться рентгеновское и оптическое излучение. С другой стороны, ряд соображений указывает на то, что в ядре не могут находиться электроны:

1) магнитный момент одного электрона в сотни раз больше магнитного момента ядер;

2) если радиус протона, рассчитываемый на основании предположения, что энергия заряженной частицы mc^2 равна энергии, содержащейся в его электрическом поле (см. ч. III, § 7),

$$mc^2 = \frac{e^2}{8\pi\varepsilon_0 r} = k\frac{e^2}{2r}; \quad r = k\frac{e^2}{2mc^2}, \tag{4.14}$$

оказывается значительно меньше измеренных размеров ядра ($r_{прот} \approx$ $\approx 0{,}7 \cdot 10^{-16}$ см, $r_{ядра} \approx 10^{-13} - 10^{-12}$ см), то электрон по той же формуле имеет размеры почти в 2000 раз больше ($r_{эл} \approx 1{,}4 \cdot 10^{-13}$ см) и поэтому в объеме ядра уместиться не может;

3) согласно соотношению Гейзенберга (см. § 11), одна только неопределенность в скорости электрона, вынужденного пребывать в пределах ядра $\Delta x \approx 10^{-14}$ м, оказывается чрезвычайно большой:

$$\Delta v = \frac{h}{2\pi m \Delta x} = \frac{6{,}62 \cdot 10^{-34}}{6{,}28 \cdot 9{,}1 \cdot 10^{-31} \cdot 10^{-14}} \approx 10^{10} \text{ м/с}$$

(для протона $\Delta v \approx 5 \cdot 10^6$ м/с, что допустимо).

Существенное затруднение возникло при объяснении суммарного заряда и спина многих ядер. Например, ядро азота имеет заряд 7 и спин, равный единице. Если допустить, что оно состоит из 14 протонов и 7 электронов, то мы получим систему из нечетного числа частиц со спином 1/2; суммарный спин такой системы никак не может быть равным единице. Точно так же ядро дейтерия (дейтрон — ${}^{2}_{1}H_1$) не могло бы иметь спин, равный единице, если бы было составлено из двух протонов и одного электрона. Эти затруднения устраняются в протонно-нейтронной теории ядра, поэтому вопрос о происхождении электронов нужно было решать в рамках этой теории.

Другой важный результат был получен при измерениях энергии и импульсов электронов, вылетающих из ядер. Экспериментальные возможности современной физики позволяют измерять скорости и энергии не только частиц, вылетающих из ядер, но и самого ядра, испытывающего отдачу. Предварительно было установлено, что при распаде одно ядро выбрасывает один электрон. Тогда, согласно закону сохранения импульса, после каждого такого акта скорости ядра и электрона должны были лежать на одной прямой, тогда как, по наблюдениям, они составляли между собой разные углы. Это означает, что в процессе бета-распада возможно участвуют и другие частицы.

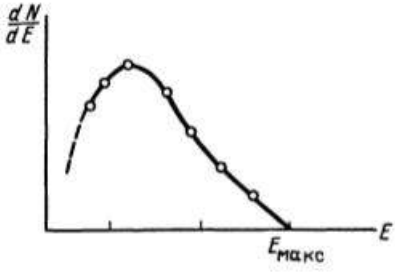


Рис. IV.105

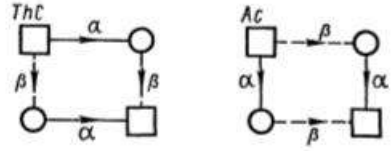


Рис. IV.106

Измерения показали, что электроны, появляющиеся при бета-распаде одинаковых ядер, имеют различные скорости (и энергии) от нуля до некоторого максимума, характерного для этого ядра. На рис. IV.105 показан типичный непрерывный спектр энергий электронов бета-распада. Если радиоактивный препарат поместить в калориметр и измерить выделившуюся энергию, то энергия, приходящаяся на каждый электрон, оказывается равной той средней энергии, которая определяется указанным на рис. IV.105 распределением электронов по энергиям. Таким образом, казалось бы, эти измерения свидетельствуют о том, что электроны вылетают из ядра с различными энергиями. Однако другие данные не согласуются с этим предположением. Известно, что некоторые ядра превращаются в другие по двум каналам: либо сначала выбрасывают альфа-частицу и затем (через некоторое время) электрон, либо, наоборот, сначала испускают электрон, а затем альфа-частицу. Схемы двух таких превращений, совершаемых изотопами тория и актиния, приведены на рис. IV.106.

Согласно закону сохранения энергии, при переходе системы из одного определенного состояния в другое через различные последовательности промежуточных состояний разность энергии должна быть одной и той же. В указанных выше переходах эта разность оказывается одинаковой только в том случае, если полагать, что при бета-распаде электроны вылетают из ядра не со средней энергией $E_{ср}$, а с максимальной энергией $E_{макс}$. Таким образом, эти измерения показывают, что при бета-распаде выделяется определенное количество энергии и остается неясным вопрос, почему энергия вылетающих электронов оказывается меньше $E_{макс}$.

Перечисленные выше затруднения привели В. Паули к гипотезе о существовании новой частицы — нейтрино, вылетающей из ядра вместе с электроном. Энергия, освобождающаяся при одном акте бета-распада, распределяется между электроном и нейтрино в различной пропорции. Нейтрино не имеет заряда и массы покоя, поэтому он обладает исключительно большой проникающей способностью и не задерживается стенками калориметра, вследствие чего вместо измерения полной энергии бета-распада удается определить только энергию электронов. Для устранения затруднений со спином ядер было предположено, что нейтрино имеет спин, равный половине. На основе этой гипотезы Э. Ферми разработал теорию бета-распада.

При бета-распаде общее число нуклонов в ядре остается постоянным, но числа протонов и нейтронов изменяются. Предполагается, что в условиях ядра возможны взаимные превращения нуклонов, сопровождающиеся рождением электронов, позитронов и нейтрино (обозначаемое ν_e):

$$n \to p + e^- + \bar{\nu}_e; \quad p \to n + e^+ + \nu_e.$$

Впоследствии оказалось, что нейтрино, выделяющиеся при этих превращениях, несколько отличаются друг от друга, поэтому ν_e получило название «электронное нейтрино», $\bar{\nu}_e$ — «электронное антинейтрино» (кроме них существуют еще «мюонные» нейтрино и антинейтрино; см. § 23). Так как масса нейтрона больше массы протона, то первая реакция идет с выделением энергии, а вторая требует затраты некоторого количества ядерной энергии.

У многих ядер (главным образом тяжелых) превращение протона в нейтрон происходит путем захвата электрона из электронной оболочки атома по схеме

$$p + e^- \to n + \nu_e.$$

Этот процесс называется *электронным захватом* (иногда K-захватом, так как электрон втягивается в ядро преимущественно из ближайшего к нему K-слоя; вероятность захвата электрона из L-слоя примерно в 100 раз меньше). Электронный захват обнаруживается по испусканию характеристического рентгеновского излучения, возникающего при заполнении образовавшихся вакансий в электронной оболочке атома. Радиоактивные ядра, испускающие протоны или нейтроны, в природе не встречаются и могут быть получены только искусственно. Кратковременное образование таких ядер устанавливается косвенно, по течению некоторых реакций. Например, появление ядер $^5_3Li_2$ (испус-

кающего протон) и $^{5}_{2}He_3$ (испускающего нейтрон) обнаруживается при изучении реакций:

$$^{3}_{2}He_1 + {}^{3}_{1}H_1 \rightarrow {}^{5}_{3}Li_2 \rightarrow {}^{4}_{2}He_2 + p;$$
$$^{3}_{1}H_2 + {}^{3}_{1}H_1 \rightarrow {}^{5}_{2}He_3 \rightarrow {}^{4}_{2}He_2 + n.$$

В последнее время Г. Н. Флеровым и его сотрудниками была обнаружена «протонная радиоактивность» у ядер, полученных при бомбардировке никелевой мишени ускоренными ядрами неона-20. Однако важнейшим естественным радиоактивным процессом, при котором изменяется число нуклонов и состав ядра, является альфа-распад.

ЭНЕРГИЯ СВЯЗИ ЯДЕР

В результате радиоактивного распада может быть получено либо стабильное, либо новое радиоактивное ядро. У некоторых ядер радиоактивное превращение сопровождается увеличением энергии связи. В этом случае можно утверждать, что исходное ядро является менее прочным, чем полученное, следовательно, радиоактивность ведет к образованию более прочных ядер. Однако можно указать на большое число радиоактивных превращений, при которых энергия связи не увеличивается, а уменьшается. Имея в виду также превращения, нельзя утверждать, что все радиоактивные ядра по значению энергии связи являются менее прочными, чем стабильные.

Ниже приводится перечень ряда радиоактивных превращений, с указанием периода полураспада T, энергии связи E, изменения этой энергии ΔE и энергии вылетающих частиц W. Знак минус у ΔE означает, что энергия связи при превращении уменьшается. Стрелкой показано направление процесса; рядом указан вид распада: испускание электрона (—), позитрона (+) или электронный захват (э. з.).

	T	E, кэВ	ΔE	W
$^{3}_{1}H_2$	12,262 лет	8 482		
↓(—)			—764	18
$^{3}_{2}He_1$	Стабилен	7 718		
$^{6}_{2}He_4$	0,83 с	29 265		
↓			+2727	3515
$^{6}_{3}Li_3$	Стабилен	31 992		
$^{10}_{4}Be_6$	$2{,}5 \cdot 10^6$ лет	64 997	—247	550
↓(—)				
$^{10}_{5}B_5$	Стабилен	64 750		
↑(+)				
$^{10}_{6}C_4$	19,1 с	60 631	+4389	2200
$^{67}_{29}Cu_{38}$	59 ч	585 386	—211	577
↓(—)				
$^{67}_{30}Zn_{37}$	Стабилен	585 175		
↑(э. з)				
$^{67}_{31}Ga_{36}$	78 ч	583 395	+1780	
↑(+)				
$^{67}_{32}Ce_{35}$	20 мин	578 210	+5185	3400

Из этой таблицы видно, что при β^--распаде наблюдается как увеличение, так и уменьшение энергии связи; при β^+-распаде и электронном захвате энергия связи только увеличивается. Величина ΔE во всех видах распада имеет большой разброс.

Найдем связь между энергией W, выделяющейся при радиоактивном распаде (в виде кинетической энергии частиц и энергии испускаемых фотонов) и разностью энергий связи исходного (E_1) и полученного (E_2) ядер. Имеем

$$E_1 = [(Zm_p + Nm_n) - M_1]c^2.$$

Если происходит β^--распад, то число нейтронов уменьшается, а число протонов увеличивается на единицу, поэтому

$$E_2 = [(Z+1)m_p + (N-1)m_n - M_2]c^2.$$

Из этих уравнений получаем

$$E_2 - E_1 = (M_1 - M_2)c^2 - (m_n - m_p)c^2.$$

При β^+-распаде и электронном захвате

$$E_2 = [(Z-1)m_p + (N+1)m_n - M_2]c^2.$$

Поэтому

$$E_2 - E_1 = (M_1 - M_2)c^2 + (m_n - m_p)c^2.$$

Следует иметь в виду, что при β^--распаде из ядра вылетает электрон, на рождение которого должна быть затрачена энергия m_0c^2. Поэтому величину $(M_1 - M_2)c^2$ следует приравнять сумме m_0c^2 и энергии W, выделяющейся при распаде:

$$(M_1 - M_2)c^2 = m_0c^2 + W.$$

Эта формула сохраняется и при β^+-распаде, так как на рождение позитрона необходима такая же энергия m_0c^2. Однако, при электронном захвате в ядро вносится энергия поглощенного электрона, поэтому

$$(M_1 - M_2)c^2 + m_0c^2 = W.$$

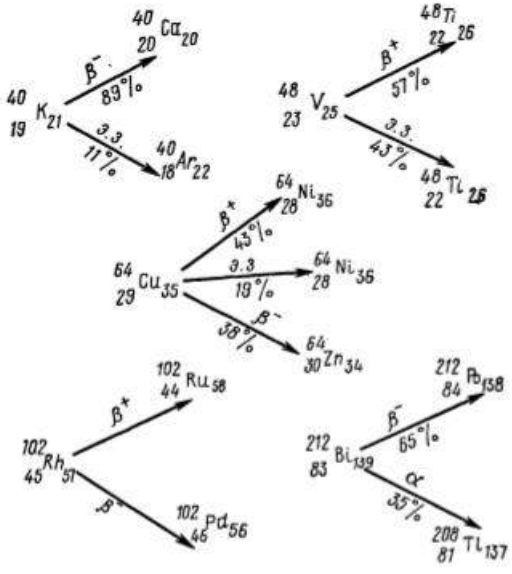
Воспользовавшись приведенными выше соотношениями между $\Delta M = M_1 - M_2$ и $\Delta E = E_2 - E_1$ для различных видов распада и учитывая, что $(m_n - m_p)c^2 = 1{,}29$ МэВ, $m_0c^2 = 0{,}51$ МэВ, получим, МэВ:

Для β^--распада	$W = \Delta E + 0{,}78$
» β^+-распада	$W = \Delta E - 1{,}8$
» электронного захвата	$W = \Delta E - 0{,}78$

ИЗОМЕРЫ

Большинство радиоактивных ядер испытывают распад единственным образом и превращаются в определенное конечное ядро. Однако существует много ядер, которые распадаются двумя или тремя спосо-

бами и образуют различные ядра-продукты, например



Существование таких ядер, имеющих несколько «каналов распада», можно объяснить двумя предположениями:

1) возможно, что радиоактивное ядро данного состава имеет одну определенную внутреннюю структуру, в которой уже заложена возможность распада по различным каналам, с определенным соотношением между вероятностями распада по этим каналам;

2) возможно, что при одном и том же составе ядра могут образоваться несколько различных структур (названных Содди *изотопами высшего порядка*). Каждая из этих структур имеет один определенный канал распада; распределение интенсивности распада между различными каналами соответствует процентному содержанию каждой структуры в естественном изотопе.

Экспериментальная проверка существования ядер, имеющих одинаковый состав, но различную внутреннюю структуру в основных (невозбужденных) состояниях, затруднена тем, что их невозможно отделить друг от друга известными физическими и химическими методами. Однако при помощи специальных косвенных методов в настоящее время установлено существование свыше 120 ядер, имеющих по две (а некоторые — и три) разновидности, названные изомерами. Радиоактивные изомеры имеют одинаковый состав (Z, N), но испытывают распад по двум (или трем) различным каналам. Например, одна из первых изомерных пар, открытая И. В. Курчатовым и др., — изомеры брома $^{80}_{35}Br_{45}$ обладают следующими свойствами:

1) основное ядро брома-80 является β^--радиоактивным и превращается в ядро $^{80}_{36}Kr_{44}$ с периодом полураспада, равным 18 мин;

2) второе ядро брома-80 имеет такой же состав, но оно непосредственно не превращается в ядро криптона-80, а переходит в указан-

ное выше основное ядро; период полураспада этого перехода равен 4,4 ч. После осуществления этого перехода происходит превращение полученного основного ядра брома-80 в ядро криптона-80.

Условились первое ядро, испытывающее радиоактивное превращение, называть *основным*, а второе ядро, которое непосредственно не распадается, а переходит в основное ядро без изменения состава, называть *изомером*. Переход изомера в основное ядро может осуществляться излучением гамма-фотонов или испусканием электронов *внутренней конверсии* (т. е. передачей энергии превращения одному из электронов оболочки атома). Так как освобождение ядром энергии при постоянном составе рассматривается как переход из возбужденного состояния в основное, то можно полагать, что изомер есть возбужденное состояние основного ядра. Однако, время существования изомеров значительно превосходит время жизни возбужденных состояний, поэтому полагают, что *изомер есть метастабильное состояние основного ядра.*

§ 22. ЯДЕРНЫЕ РЕАКЦИИ ДЕЛЕНИЯ И СИНТЕЗА. ТЕРМОЯДЕРНЫЕ РЕАКЦИИ

Обозначим условно буквами A, B, C, ... ядра различных элементов, а также различные частицы вещества: протон, нейтрон, электрон и т. д., обладающие массой покоя. Всевозможные реакции, в которых участвуют эти объекты, можно разделить на три группы:

1) реакции *синтеза*

$$A + B + \ldots \rightarrow C;$$

2) реакции *разложения*

$$A \rightarrow B + C + \ldots;$$

3) реакции *обмена* частицами (образование новых ядер из прореагировавших)

$$A + B + \ldots \rightarrow C + D + \ldots$$

Обозначим сумму масс всех тел, участвующих в ядерных реакциях: до реакции — через $M_0 = \Sigma M_i^0$, после реакции — через $M = \Sigma M_i$. Изменение массы в процессе реакции будет равно

$$\Delta M = M_0 - M.$$

Если $\Delta M > 0$, то реакция протекает с выделением энергии (экзотермическая); при $\Delta M < 0$ — с поглощением энергии (*эндотермическая*). Произведение $\Delta M c^2 = W$ является составной частью энергетического баланса реакции. Если рассматривать самый общий случай, когда в реакции участвуют не только частицы, имеющие массу покоя, но и частицы, не имеющие ее (фотоны и нейтрино), то на основании закона сохранения энергии должно быть написано:

$$M_0 c^2 + W_к^0 + W_{фот}^0 + W_ν^0 = M c^2 + W_к + W_{фот} + W_ν,$$

где $W_к$ — кинетическая энергия частиц, имеющих массу покоя, $W_{фот}$ — энергия фотонов, $W_ν$ — энергия нейтрино; индексом «ноль» обозначены начальные значения энергий. Однако нейтрино вследствие

своей большой проникающей способности не участвует в ядерных реакциях как одна из первичных частиц, поэтому в левой части этого равенства энергия нейтрино будет отсутствовать. Таким образом,

$$W = c^2\Delta M = (W_к - W_к^0) + (W_{фот} - W_{фот}^0) + W_ν. \qquad (4.15)$$

Расчеты упрощаются, если начальной кинетической энергией ядер можно пренебречь, а фотоны присутствуют либо только как возбудители реакций, либо только как ее продукты.

Массы частиц и ядер принято выражать в «атомных единицах массы» (сокращенно а. е. м.): 1 а. е. м. есть $^1/_{12}$ массы изотопа углерода $^{12}_6C_6$ и равна $1{,}66 \cdot 10^{-27}$ кг. Иногда удобно пользоваться энергетическим эквивалентом а. е. м., равным $1{,}66 \cdot 10^{-27} \cdot (3 \cdot 10^8)^2 = 931{,}8$ МэВ.

Возможность выделения ядерной энергии в реакциях разложения и синтеза связана с различными значениями «энергии связи одного нуклона» в легких и тяжелых ядрах (см. § 19, рис. IV.101).

Ядра, расположенные в начале и в конце периодической системы Менделеева, имеют меньшие энергии связи (одного нуклона), чем ядра, находящиеся в средней части (эта энергия имеет максимум для железа, никеля и соседних элементов). Поэтому если, например, ядро урана, для которого эта энергия равна 7,6 МэВ, разбить на два ядра, расположенные в средней части периодической системы и имеющие энергию связи порядка 8,6 МэВ, то можно получить в ы д е л е н и е энергии, равное 1 МэВ на каждый нуклон или около 200 МэВ на каждое разбитое ядро урана. Этот способ извлечения ядерной энергии используется в урановых реакторах. С другой стороны, если соединить легкие ядра с малыми энергиями связи (на один нуклон) в тяжелые — с большими, то также будет выделена ядерная энергия. Например, при синтезе двух изотопов водорода — дейтерия $^2_1H_1$ и трития $^3_1H_2$, имеющих энергии связи 1,112 и 2,827 МэВ, в ядро гелия $^4_2H_2$, у которого эта энергия равна 7,05 МэВ, выделится около 4 МэВ на каждый нуклон. Этот способ извлечения ядерной энергии использован в водородных бомбах.

БОМБАРДИРОВКА ЯДЕР ЧАСТИЦАМИ

Наиболее простые ядерные реакции осуществляются путем бомбардировки ядер-мишеней какими-нибудь частицами: протонами, нейтронами, альфа-частицами или ядрами других элементов. При внедрении этих частиц в ядро-мишень сначала образуется новое ядро, которое, как правило, оказывается очень неустойчивым и испытывает радиоактивный распад. При бомбардировке частицами, имеющими положительный заряд, существенно преодоление кулоновского отталкивания от ядра-мишени. Например, кинетическая энергия, которая необходима протону только для преодоления электрического отталкивания при приближении к ядру кислорода на расстояние порядка его радиуса $r \approx 4 \cdot 10^{-15}$ м, должна быть равна

$$W = k\frac{Ze^2}{r} = 9 \cdot 10^9 \frac{8\,(1{,}6 \cdot 10^{-19})^2}{4 \cdot 10^{-15}} = 1{,}8 \cdot 10^{-15}\ \text{Дж} \approx 10^6\ \text{эВ} = 1\text{МэВ}.$$

Чтобы иметь такую энергию, протон должен предварительно пробежать в электрическом поле разгоняющую разность потенциалов в 1 млн. В.

Для того чтобы проникнуть в ядро, энергия протона должна быть еще больше. Ядерные силы притяжения, действующие на протон, становятся соизмеримыми с электрическими силами отталкивания только на очень малых расстояниях от поверхности ядер.

Ядра окружены *ядерным потенциальным барьером*, препятствующим вылету нуклонов, не имеющих достаточной энергии; радиоактивный выход альфа-частиц, протонов и нейтронов из неустойчивых ядер объясняется «туннельным эффектом» (см. ч. IV, § 11). С другой стороны, частица (протон, нейтрон и др.), вошедшая в ядро с некоторой кинетической энергией, приобретает под действием ядерных сил притяжения некоторую дополнительную энергию. Эта энергия могла бы обеспечить обратный вылет частицы из ядра, но вследствие взаимодействия часть этой энергии передается другим нуклонам. Таким образом, поступление новой частицы в ядро сопровождается увеличением средней энергии нуклонов в ядре. Эта энергия может быть уменьшена при удалении каких-нибудь частиц из ядра (радиоактивность), а также путем испускания фотонов.

Рассмотрим несколько примеров ядерных реакций такого типа. Исторически первой была реакция, осуществленная Э. Резерфордом (1919 г.) при бомбардировке ядра азота альфа-частицами;

$$^{14}_{7}N_7 + ^{4}_{2}He_2 \rightarrow ^{17}_{8}O_9 + p,$$

т. е. ядро азота вместе с ядром гелия образуют сначала ядро изотопа фтора, которое оказывается весьма неустойчивым и почти мгновенно распадается на стабильный изотоп кислорода и протон. В этой реакции кинетическая энергия альфа-частиц превосходит кинетическую энергию продуктов реакции на 1,26 МэВ. (Сумма масс ядер азота и гелия равна $M_0 = 18{,}01139$, ядра кислорода и протона имеют массу $M = 18{,}01263$, изменение массы отрицательное: $\Delta M = -0{,}00124$, а $c^2\Delta M = -1{,}16$ МэВ.) В другой реакции

$$^{19}_{9}F_{10} + ^{4}_{2}He_2 \rightarrow (^{23}_{11}Na_{12}) \rightarrow ^{22}_{10}Ne_{12} + p$$

выделяется энергия, равная 1,58 МэВ.

Приведем несколько ядерных реакций, полученных при бомбардировке протонами:

$$^{13}_{6}C_6 + p \rightarrow ^{13}_{7}N_6 \rightarrow ^{13}_{7}N_6 + \gamma$$

(образовавшееся ядро азота оказывается в возбужденном состоянии, из которого возможен переход в нормальное путем излучения гамма-фотона);

$$^{63}_{29}Cu_{34} + p \rightarrow ^{64}_{30}Zn_{34} \rightarrow ^{63}_{30}Zn_{33} + n$$

(образовавшееся ядро цинка «перегружено» нейтронами и переходит в стабильное состояние, выбросив лишний нейтрон);

$$^{19}_{9}F_{10} + p \rightarrow ^{20}_{10}Ne_{10} \rightarrow ^{16}_{8}O_8 + ^{4}_{2}He_2$$

(полученные при реакции альфа-частицы имеют различные значения энергии, вследствие чего ядро кислорода оказывается или в нормальном, или в возбужденном состоянии; в последнем случае при переходе в нормальное состояние испускается гамма-фотон).

Интересное явление установлено при бомбардировке ядер дейтонами (ядром водорода, содержащим протон и один нейтрон). Вблизи поверхности ядра на нейтрон действуют только ядерные силы притяжения, тогда как на протон действуют значительные электрические силы отталкивания (возле ядра напряженность электрического поля достигает колоссальных значений порядка 10^{20}—10^{22} В/м). Вследствие этого дейтон, у которого энергия связи нуклонов (1,225 МэВ) очень мала, не проникает внутрь ядра-мишени, а разрывается вблизи него, причем нейтрон втягивается в ядро, а протон отбрасывается обратно:

$$^{13}_{6}C_7 + {}^{2}_{1}H_1 \rightarrow {}^{14}_{6}C_8 + p.$$

Напишем несколько реакций, вызываемых нейтронами:
1) образование дейтерия из обыкновенного водорода с выделением энергии 2,18 МэВ на каждое ядро:

$$^{1}_{1}H_0 + n \rightarrow {}^{2}_{1}H_1 + \gamma.$$

Вероятность осуществления такой реакции очень мала (из очень большого числа столкновений между протонами и нейтронами лишь небольшая часть приводит к образованию дейтона);

2) поглощение нейтрона кадмием, которое для медленных («тепловых») нейтронов (с энергией порядка 0,025 эВ) проходит с большой вероятностью:

$$^{113}_{48}Cd_{65} + n \rightarrow {}^{114}_{48}Cd_{66} + \gamma.$$

Обе указанные выше реакции относятся к *радиационным захватам*, при которых внедрение новой частицы в ядро сопровождается излучением фотона;

3) реакции отражения нейтронов ядрами углерода

$$^{12}_{6}C_6 + n \rightarrow {}^{13}_{6}C_7 \rightarrow {}^{12}_{6}C_6 + n;$$

4) реакции извлечения из ядер протона или альфа-частицы:

$$^{27}_{13}Al_{14} + n \rightarrow {}^{27}_{12}Mg_{15} + p; \quad {}^{27}_{13}Al_{14} + n \rightarrow {}^{24}_{11}Na_{13} + {}^{4}_{2}He_2;$$

5) если энергия нейтронов очень велика (10—20 МэВ), то возможны реакции с выходом нескольких нейтронов, например

$$^{63}_{29}Cu_{34} + n \rightarrow {}^{62}_{29}Cu_{33} + 2n.$$

Новое ядро, образовавшееся при проникновении в мишень бомбардирующей частицы (называемое иногда *составным ядром*), может иметь различные каналы распада в зависимости от энергии налетающей частицы. Характерным примером является поведение изотопа фтора $^{20}_{9}F_{11}$ при энергии бомбардирующего нейтрона, МэВ:

$$^{19}_{9}F_{10} + n \longrightarrow {}^{20}_{9}F_{11} \longrightarrow \begin{cases} {}^{20}_{9}F_{11} + \gamma \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \text{до } 1{,}5 \\ {}^{16}_{8}O_{11} + p \ldots\ldots\ldots\ldots\ldots\ldots\ldots & 4\text{—}10 \\ {}^{18}_{9}F_9 + 2n \ldots\ldots\ldots\ldots\ldots\ldots\ldots & 10\text{—}20 \\ {}^{15}_{8}O_7 + p + 4n \ldots\ldots\ldots\ldots\ldots\ldots & 20\text{—}40 \\ {}^{13}_{7}N_6 + 2p + 5n \ldots\ldots\ldots\ldots\ldots\ldots & 60\text{—}80 \end{cases}$$

Время существования составного (нестабильного) ядра может быть различным; у некоторых оно равно 10^{-14}—10^{-15} с. Обычно это время сравнивают с временем, которое необходимо, чтобы нуклон с энергией 10 МэВ (средняя энергия нуклонов в ядре) прошел путь, равный диаметру ядра. Это время имеет порядок 10^{-22}с, поэтому условились принять $T = 10^{-22}$ с за «естественное», или «характерное», ядерное время. Таким образом, время существования нестабильных составных ядер значительно больше ядерного времени, что дает основание рассматривать их как определенные промежуточные системы в цепи ядерных превращений.

При бомбардировке частицами ядер-мишеней вероятность попадания очень мала. Так, например, в опыте Резерфорда из 500 000 альфа-частиц лишь восемь проникли в ядра азота и вызвали ядерное превращение. Поэтому ядерные реакции такого типа весьма нерентабельны: они требуют затраты огромной энергии на сообщение частицам-снарядам необходимой кинетической энергии для проникновения в ядро-мишень и дают весьма незначительную полезную продукцию.

ЭФФЕКТИВНЫЕ СЕЧЕНИЯ РЕАКЦИЙ

Для оценки вероятности осуществления какой-нибудь ядерной реакции при бомбардировке ядер-мишеней различными частицами необходимо отдельно вычислить: 1) вероятность попадания бомбардирующей частицы в ядро-мишень и 2) вероятность реализации той реакции, которая нас интересует при каждом прямом попадании «снаряда» в «мишень». Например, результатом столкновения может быть упругое или неупругое (с поглощением энергии) рассеяние, внедрение снаряда в ядро с последующим выходом другой частицы или фотона и т. д. Из всевозможных течений процесса столкновения нас может интересовать только один определенный результат, и поэтому необходимо выяснить, как часто реализуется этот эффект среди множества попыток получить его.

Вероятность реализации определенного «акта взаимодействия» между бомбардирующей частицей и мишенью измеряется при помощи особой величины, получившей название *эффективного сечения*. Допустим, что частица-снаряд движется в объеме V, в котором имеется N одинаковых мишеней. Нас интересует число столкновений Δz, испытываемых этой частицей на длине пути Δx. Очевидно, что Δz будет пропорционально плотности $N/V = n$ и толщине слоя Δx:

$$\Delta z = \sigma n \Delta x.$$

Для пучка одинаковых частиц, движущихся в этой среде с одинаковыми скоростями, Δz означает число частиц, испытавших столкновение в слое толщиной Δx. Величина $\sigma = \Delta z/n\Delta x$ имеет размерность площади и называется *эффективным сечением* данного процесса (σ измеряется в барнах: 1 барн $= 10^{-24}$ см^2 $= 10^{-28}$ м^2): это есть число «актов взаимодействия», которое произошло бы в единичной толщине среды при единичной плотности мишеней. Эффективное сечение зависит от свойств и параметров как частицы, так и мишени и, кроме

того, от их энергии (главным образом от энергии частицы). Если столкновение — упругое и заканчивается изменением энергии и импульса частиц и мишени (в согласии с соответствующими законами сохранения), то σ называют *эффективным сечением упругого рассеяния*. Если «акт взаимодействия» приводит, например, к ядерной реакции, то σ называют *эффективным сечением этой реакции*.

Эффективное сечение некоторых реакций при определенных значениях энергии бомбардирующих частиц (обычно лежащих в узких пределах) имеет резкий максимум, что соответствует резонансу между дебройлевскими колебаниями налетающей частицы и собственными колебаниями нуклонов в ядре-мишени; это означает также, что энергия бомбардирующей частицы равна разности энергии двух квантовых состояний ядра-мишени. Таким образом, в общем случае σ должно учитывать не только вероятность самого столкновения при данных условиях, но и вероятность реализации определенного эффекта от этого столкновения.

ЯДЕРНЫЕ РЕАКТОРЫ

Для ядерной энергетики наиболее эффективными оказались реакции деления тяжелых ядер. Такое деление может происходить *спонтанно* (самопроизвольно, без внешнего воздействия, за счет «туннельного эффекта») или *вынужденно* при попадании в ядро подходящих частиц (такими частицами оказались нейтроны). Период полураспада спонтанного деления одного из наиболее подходящих для ядерной энергетики веществ — урана-235 — равен $7{,}07 \cdot 10^8$ лет, тогда как, бомбардируя его ядра потоком нейтронов, можно получить распад ядра за весьма короткое время. Ввиду этого спонтанное деление тяжелых ядер имеет вспомогательное значение; особое значение приобрели вынужденные реакции деления, возбужденные нейтронами.

Для осуществления ядерных реакций, возбуждаемых нейтронами, вовсе нет необходимости иметь «пучки нейтронов», направляемых на бомбардируемые ядра-мишени; вероятность попадания нейтронов в ядра будет при этом очень малой. Значительно более эффективным является помещение бомбардируемых мишеней в «нейтронный газ», который образуется в так называемых *ядерных реакторах*, имеющих стенки, хорошо отражающие нейтроны. В этом случае вследствие беспорядочного теплового движения самих нейтронов вероятность осуществления «акта взаимодействия» между ядром и нейтроном значительно выше, чем в пучках. Не менее важным преимуществом такого способа осуществления ядерных реакций является возможность у п р а в л е н и я ими. Если в объем, в котором существует нейтронный газ, ввести или вывести из него тела, сильно поглощающие нейтроны (например, кадмиевые стержни), то плотность нейтронного газа (их число в единице объема) может быть изменена и тем самым интенсивность ядерной реакции контролируема.

Для того чтобы при бомбардировке вызвать д е л е н и е ядра, необходимо сообщить ему некоторое минимальное количество энергии,

которое называется *энергией активации* или порогом *деления*. Если сообщаемая ядру энергия меньше энергии активации, то ядро перейдет из возбужденного состояния в нормальное, излучив фотон, т. е. сохранив свой состав. Энергия, вносимая частицей в ядро, равна сумме ее кинетической энергии (до попадания в ядро) и энергии связи, которую он приобретает внутри ядра под действием ядерных сил притяжения (средняя энергия связи на один нуклон для различных ядер приведена на рис. IV. 101). Поэтому минимальная кинетическая энергия частицы, способной вызвать реакцию деления, должна быть равна

$$E_к = E_{актив} - E_{связи}.$$

Все эти величины зависят от состава и структуры атомного ядра, поэтому, например, ядра урана-233, 235 и плутония-232 (имеющие нечетное число нуклонов) разбиваются на осколки при попадании в ядро «медленных» нейтронов (кинетическая энергия которых составляет доли электронвольта), а ядра урана-234, 238, тория-232 (имеющие четное число нуклонов) делятся только быстрыми нейтронами с энергиями не менее 1 МэВ.

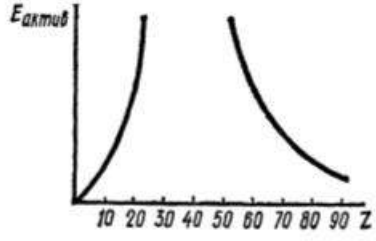


Рис. IV.107

Зависимость энергии активации от порядкового номера элемента показана на рис. IV. 107. С увеличением Z эта энергия растет у легких элементов и убывает у тяжелых. Сопоставляя эти значения $E_{актив}$ со значениями $E_{связи}$, данными на рис. IV. 101, можно найти кинетическую энергию частицы, способной вызвать ядерную реакцию деления.

Рассмотрим реакцию деления урана-235, вызванную нейтроном. Имеется много каналов, по которым это деление может осуществляться; укажем два способа деления:

$$^{235}_{92}U_{143} + n \rightarrow {}^{139}_{54}Xe_{85} + {}^{95}_{38}Sr_{57} + 2n,$$
$$^{235}_{92}U_{143} + n \rightarrow {}^{139}_{56}Ba_{83} + {}^{94}_{36}Kr_{58} + 3n.$$

Существенно, что поглощение одного нейтрона вызвало появление нескольких новых нейтронов; усредненное для всего множества различных каналов распада число новых нуклонов называется *коэффициентом размножения* (k) и оказывается больше единицы. Ввиду этого возможна *цепная реакция*, когда появившиеся нейтроны продолжают процесс, их вызвавший.

Допустим, что в реакторе в начальный момент времени находится N нейтронов. В течение времени Δt часть из них ΔN_1 поглотится активным веществом и вызовет деление, вследствие чего появятся $k\Delta N_1$ новых нейтронов. Другая часть ΔN_2 поглотится (без деления) ядрами стенок реактора и других тел, введенных в реактор для управления, отвода энергии и т. д. Наконец, третья часть ΔN_3 может покинуть пределы реактора вследствие некоторой прозрачности его

стенок. Если $\Delta N = k\Delta N_1 - \Delta N_2 - \Delta N_3$ будет положительной, то число нейтронов в реакторе, а следовательно, и интенсивность процесса деления будут увеличиваться; при $\Delta N < 0$ процесс деления будет затухать, а при условии $k\Delta N_1 = \Delta N_2 + \Delta N_3$ в реакторе будет установлено равновесное состояние, при котором число появляющихся нейтронов будет равно их потерям. Это условие не зависит от N, т. е. равновесие может быть достигнуто при любой плотности нейтронного газа в реакторе, т. е. при любой интенсивности процессов деления. Если W — энергия, выделяющаяся при одной реакции деления, то мощность реактора (энергия, выделяющаяся в единицу времени) будет равна

$$P = W\frac{\Delta N_i}{\Delta t},$$

т. е. определяется ежесекундным числом ядерных процессов деления. Предельно допустимое значение P определяется интенсивностью отвода выделяющейся энергии из реактора.

Отношение $\Delta N_1/N$, выражающее эффективное сечение процесса деления, зависит от энергии нейтронов. Для урана-235 оно достигает максимума при кинетической энергии нейтронов 0,025 эВ; при попадании нейтрона с такой энергией в ядро урана-235 вероятность деления оказывается равной 84%. Однако вторичные нейтроны, появившиеся при делении, имеют большие энергии — от 1 до 2 МэВ. Ввиду этого для осуществления цепной реакции (т. е. повышения вероятности деления при каждом попадании нейтрона в ядро) необходимо уменьшить энергию вторичных нейтронов. Этот процесс (названный *замедлением* нейтронов) осуществляется помещением в объем реактора, где находятся урановые стержни некоторых веществ (тяжелая вода, графит и др.); при упругих столкновениях с ядрами этих веществ нейтрон постепенно теряет кинетическую энергию до значений, соответствующих температуре реактора.

В реакторах в качестве источника ядерной энергии («ядерного топлива») должны быть использованы вещества, ядра которых имеют большую вероятность деления при попадании нейтрона. Например, у урана-235 эффективные сечения двух процессов, вызванных медленными нейтронами: деления и излучения фотона, — имеют значения 550 и 110 барн, тогда как у урана-238 деления не происходит, а излучение фотона имеет сечение 2,8 барн. Быстрые нейтроны поглощаются ураном-238 с большой вероятностью, однако и при этом вместо деления в основном происходит испускание фотона по схеме

$$^{238}_{92}U_{146} + n \rightarrow {}^{239}_{92}U_{147} + \gamma.$$

Для ядерной энергетики очень ценным оказался процесс распада урана-239, приведшего к образованию нового «ядерного горючего» плутония:

$$^{239}_{92}U_{147} \rightarrow {}^{239}_{93}Np_{146} + e^-; \quad {}^{239}_{93}Np_{146} \rightarrow {}^{239}_{94}Pu_{145} + e^-$$

(периоды полураспада: урана — 23 мин., нептуния — 2,3 дня, плутония — около 24 000 лет). Ядро плутония оказалось таким же «ядер-

ным горючим», как и уран-235: плутоний легко делится под действием медленных нейтронов и способен обеспечить цепную реакцию.

На использовании этих реакций основаны «бридерные реакторы», в которых из урана-238 получается химически легко отделимый от него плутоний. В реактор закладываются урановые стержни, в которых содержание урана-235 доведено до 5%. В результате деления урана-235 выделяются нейтроны с энергиями 1—2 МэВ. Эти нейтроны поглощаются ядрами урана-238, а последовательность указанных выше распадов приводит к образованию ядер плутония. В среднем на одно исчезнувшее ядро урана-235 появляются 1,5 ядер плутония-239. Так как уран-238 поглощает только б ы с т р ы е н е й т р о н ы, то в таких реакторах замедлители не нужны. Таким образом, в бридерных реакторах расходуется один вид ядерного горючего и воспроизводится (в бóльшем количестве) другой вид.

ЯДЕРНЫЙ СИНТЕЗ

Ядерные реакции синтеза были использованы для получения искусственных трансурановых элементов. В качестве бомбардирующих частиц использовались ядра, например, неона, а в качестве мишеней — тяжелые ядра, расположенные в конце периодической системы элементов. Так, в лаборатории Г. Н. Флерова (Дубна, 1964 г.) был получен курчатовий — элемент с атомным номером 104:

$$^{242}_{94}\mathrm{Pu}_{148} + {}^{22}_{10}\mathrm{Ne}_{12} \rightarrow {}^{260}_{104}\mathrm{Ku}_{156} + 4n.$$

Ядерные реакции могут быть вызваны также при бомбардировке ядер фотонами бóльших энергий (порядка средней энергии связи одного нуклона в ядре). Например, Д. Чадвик (1934) установил, что гамма-фотоны с энергией $h\nu \approx 2{,}23$ МэВ разрушают ядро дейтерия на протон и нейтрон (энергия связи нуклонов в дейтроне равна 2,225 МэВ). При помощи фотонов бóльших энергий (8—100 МэВ) можно извлечь из ядер протоны и нейтроны, например

$$^{16}_{8}\mathrm{O}_{8} + \gamma \rightarrow {}^{15}_{8}\mathrm{O}_{7} + n; \quad {}^{26}_{12}\mathrm{Mg}_{14} + \gamma \rightarrow {}^{25}_{11}\mathrm{Na}_{14} + p.$$

Одной из важных проблем современной энергетики является осуществление управляемых реакций соединения ядер легких элементов в ядро тяжелого элемента, например синтез гелия из дейтерия и трития:

$$^{3}_{1}\mathrm{H}_{1} + {}^{2}_{1}\mathrm{H}_{2} \rightarrow {}^{4}_{2}\mathrm{He}_{2} + n,$$

идущий с выделением 17,6 МэВ энергии. Подобного рода соединения можно произвести различными способами; при их выборе определяющим фактором является необходимость преодоления электрического отталкивания ядер, для чего синтезируемым ядрам должны быть сообщены достаточно большие кинетические энергии. Бомбардировка одних ядер другими (в условиях, когда ядра мишени покоятся или когда пучки синтезируемых ядер летят навстречу) была бы нерентабельной ввиду малой вероятности столкновений и малой вероятности осуществления самой реакции синтеза при отдельном стол-

кновении; затрата энергии на ускорение ядер значительно превосходит энергетический эффект такого способа. Кроме того, реакция должна быть управляема (неуправляемая реакция такого типа осуществлена в водородной бомбе), а отвод выделяемой энергии — удобным.

Для увеличения вероятности столкновения ядер можно смешать оба газа, а для обеспечения их сближения до необходимых для реакции расстояний — повысить температуру T смеси до значений, при которых кинетическая энергия ядер $W_к$ была бы достаточной (элементарный расчет дает W около 1 МэВ, T — порядка нескольких сотен миллионов градусов). Одно из известных направлений в решении этой проблемы заключается в том, чтобы вызвать в смеси легких газов кратковременные очень мощные электрические разряды и воспользоваться «пинч-эффектом» — сжатием образовавшегося плазменного шнура собственным магнитным полем. В этом направлении достигнуты известные успехи, однако проблема еще далека до удовлетворительного решения. В последнее время при помощи мощного лазерного излучения, сфокусированного в малом объеме среды, удалось получить сверхвысокие температуры ($\sim 10^8$ К) и вызвать термоядерную реакцию.

Ядерные цепные реакции протекают весьма быстро и могут иметь взрывной характер. В куске урана-235 энергия, выделяющаяся в единицу времени, зависит не только от его массы (числа ядер), но и от формы этого куска. Допустим, например, что кусок чистого урана-235 имеет форму очень тонкого диска и при массе M содержит N ядер. Далее предположим, что в этом диске ежесекундно происходит N_1 делений и появляется $kN_1 = n$ новых нейтронов (k — коэффициент размножения). Однако из этих нейтронов некоторая часть ($n_{исп}$) испарится с поверхности диска и лишь остальная часть ($n - n_{исп} = n_{раб}$) будет производить дальнейшее деление. Так как диск при данной массе имеет большую поверхность, то отношение $n_{исп}/n$ будет большим, а $n_{раб}/n$ — малым, поэтому количество ежесекундно выделяющейся энергии $P = n_{раб}\varepsilon = (n - n_{исп})\,\varepsilon = \varepsilon n(1 - n_{исп}/n) = \varepsilon k N_1(1 - \alpha)$ будет небольшим (ε — энергия, освобождающаяся при делении одного ядра; $\alpha = n_{исп}/n$).

Если к этому диску присоединить второй, то масса полученного куска будет вдвое больше, но поверхность, с которой происходит испарение, увеличится незначительно. В новом куске урана N_1 и n будут вдвое больше, а α — почти вдвое меньше (так как $n_{исп}$ останется почти неизменным). Вследствие этого интенсивность процесса деления в двух дисках будет больше, чем $2P$, т. е. при увеличении числа дисков мощность цепной реакции деления растет быстрее, чем масса делящегося вещества.

Для того чтобы рассматриваемый кусок урана не перегревался, необходимо отводить выделяющуюся энергию. Эта энергия может не только нагреть, но и испарить делящееся вещество. Если сближение дисков произвести очень быстро, то можно осуществить выделение энергии в виде взрыва. Очевидно, что при данной форме (например, цилиндра или шара) существует предельная масса вещества, называ-

емая *критической*, которая еще может сохраниться в твердом или жидком состоянии при полном отводе выделяющейся энергии, поэтому иметь кусок урана-235 с массой $m > m_{крит}$ невозможно. В ядерных бомбах сближение двух кусков делящегося вещества производится при помощи выстрела и поэтому выделение энергии носит взрывной характер. Заметим кстати, что при взрыве бомбы лишь незначительная часть ядер участвует в цепной реакции.

§ 23. ЭЛЕМЕНТАРНЫЕ ЧАСТИЦЫ

ОСНОВНЫЕ ХАРАКТЕРИСТИКИ

Отчетливого определения понятия «элементарная частица» не существует; обычно указывается только некоторый набор значений физических величин, характеризующих эти частицы, и их некоторые весьма важные отличительные свойства. Элементарные частицы имеют:

1) электрический заряд (e);
2) собственный момент импульса или спин (S);
3) магнитный момент (μ);
4) собственную массу — «массу покоя» (m_0).

В дальнейшем могут обнаружиться другие величины, характеризующие частицы, поэтому этот список основных свойств элементарных частиц не следует полагать законченным.

Однако не все элементарные частицы (список их приводится ниже) обладают полным комплектом указанных выше свойств. Некоторые из них имеют только электрический заряд и массу, но не имеют спина (заряженные пионы и каоны); другие частицы имеют массу, спин и магнитный момент, но не имеют электрического заряда (нейтрон, лямбда-гиперон); третьи — имеют только массу (нейтральные пионы и каоны) или только спин (фотоны, нейтрино). Обязательным для элементарных частиц является наличие хотя бы одного из перечисленных выше свойств. Заметим, что важнейшие частицы вещества — протоны и электроны — характеризуются полным комплектом этих свойств. Необходимо подчеркнуть: электрический заряд и спин являются *фундаментальными* свойствами частиц вещества, т. е. их численные значения сохраняются постоянными во всех условиях.

ЧАСТИЦЫ И АНТИЧАСТИЦЫ

У каждой элементарной частицы имеется ее противоположность — «античастица». Масса, спин и магнитный момент частицы и античастицы одинаковы, но если частица имеет электрический заряд, то ее античастица имеет заряд противоположного знака. У протона, позитрона и антинейтрона магнитные моменты и спины имеют одинаковые, а у электрона, нейтрона и антипротона — противоположные ориентации.

Взаимодействие частицы со своей античастицей существенно отличается от взаимодействия с другими частицами. Это отличие выражается в том, что частица и ее античастица способны к *аннигиляции*, т. е. к процессу, в результате которого они исчезают, а вместо них появляются другие частицы. Так, например, в результате аннигиляции электрона и позитрона появляются фотоны, протона и антипротона — пионы и т. д.

ВРЕМЯ ЖИЗНИ

Стабильность не является обязательным признаком элементарных частиц. Стабильными являются только электрон, протон, нейтрино и их античастицы, а также фотоны. Остальные частицы превращаются в стабильные либо непосредственно, как это происходит, например, у нейтрона, или через цепочку последовательных превращений; например, нестабильный отрицательный пион сначала превращается в мюон и нейтрино, а затем мюон превращается в электрон и другое нейтрино:

$$\pi^- \to \mu^- + \bar{\nu}_\mu; \quad \mu \to e + \nu_e + \nu_\mu.$$

Символами ν_μ и $\bar{\nu}_\mu$ обозначены «мюонные» нейтрино и антинейтрино, которые отличаются от ν_e и $\bar{\nu}_e$ — «электронных» нейтрино и антинейтрино.

Нестабильность частиц оценивается по продолжительности времени их существования от момента «рождения» до момента распада; оба эти момента времени отмечаются по трекам частиц в измерительных установках. При наличии большого числа наблюдений за частицами данного «сорта» вычисляется либо «среднее время жизни» τ, либо полупериод распада T. Допустим, что в некоторый момент времени $t = 0$ число распадающихся частиц равно N_0, а в момент t это число сделалось равным N. Полагая, что распад частиц подчиняется вероятностному закону

$$N = N_0 e^{-kt},$$

можно вычислить среднее время жизни $\tau = 1/k$ (в течение которого число частиц убывает в e раз) и период полураспада

$$T = \frac{0{,}693}{k} = 0{,}693\,\tau$$

(в течение которого это число уменьшается в два раза).

Интересно отметить, что:

1) все незаряженные частицы, кроме нейтрино и фотона, нестабильны (нейтрино и фотоны выделяются среди других элементарных частиц тем, что не имеют собственной массы покоя);

2) из заряженных частиц только электрон и протон (и их античастицы) являются стабильными.

Приведем список важнейших частиц (их число продолжает увеличиваться и в настоящее время) с указанием обозначений и основных

свойств; электрический заряд обычно указывается в элементарных единицах $e_0 = 1{,}602 \cdot 10^{-19}$ Кл, масса — в единицах массы электрона $m_0 = 9{,}10 \cdot 10^{-31}$ кг, спин — в единицах $h/2\pi = 1{,}054 \cdot 10^{-32}$ Дж·с:

Частица	Заряд	Масса	Спин	Среднее время жизни, с
Фотон γ	0	0	1	Стабильное
Нейтрино ν	0	0	1/2	»
Электрон e^- и позитрон e^+	± 1	1	1/2	»
Мюоны μ	± 1	206,8	1/2	$2{,}2 \cdot 10^{-6}$
Пионы π:				
заряженные	± 1	273,1	0	$2{,}61 \cdot 10^{-8}$
нейтральные	0	264,1	0	$0{,}89 \cdot 10^{-16}$
Каоны K:				
заряженные	± 1	966,4	0	$1{,}235 \cdot 10^{-8}$
нейтральные	0	974,2	0	$0{,}87 \cdot 10^{-10}$ и $5{,}68 \cdot 10^{-8}$
Протон p	± 1	1836	1/2	Стабильное
Нейтрон n	0	1838,6	1/2	$1{,}1 \cdot 10^{3}$
Гипероны:				
лямбда Λ	0	2183,1	1/2	$2{,}51 \cdot 10^{-10}$
сигма плюс Σ^+	$+1$	2327,7	1/2	$0{,}81 \cdot 10^{-10}$
сигма ноль Σ^0	0	2333,8	1/2	10^{-14}
сигма минус Σ^-	-1	2343,3	1/2	$1{,}65 \cdot 10^{-10}$
кси ноль	0	2572,7	1/2	$3 \cdot 10^{-10}$
кси минус	-1	2586,5	1/2	$1{,}74 \cdot 10^{-10}$
омега Ω	± 1	3276,9	3/2	$1{,}5 \cdot 10^{-10}$

КЛАССИФИКАЦИЯ ЧАСТИЦ

Изучение элементарных частиц показало, что группировка их по значениям основных свойств (заряд, масса, спин) недостаточна. Оказалось необходимым разделить эти частицы на существенно различные «семейства»:

1) фотоны, 3) мезоны,
2) лептоны, 4) барионы

и ввести новые характеристики частиц, которые показали бы принадлежность данной частицы к одному из этих семейств. Эти характеристики получили условное название «зарядов» или «чисел». Различают три сорта зарядов:

1) лептонно-электронный заряд (l_e);
2) лептонно-мюонный заряд (l_μ);
3) барионный заряд (B).

Этим зарядам придаются числовые значения: $+1$, 0 и -1 (знак плюс имеют частицы, минус — античастицы; фотоны и мезоны имеют нулевые заряды).

Элементарные частицы подчиняются следующим двум правилам:

каждая элементарная частица принадлежит только одному семейству и характеризуется только одним из указанных выше зарядов (чисел).

Например:

Электрон	$l_e=+1$	$l_\mu=0$	$B=0$
Позитрон	-1	0	0
Протон	0	0	$+1$
Нейтрон	0	0	$+1$

Однако одному семейству элементарных частиц может принадлежать некоторое множество различных частиц; например, к группе барионов относятся протон, нейтрон и большое число гиперонов. Приведем разделение элементарных частиц на семейства:

лептоны «электронные»: $l_e = \pm 1$, $l_\mu = 0$, $B = 0$. К ним относятся электрон (e^-), позитрон (e^+), электронное нейтрино (ν_e) и электронное антинейтрино ($\bar{\nu}_e$);

лептоны «мюонные»: $l_e = 0$, $l_\mu = \pm 1$, $B = 0$. К ним относятся мюоны с отрицательным (μ^-) и положительным (μ^+) электрическим зарядом и мюонные нейтрино (ν_μ) и антинейтрино ($\bar{\nu}_\mu$);

барионы: $l_e = 0$, $l_\mu = 0$, $B = \pm 1$. К ним относятся протон, нейтрон, гипероны и все их античастицы.

Существование или отсутствие электрического заряда не связано с принадлежностью к какому-нибудь из перечисленных семейств. Замечено, что все частицы, спин которых равен 1/2, обязательно имеют один из указанных выше зарядов. Фотоны (имеющие спин, равный единице), мезоны — пионы и каоны (спин которых равен нулю) не имеют ни лептонных, ни барионных зарядов.

Во всех физических явлениях, в которых участвуют элементарные частицы — в процессах распада, рождения, аннигиляции и взаимных превращений, — соблюдается второе правило:

алгебраические суммы чисел для каждого вида заряда в отдельности всегда сохраняются постоянными.

Это правило эквивалентно трем законам сохранения:

$\Sigma(\pm l_e) = \text{const}$ — закон сохранения числа электронных лептонов;

$\Sigma(\pm l_\mu) = \text{const}$ — закон сохранения числа мюонных лептонов;

$\Sigma(\pm B) = \text{const}$ — закон сохранения числа барионов.

Эти законы означают также, что *взаимные превращения между частицами, принадлежащими различным семействам, запрещены.*

Для некоторых частиц — каонов и гиперонов — оказалось необходимым дополнительно ввести еще одну характеристику, названную *странностью* и обозначаемую через S. Каоны имеют $S = \pm 1$, лямбда- и сигма-гипероны — $S = \mp 1$, кси-гипероны — $S = \mp 2$ (верхний знак — у частиц, нижний — у античастиц). В процессах, в которых наблюдается появление (рождение) частиц, обладающих странностью, соблюдается следующее правило:

$\Sigma(\pm S) = \text{const}$ — закон сохранения странности. Это означает, что появление одной странной частицы должно обязательно сопровождаться появлением еще одной или нескольких странных античастиц, с тем чтобы алгебраическая сумма чисел S до и после

процесса рождения оставалась постоянной. Замечено также, что при распаде странных частиц закон сохранения странности не соблюдается, т. е. этот закон действует только в процессах рождения странных частиц. Таким образом, *для странных частиц процессы рождения и распада необратимы.* Например, лямбда-гиперон (странность равна +1) распадается на протон и отрицательный пион:

$$\Lambda \to p + \pi^-.$$

В этой реакции закон сохранения странности не соблюдается, так как полученные после реакции протон и пион имеют странности, равные нулю. Однако в обратной реакции, при столкновении отрицательного пиона с протоном, одиночный лямбда-гиперон не появляется; реакция идет с образованием двух частиц, имеющих странности противоположных знаков:

$$p + \pi^- \to \Lambda + K^0.$$

Следовательно, в реакции рождения лямбда-гиперона закон сохранения странности соблюдается: до и после реакции алгебраическая сумма «странных» чисел равна нулю. Известна только одна реакция распада, в которой выполняется постоянство суммы странных чисел, — это распад нейтрального сигма-гиперона на лямбда-гиперон и фотон:

$$\Sigma^0 \to \Lambda + \gamma.$$

Другой особенностью странных частиц является резкое различие между продолжительностью процессов рождения (порядка 10^{-23} с) и средним временем их существования (около 10^{-9} с); для других (не странных) частиц эти времена имеют один порядок.

Заметим, что необходимость введения лептонных и барионных чисел или зарядов и существование указанных выше законов сохранения заставляют предполагать, что эти заряды выражают качественное различие между частицами различных сортов, а также и между частицами и античастицами. То обстоятельство, что частицам и античастицам необходимо приписать заряды противоположных знаков, указывает на невозможность взаимных превращений между ними.

РАСПАДЫ ЧАСТИЦ

Важным свойством элементарных частиц является их способность к взаимным превращениям. Эти превращения наблюдаются в процессах:

1) распада нестабильных частиц;

2) столкновения частиц, обладающих большими энергиями (в результате которых появляются новые частицы);

3) взаимодействия одних частиц, ведущих к появлению других.

Во всех этих процессах наряду с указанными выше законами сохранения соответствующих зарядов соблюдаются фундаментальные законы сохранения:

1) закон сохранения энергии, включая энергию, связанную с массой покоя (mc^2);

2) закон сохранения импульса;

3) закон сохранения механического момента (включая спин частиц).

Рассмотрим наиболее характерные процессы р а с п а д а частиц. Предварительно заметим, что некоторые нестабильные частицы распадаются различными способами (*каналами распада*), каждый из которых характеризуется средней вероятностью реализации (%). Кроме того, интересно заметить, что многие реакции распада сопровождаются появлением «пары частиц», принадлежащих к одному семейству, из которых одна является частицей, другая — античастицей, например $e^- + \bar{\nu}_e$, $e^+ + \nu_e$; $\mu^+ + \nu_\mu$, $\mu^- + \bar{\nu}_\mu$ и т. д. (в дальнейшем эти пары заключены в скобки).

Очень важным в атомной и ядерной физике является р а с п а д н е й т р о н а на протон, электрон и электронное нейтрино:

$$n \to p + e^- + \bar{\nu}_e.$$

Этот процесс в соответствии с законом сохранения числа барионов следует рассматривать как превращение нейтрона в протон, сопровождающееся появлением пары легких частиц: электрона и антинейтрино; обе эти частицы принадлежат к одному семейству — электронных лептонов.

Распад отрицательного мюона происходит по схеме

$$\mu^- \to e^- + \bar{\nu}_e + \nu_\mu.$$

Этот процесс можно трактовать либо как «превращение отрицательного мюона в электрон, сопровождающееся появлением двух нейтрино», либо же как «превращение мюона в соответствующее мюонное нейтрино, сопровождающееся появлением пары электрона и электронного нейтрино». Обе эти трактовки допустимы. Первая трактовка может быть обусловлена тем, что отрицательный мюон и электрон имеют очень много общего и отличаются только по величине массы; однако появившиеся в этом процессе нейтрино и антинейтрино не составляют «пару», так как нейтрино является м ю о н н ы м лептоном, а антинейтрино — э л е к т р о н н ы м. Вторая трактовка удобна тем, что в ней подчеркивается появление «пары», состоящей из лептона и антилептона одного семейства ($e^- + \bar{\nu}_e$); однако в этой трактовке допускается превращение заряженной частицы, имеющей собственную массу, в нейтральную частицу, не имеющую собственной массы.

Р а с п а д положительного п и о н а: происходит по двум каналам:

99,99% $\pi^+ \to \mu^+ + \nu_\mu$;

0,01% $\pi^+ \to e^+ + \nu_e$.

Существенное отличие этого процесса от распада мюона заключается в том, что пион распадается на д в е частицы, составляющие пару «лептон + антилептон». Так как μ^+ является нестабильной частицей, то, учитывая его распад, получим

$$\pi^+ \to (e^+ + \nu_e) + (\nu_\mu + \bar{\nu}_\mu).$$

Следовательно, можно сделать два предположения:

1) либо «пион превращается в мюон с выделением нейтрино»,

2) либо «пион распадается на пары, состоящие из лептона и антилептона».

Распад нейтрального пиона также имеет два канала:

98,8% $\pi^0 \to 2\gamma$;

0,2% $\pi^0 \to \gamma + e^- + e^+$.

Исчезновение нейтрального пиона и появление двух фотонов является примером непосредственного превращения частицы вещества, имеющей собственную массу, в частицы, лишенные этой массы.

Множество каналов распада имеют заряженные и нейтральные к а о н ы:

$K^+ \to (\mu^+ + \nu_\mu)$ (63,1%)
$\pi^+ + \pi^0$ (21,5%)
$\pi^+ + \pi^- + \pi^0$ (5,5%)
$(\mu^+ + \nu_\mu) + \pi^0$ (3,4%)
$(e^+ + \nu_e) + \pi^0$ (4,8%)

$K_1^0 \to \pi^+ + \pi^-$ (69%)
$\pi^0 + \pi^0$ (31%)

$K_2^0 \to \pi^+ + (e^- + \bar{\nu}_e)$, $\pi^- + (e^+ + \nu_e)$ } (35%)
$\pi^+ + (\mu^- + \bar{\nu}_\mu)$, $\pi^- + (\mu^+ + \nu_\mu)$ } (27%)
$\pi^0 + \pi^0 + \pi^0$ (27%)
$\pi^+ + \pi^- + \pi^0$ (11%)

Таким образом, мезоны распадаются на н е с т а б и л ь н ы е частицы; если учесть дальнейший распад полученных частиц, то можно установить, что конечными стабильными продуктами оказываются п а р ы лептонов и антилептонов. При распаде заряженного мезона в конечных продуктах обязательно присутствует электрон или позитрон (остальные продукты распада — нейтрино и фотоны).

Для распада нейтрального л я м б д а - г и п е р о н а напишем также и конечные продукты:

68,4% $\Lambda \to p + \pi^- \to p + (e^- + \bar{\nu}_e) + (\nu_\mu + \bar{\nu}_\mu)$;

31,6% $\Lambda \to n + \pi^0 \to p + (e^- + \bar{\nu}_e) + 2\gamma$.

Различие между этими двумя возможными каналами распада выявляется и в конечных продуктах. Лямбда-гиперон превращается в протон с выделением двух пар частиц; по второй схеме вместо пары «нейтрино + антинейтрино» выделяются два фотона. Если бы эти фотоны составляли пару «фотон + антифотон», то обе схемы были бы более похожими.

Несколько иную картину представляет распад сигма-гиперонов:

50% $\Sigma^+ \to p + \pi^0 \to p + 2\gamma$;

50% $\Sigma^+ \to n + \pi^+ \to p + (e^- + \bar{\nu}_e) + (e^+ + \nu_e) + (\nu_\mu + \bar{\nu}_\mu)$.

Заметим, что первая схема означает превращение сигма-гиперона в протон (через промежуточный этап, с участием нейтрального пиона). Вторая схема резко отличается от первой тем, что содержит в конечных продуктах три пары лептонов и антилептонов.

Для отрицательного сигма-гиперона имеется только один канал распада:

100% $\Sigma^- \to n + \pi^- \to p + 2(e^- + \bar{\nu}_e) + (\nu_\mu + \bar{\nu}_\mu)$.

Распад нейтрального сигма-гиперона:

100% $\Sigma^0 \to \Lambda + \gamma$.

Этот процесс отличается тем, что он представляет собой превращение одной тяжелой (имеющей собственную массу) частицы в другую без промежуточных этапов.

Схемы распадов отрицательного и нейтрального кси-гиперонов в первом этапе имеют простой вид:

100% $\Xi^- \to \Lambda + \pi^-$;

100% $\Xi^0 \to \Lambda + \pi^0$.

Однако в последующих этапах распад усложняется тем, что у лямбда-гиперонов существуют два различных канала распада.

В следующих реакциях две взаимодействующие частицы исчезают, вместо них появляются другие:

1) аннигиляция пары «частица + античастица», например

$$e^- + e^+ - \begin{cases} \gamma + \gamma, \\ \gamma + \gamma + \gamma. \end{cases}$$

Превращение этой пары в один фотон запрещается законом сохранения импульса (суммарный импульс пары относительно центра масс равен нулю, тогда как импульс одного фотона не может равняться нулю). Верхний процесс протекает, когда спины электрона и позитрона антипараллельны, нижний — параллельны; эти условия необходимы, чтобы суммарные спины частиц до и после аннигиляции могли быть одинаковыми.

Аннигиляция протона и антипротона проходит через несколько этапов: сначала появляются два пиона и новая частица ω («омега ноль-мезон»):

$$p + \bar{p} \to \pi^+ + \pi^- + \omega;$$

затем за время $\approx 10^{-22}$ с происходит распад:

$$\omega \to \pi^+ + \pi^- + \pi^0.$$

Появившиеся пионы затем распадаются по приведенным выше каналам; конечными (стабильными) продуктами аннигилации будут электроны, позитроны, нейтрино и фотоны;

2) взаимодействие нуклонов с пионами. Внутри ядра этим взаимодействием объясняется природа ядерных сил; вне ядра — приводит к образованию гиперонов и каонов:

$$\pi^- + p \to \Lambda^0 + K^0,$$
$$\Sigma^0 + K^0,$$
$$\Sigma^- + K^+.$$

Во всех написанных выше реакциях распада можно проследить соблюдение законов сохранения фундаментальных свойств частиц: электрического заряда, лептонных и баритонных чисел (законы сохранения энергии и импульсов относятся к состояниям частиц). Можно установить также сохранение спиновых квантовых чисел с учетом того, что спины частиц до и после распада могут иметь или параллельные, или антипараллельные ориентации. Закон сохранения механического момента предполагает суммирование орбитальных и спиновых моментов, однако если при распаде орбитальные моменты не появляются (например, когда траектории продуктов распада — прямые линии), то суммарный спин появившихся частиц будет равен спину распадающейся частицы.

Приведем несколько реакций, иллюстрирующих появление («рождение») новых элементарных частиц при столкновениях:

1) появление пары «электрон + позитрон» при столкновении двух электронов, кинетическая энергия которых (относительно центра масс) $W_к \geqslant 2m_0c^2$:

$$e^- + e^- + W_к \rightarrow e^- + e^- + (e^- + e^+);$$

2) при столкновении протонов, имеющих достаточно большие энергии ($W_к \geqslant c^2\Delta M$, где ΔM — разность масс частиц до и после столкновения), наблюдается появление различных частиц:

$$p + p \longrightarrow \begin{cases} p + p + (p + \bar{p}) \\ p + p + (n + \bar{n}) \\ p + K^+ + \Lambda^0 \\ p + \Sigma^+ + K^0 \\ p + \pi^+ + n \\ K^+ + \pi^+ + \Lambda^0 + n \\ p + K^+ + \Xi^0 + K^0 \end{cases}$$

ВИДЫ ВЗАИМОДЕЙСТВИЯ

Для образования физических систем из элементарных частиц необходимо взаимодействие между ними. Различают:

1) *сильное* взаимодействие; оно имеет место между элементарными частицами, принадлежащими к определенному классу. Эти частицы названы *адронами*; к ним относятся: протон, нейтрон, гипероны (с их античастицами); а также пионы, каоны и некоторые «тяжелые» мезоны. Их взаимодействие приводит к образованию атомных ядер и определяет их свойства;

2) *электромагнитное* взаимодействие между частицами, имеющими электрические заряды, электрические и магнитные моменты. Это взаимодействие приводит к образованию атомов и молекул и определяет их свойства. Электромагнитное взаимодействие существует и между электроном и фотоном, определяя течение процессов излучения и поглощения энергии атомами. Оно определяет также течение процессов, в которых одна элементарная частица превращается в другую путем одного только излучения фотонов ($\pi^0 \rightarrow \gamma + \gamma$; $\Sigma^0 \rightarrow \Lambda^0 + \gamma$), рождение и аннигиляцию электронно-позитронных пар и др.;

3) *слабое* взаимодействие между элементарными частицами, приводящее к образованию метастабильных физических систем и определяющее их свойства. Например, слабое взаимодействие определяет законы радиоактивного распада. Существенным для слабого взаимодействия является участие нейтрино;

4) *гравитационное* взаимодействие, имеющее универсальный характер и приводящее к образованию систем космического масштаба. Гравитационное взаимодействие определяет течение процессов, в которых участвуют макроскопические тела: движение тел в поле тяготения Земли, Солнца и т. п.

§ 1. ФИЗИЧЕСКИЕ ИЗМЕРЕНИЯ. СВЯЗЬ МЕЖДУ СПОСОБАМИ ИЗМЕРЕНИЯ ФИЗИЧЕСКИХ ВЕЛИЧИН И ФОРМУЛИРОВКАМИ ЗАКОНОВ ФИЗИКИ

Для изучения физических явлений и нахождения законов, управляющих ими, необходимо:

1) выбрать *физические понятия и величины,* при помощи которых должно производиться описание изучаемых явлений;

2) выбрать *систему отсчета*, т. е. тела, с которыми связываются координатные оси.

Физические понятия и величины должны иметь отчетливые (однозначные) определения; точная формулировка этих определений имеет весьма важное значение в физике. Для каждой физической величины должно быть дано подробное описание однозначного способа измерения (включая необходимую для этого аппаратуру) во всех условиях, при которых эта величина должна быть измерена. Необходимо точно указать, какие операции следует проделать в процессе измерения и что является результатом измерения.

Для измерения первых — *основных* — величин (размеры тел, расстояния, время, масса, заряд и др.) выбираются *эталоны.* Измерение других величин (скорость, ускорение, энергия, работа и др.), согласно их определению, сводится к измерению основных величин. Ввиду этого способы измерения основных величин имеют весьма важное значение и их следует проанализировать в первую очередь. Необходимо выяснить, не сделаны ли при определении результатов измерения этих величин какие-либо предположения, например, относительно «поведения» единичных эталонов или измерительных приборов. Остаются ли эти эталоны и измерительная аппаратура абсолютно тождественными при всех условиях их применения? Если допускается, что данный эталон (как и любое физическое тело) может претерпевать какие-либо изменения (например, эталонная линейка при повороте от горизонтального положения к вертикальному может изменять свои размеры), то возможно ли обнаружение этих изменений, с тем чтобы можно было внести необходимые коррективы в результаты измерений?

Система отсчета необходима для определения местонахождения и ориентации не только изучаемых тел, но и измерительной аппаратуры, в частности и единичных эталонов, так как их расположение или перемещение в данной системе отсчета может, вообще говоря, иметь некоторое влияние на результаты измерений. Выбирая то или иное тело отсчета — корпус экспериментальной установки, стены лаборатории, поверхность Земли и т. п., мы *условно считаем их неподвижными.*

Можно предполагать, что существует тело отсчета, которое не условно, а на «самом деле» является неподвижным; таким телом отсчета могло быть окружающее нас пространство, или так называемый «мировой эфир», в котором существуют гравитационные, электрические и магнитные поля и происходит распространение электромагнитных (и световых) волн. Однако измерения физических величин по отношению к «мировому эфиру» можно было бы осуществить только в том случае, если бы удалось найти физическую систему, которая покоилась бы относительно этого «эфира». Это было бы возможно, если бы движение тел относительно «эфира» сопровождалось явлениями, которые позволили бы определить скорость этого движения. Возникает проблема обнаружения и измерения так называемого «абсолютного движения», т. е. движения тел относительно «мирового эфира». Этой проблеме были посвящены целый ряд весьма тонких экспериментов, однако найти признаки, по которым можно было бы установить, движется ли или покоится данное тело (вернее, измерительная аппаратура) относительно «мирового эфира», не удалось.

Рассмотрим теперь связь между законами физики и определениями (способами измерения) величин, входящих в эти законы. Перечислим сначала основные очевидные положения:

1) законы физических явлений могут быть найдены только после выбора способов измерения всех величин, необходимых для описания этих явлений, а также после производства измерений, их обработки и анализа.

Следовательно, при выборе определений и способов измерения физических величин нельзя ссылаться на какие-либо ранее найденные законы физики;

2) если для измерения какой-нибудь величины (например, времени) используется определенное физическое явление, то законы, управляющие этим явлением, должны постулироваться.

Эти законы не могут быть известны заранее, так как они содержат ту самую физическую величину, способ измерения которой еще только выбирается;

3) выбор определений и способов измерения физических величин должен производиться в какой-то последовательности. Например, сначала можно установить способ измерения размеров, затем массы, зарядов и т. д.

Очевидно, что при выборе измерительных операций для первой физической величины (например, длины) нельзя пользоваться измерением других величин, так как для них способы измерений еще не установлены. Точно так же нельзя ссылаться на какие-либо опытные данные, получение которых потребовало бы измерения других величин. Ввиду этого для первой физической величины «поведение» единичного эталона (а также измерительной аппаратуры) в различных условиях придется постулировать. Такое постулирование оказывается необходимым и для последующих физических величин.

Рассмотрим простой пример — измерение размеров тел путем последовательного перекладывания эталонной линейки; результатом измерения является только число перекладываний. Допустим, что

измерение производится в неоднородном поле тяготения. В общем случае при n перекладываниях следует полагать, что определяемый размер равен

$$L = l_{01} + l_{02} + \ldots + l_{0n}, \tag{1}$$

где l_{0i} означает длину эталона при i-м прикладывании к измеряемому телу. Однако никакой возможности контролировать размеры эталона не имеется (L является первой величиной, способ измерения которой только что выбран; остальные величины еще не определены и никаких контрольных опытов произвести нельзя); поэтому мы вынуждены полагать, что $l_{0i} = l_0$ и $L = nl_0$. Таким образом, в утверждении, что число n выражает результат измерения размера L (по определению), неявно содержится предположение о неизменяемости эталона в процессе его использования. Заметим, что постоянство числа n для данного тела при всевозможных ориентациях этого тела в поле тяготения еще не означает независимости размеров тела от его расположения, так как возможно, что и тело и эталонная линейка одновременно изменяют свои размеры.

Одной из важнейших физических величин является время. Для измерения времени необходимо выбрать эталонные часы, в которых происходит некоторый периодически повторяющийся процесс; период τ этого процесса принимается за единицу времени. Очевидно, что «постоянство» этого периода можно было бы контролировать только при помощи других часов, которые полагаются более точными. Но такой контроль, по существу, означает, что первые часы уже не считаются эталонными и заменены новыми. Неизменность периода повторения любого физического процесса, взятого в качестве эталона для измерения времени, является п р е д п о л о ж е н и е м, необходимым для определения «результата измерения» времени. В общем случае, если отмечено n возвращений часов в исходное состояние, прошедшее время t должно быть приравнено сумме:

$$t = \tau_{01} + \tau_{02} + \ldots \tau_{0n}. \tag{2}$$

Обычно же мы принимаем

$$t = n\tau_0,$$

полагая, что $\tau_{0i} \equiv \tau_0$. Очевидно, что выбор «хороших часов» (не только удобных для измерения, но и менее подверженных воздействию внешних факторов) имеет большое значение. Так, например, ход механических (маятниковых) часов может зависеть от температуры, напряженности гравитационного поля и ориентации оси часов относительно направления этого поля и т. д. В этом отношении различные «атомные часы» являются более надежными, однако нет уверенности, что на ход этих часов а б с о л ю т н о не влияют другие поля (или физические факторы). Допущение, что промежуток времени между двумя повторяющимися событиями в часах (т. е. «единица времени») остается постоянной величиной в о в с е х у с л о в и я х, в которых производится измерение времени при помощи данных часов, является произвольным, условным, постулативным. Можно производить за-

мену одних часов другими, более удобными или менее подверженными внешним воздействиям, однако контролировать постоянство «единицы времени» у эталонных часов принципиально невозможно.

Подробный и достаточно строгий анализ способов измерения других физических величин — массы, электрического заряда, энергии, работы и т. д. — также приводит к выводу о неизбежности использования произвольных предположений при определении результатов измерений каждой физической величины.

Из изложенного выше вытекает важное следствие:

формулировка объективных законов физики, полученных при теоретической обработке экспериментальных данных, содержит в себе все предположения, которые используются при определении результатов измерения физических величин.

При переходе от одних способов измерения к существенно другим вид объективных законов физики может измениться. Если же мы из каких-либо соображений придадим какому-нибудь закону физики определенный вид, то этим будут предопределены способы измерения физических величин, входящих в эти законы. Так, например, если сначала выбрать независимые друг от друга (применимые к любым системам отсчета) способы измерения расстояний и времени, то, производя измерения, можно найти законы распространения света во всех условиях относительно любых систем отсчета. Но можно поступить и иначе: можно согласиться придать этому закону определенный вид, например постулировать постоянство скорости света относительно всех инерциальных систем отсчета; тогда можно найти те способы измерения расстояний и времени, которые соответствуют этому постулату.

При историческом формировании различных областей физики специального согласования или даже достаточно полного исследования вопроса о способах измерения физических величин, о связи между ними и формулировкой законов физики проделано не было. В связи с опытом Майкельсона и созданием теории относительности был затронут лишь вопрос об измерениях длин (размеров твердых тел) и времени. Поэтому не исключена возможность такой ситуации, когда одна часть физики (например, механика) основана на одном «комплекте» способов измерения физических величин, а другая часть (например, электродинамика) — на другом, существенно отличном от первого. Подтверждением этого предположения можно считать то обстоятельство, что уравнения классической механики оказались инвариантными относительно преобразований Галилея, тогда как уравнения классической электродинамики потребовали преобразований Лоренца.

При производстве физических измерений, для того чтобы найти значение интересующей нас величины x, необходимо привести измерительный прибор во взаимодействие с изучаемым объектом. В некоторых случаях это взаимодействие оказывается настолько слабым, что влиянием процесса измерения на искомую величину x можно

пренебрегать. Однако в общем случае следует иметь в виду, что значение величины x до процесса измерения может отличаться от того значения x', которое показывает аппаратура, по двум причинам:

1) вследствие воздействия приборов на изучаемый объект,

2) вследствие изменения параметров измерительной установки, вызванного обратным воздействием объекта на прибор (например, при прохождении электрического тока через гальванометр происходит нагревание его обмотки и деталей, вследствие чего может измениться цена деления его шкалы).

Необходимыми поправками к измеренному значению x' для нахождения истинного значения x во многих измерительных операциях (главным образом макроскопического масштаба) можно пренебрегать. Однако если и объект и прибор представляют собой *микрофизические* системы, то при их взаимодействии искажения величины x могут быть такого же порядка, как и сама измеряемая величина. Такая ситуация возникает при измерительных операциях, производимых в атомных и ядерных масштабах.

§ 2. О ТЕОРИИ ОТНОСИТЕЛЬНОСТИ ЭЙНШТЕЙНА

В теории относительности Эйнштейна важное значение имеет предложенный им способ измерения времени и новое определение понятия одновременности событий, происходящих в различных местах инерциальных систем отсчета.

В классической физике до появления теории относительности (1905) предполагалось, что любой физический процесс, используемый (как «эталонный») для измерения времени, выявляет одно и то же течение «мирового» или *абсолютного времени*. Предполагалось также, что течение времени, измеренное при помощи любых часов, не зависит от того, покоятся или движутся эти часы относительно данной системы отсчета. А. Эйнштейн прежде всего указал на недостаточную обоснованность этих предположений и на необходимость предварительного выбора строго определенного (однозначного) способа измерения времени, пригодного (без изменений) для всех условий, в которых производятся эти измерения. Для того чтобы более отчетливо изложить его физические идеи, рассмотрим измерение длин, расстояний и времени в двух системах отсчета, движущихся друг относительно друга.

Допустим, что нам необходимо измерить расстояние между точками A и B, отмеченными на каком-нибудь твердом предмете.

Каждый «способ измерения» должен содержать в себе указание о том, как следует измерять данную физическую величину *во всех условиях,* в которых приходится производить измерение. В частности, приходится измерять размеры не только покоящихся (относительно системы отсчета), но и движущихся тел. В этой весьма важной части «способа измерения» возможны два приема:

1) можно каждый раз сообщить измерительному прибору (например, эталону длины или интерферометру) такие же скорости, какие имеет измеряемый предмет. В этом случае измерительный прибор в процессе измерений всегда покоится относительно предмета, но дви-

жется относительно системы отсчета. Нам необходимо знать, не изменяется ли эталон длины вследствие своего движения и не зависят ли результаты измерения от скорости этого движения;

2) можно отметить те точки A' и B' системы отсчета, в которых одновременно находятся точки A и B движущегося предмета, и затем определить расстояние между A' и B'. В этом случае измерительный прибор в процессе его использования всегда покоится относительно системы отсчета. Это удобно, но оно связано с необходимостью дополнительно измерять время, так как одновременность совпадения точек A и A', B и B' должна быть отмечена по часам. Поэтому, если имеется однозначный способ измерения времени (а следовательно, способ установления одновременности совпадений точек A и A', B и B'), то второй прием будет иметь преимущество перед первым.

Заметим, что оба приема не применимы, если необходимо измерять *расстояние между двумя точками пространства*, например, атомного или космического масштаба. Расстояние, не фиксированное твердыми телами, не может быть измерено эталонными линейками. Поэтому это расстояние либо вообще не может быть основной величиной (т. е. величиной, для которой выбран эталон), либо же для него должен быть указан новый способ измерения, который необходимо отдельно проанализировать, чтобы выявить все содержащие в нем предположения.

Рассмотрим измерение времени; для этого необходимо выбрать часы, т. е. физические системы, в которых происходит какой-нибудь периодический процесс. Результатом измерения времени в данном месте является подсчет числа возвращений часов и одно и то же фиксируемое состояние. Для измерения времени в различных местах возможны два приема:

1) можно изготовить некоторое множество совершенно одинаковых часов, отрегулировать их ход в определенном месте (и при определенной ориентации их осей) и затем расставить их в интересующих нас местах, где должно производиться измерение времени. При этом необходимо знать, не изменяется ли ход часов при их переносе и работают ли они синхронно в местах назначения. Вполне возможно, что, например, в различных местах поля тяготения Земли часы имеют различный ход;

2) можно предварительно расставить эти часы и затем уже согласовать их ход между собой при помощи каких-нибудь «сигналов времени» — световых вспышек и т. д. В этом случае необходимо заранее знать законы распространения сигналов в данной системе отсчета.

Для измерения времени в движущихся физических системах можно, опять-таки, либо пользоваться часами, которые движутся вместе с системой и, следовательно, неподвижны относительно нее (в этом случае нужно знать, как зависит ход часов от скорости их движения); либо же пользоваться «сигналами времени» (в этом случае необходимо знать законы распространения этих сигналов относительно любой физической системы, в которой производится измерение времени). Нет уверенности в том, что результаты измерения времени при помощи всех этих способов будут совпадать.

Таким образом, при рассмотрении способов измерения физических величин, в частности — длины и времени, *обнаруживается следующая проблема*: при производстве измерений необходимо заранее знать, как изменяются измерительные приборы и, в частности, эталоны физических величин (линейка, часы и т. п.) в процессе измерений. Сами результаты измерений — число перекладываний линейки или подсчет числа колебаний в часах и т. д. — не могут решить этой проблемы.

Можно было бы выяснить поведение эталонов в различных условиях их применения, если были бы известны необходимые для этого законы физики, но они могут быть установлены только после производства и обработки измерений; заранее они не могут быть известны. Особенно отчетливо выявляется возникающее здесь затруднение при измерении времени при помощи «сигналов», например, световых вспышек. Для измерения времени, прошедшего с момента выхода сигнала из одной точки до момента прихода сигнала в другую точку, должны быть заранее известны законы распространения света, но эти законы могут быть получены только после экспериментального изучения распространения света, для чего способ измерения времени уже должен быть выбран.

Очевидно, указанную выше проблему можно разрешить, если удастся найти такие законы физики, установление которых не требует производства измерений. Тогда способы измерения физических величин могли быть основаны на таких законах; остальные законы физики получатся в результате обработки измерений. Таким (*не зависящим от способов измерения физических величин*) законом физики является следующий результат:

движение физических систем относительно пространства не обнаруживается.

Можно было ожидать, что движение тел относительно «мирового эфира» повлияет на форму и размеры тел, на распределение зарядов в них, на форму эквипотенциальных поверхностей силовых полей (электрических, гравитационных), на форму волновых поверхностей света, испускаемого телами, и. д. Большая группа экспериментов была поставлена с целью обнаружить влияние орбитального движения Земли на свойства тел или на течение физических процессов. Например, известные расчеты показывали, что при повороте интерферометра Майкельсона (ч. IV, § 4, рис. IV. 14) должно было наблюдаться заметное смещение интерференционных полос, пропорциональное скорости движения этого прибора вместе с Землей; однако это смещение не было обнаружено. При измерении сопротивления прямолинейного проводника (например, при помощи мостика Уитстона) ориентировали его по направлению движения Земли и перпендикулярно к этому направлению; ожидаемое изменение сопротивления также не было обнаружено. На основании многочисленных подобных экспериментов было установлено, что направление и скорость движения тел относительно «мирового эфира» обнаружить невозможно.

Среди других законов физики этот экспериментально установленный результат выделяется тем, что он *не связан с производством ка-*

ких-либо измерений, в которых было бы необходимо употребление эталонов или других измерительных приборов.

Действительно, для того чтобы показать, что движение физических систем, установленных на Земле, относительно «мирового эфира» обнаружить невозможно, нужно только констатировать *отсутствие регистрируемых изменений* в экспериментальных установках при их повороте. Например, чтобы отметить отсутствие интерференционного эффекта в приборе Майкельсона при его поворотах, вовсе нет необходимости измерять расстояния до зеркал или время распространения света в различных направлениях. Точно так же и в других экспериментах фиксируется только отсутствие ожидаемого эффекта при движении измерительной установки относительно «мирового эфира»; никаких измерений производить не приходится, поэтому в предварительном выборе способов измерения физических величин нет необходимости.

Окружающее нас пространство играет исключительную роль в физических явлениях; в нем существуют гравитационные, электрические, магнитные поля, распространяются электромагнитные волны, через него осуществляются всевозможные взаимодействия между телами и т. д. Если движение тел относительно «мирового эфира» не обнаруживается, то можно предположить, что *наблюдаемые нами физические явления происходят совершенно одинаково в системах покоящихся или как угодно движущихся относительно «мирового эфира».* Из этого утверждения можно сделать вывод, что законы физики должны иметь одинаковый вид по отношению к различным системам отсчета. В частности, если измерение времени производится при помощи «сигналов», то закон распространения этих сигналов можно выбрать одинаковым для всех систем отсчета.

Эти результаты были использованы А. Эйнштейном в разработанной им теории относительности. В первой части этой теории — в так называемой специальной теории относительности — утверждается следующее:

1) *законы физических явлений имеют одинаковый вид по отношению ко всем инерциальным системам отсчета* (принцип относительности);

2) *скорость распространения света одинакова по всем направлениям для всех инерциальных систем отсчета* (принцип постоянства скорости света). Как показал А. Эйнштейн, на этих основных принципах его теории могут быть обоснованы способы измерения всех физических величин. Важнейшая физическая величина, характеризующая все изменения в природе, — время — измеряется при помощи световых сигналов, вследствие чего скорость света вошла во все формулы теории относительности.

Допустим, что один из законов физики, полученный относительно системы отсчета S, имеет вид

$$f(x,\ y,\ z,\ t,\ \ldots)=0,$$

а относительно системы отсчета S' имеет вид

$$f'(x',\ y',\ z,\ t',\ \ldots)=0.$$

Согласно принципу относительности, функции f и f' должны иметь одинаковый вид. Это возможно, если между результатами измерения физических величин относительно S и S' существуют определенные соотношения. А. Эйнштейн показал, что из двух принципов его теории следует, что координаты движущихся тел и время, измеренные относительно S и S', связаны между собой *преобразованиями Лоренца*:

$$x' = \frac{x - vt}{\sqrt{1 - v^2/c^2}}; \quad y' = y; \quad z' = z; \quad t' = \frac{t - xv/c^2}{\sqrt{1 - v^2/c^2}} \tag{3}$$

(при выводе этих формул предполагается, что оси OX и OX' систем отсчета S и S' совпадают, v — есть скорость движения S' относительно S, а тела движутся параллельно оси иксов). Далее он показал, что уравнения Максвелла (ч. III, § 24 и 29), описывающие электромагнитные явления в вакууме, сохраняют свою форму, если в этих уравнениях, написанных для системы S, произвести замену x, y, z и t на x', y', z' и t', согласно преобразованиям Лоренца. Однако в § 5 части I курса было указано, что законы механики при переходе от одной инерциальной системы отсчета к другой сохраняют свой вид только в том случае, если результаты измерений координат и времени по отношению к этим системам связаны *преобразованиями Галилея*:

$$x' = x - vt; \quad y' = y; \quad z' = z; \quad t' = t. \tag{4}$$

Это расхождение можно объяснить тем, что *законы механических и электромагнитных явлений, по-видимому, основаны на различных способах измерения физических величин,* в частности времени. Если в преобразованиях Лоренца положить c равной бесконечности, то получатся преобразования Галилея. Следовательно, можно утверждать, что в механике Ньютона предполагается бесконечно большая скорость распространения «сигналов времени». Так как не имеет смысла пользоваться в различных областях физики различными способами измерения физических величин, то А. Эйнштейн разработал новую (релятивистскую) механику, которая основана на предложенных им способах измерения физических величин, главным образом времени. При этом несколько видоизменились формулировки некоторых законов физики, но зато открылась возможность объяснения многих явлений единой теорией, а также выявились новые, весьма важные соотношения между физическими величинами (например, между энергией и массой и т. д.).

Приведем важнейшие результаты теории относительности А. Эйнштейна; пользуясь преобразованиями Лоренца, можно простыми расчетами показать, что:

1) расстояния между двумя определенными точками твердого тела, измеренные относительно «покоящейся» (l) и «движущейся» (l') систем отсчета (S и S'), связаны соотношением

$$l' = l\sqrt{1 - v^2/c^2}, \tag{5}$$

т. е. для системы отсчета, относительно которой это тело движется, измеряемое расстояние будет меньше, чем для системы отсчета, в которой это тело покоится;

2) промежутки времени Δt и $\Delta t'$, измеренные относительно S и S', связаны соотношением

$$\Delta t' = \frac{\Delta t}{\sqrt{1 - v^2/c^2}}. \tag{6}$$

Исходя из этой формулы, иногда делают вывод, будто течение времени различно для различных систем отсчета или что движущиеся часы идут медленнее, чем покоящиеся;

3) скорости u и u' и ускорения a и a'(вдоль оси OX), измеренные относительно S и S', связаны соотношениями:

$$u' = \frac{u+v}{1+\frac{uv}{c^2}}; \quad a' = a\frac{(1-v^2/c^2)^{3/2}}{(1-u'v/c^2)^3}. \tag{7}$$

В частности, если $u = c$, то u' также равно c, т. е. скорость света, измеренная относительно S и S', одинакова.

В релятивистской механике А. Эйнштейн получил следующие результаты:

4) если тело, покоящееся относительно S, имеет массу m_0, то при движении со скоростью v его масса увеличивается согласно формуле

$$m = \frac{m_0}{\sqrt{1 - v^2/c^2}}. \tag{8}$$

Эта формула получила экспериментальные подтверждения в опытах с электронами, скорость которых близка к скорости света;

5) основное уравнение механики, сохраняющееся по отношению ко всем инерциальным системам отсчета, должно быть записано в виде

$$F = \frac{\mathrm{d}}{\mathrm{d}t}(mv), \tag{9}$$

где, в отличие от механики Ньютона, масса тела зависит от скорости его движения. Однако эта зависимость может быть замечена только при очень больших скоростях, поэтому релятивистская механика применяется главным образом при исследовании движения электронов и других частиц, скорости которых могут приближаться к скорости света.

Весьма важным для современной физики результатом теории относительности является:

6) соотношение между энергией тела и его массой:

$$E = mc^2 = \frac{m_0c^2}{\sqrt{1 - v^2/c^2}}. \tag{10}$$

Разложим это выражение в ряд и отбросим третий и последующие члены разложения (при $v \ll c$ отброшенными членами разложения можно пренебречь); тогда

$$E = m_0c^2 + \frac{m_0v^2}{2} + \ldots \tag{11}$$

Первый член m_0c^2 представляет собой энергию, связанную с покоящейся массой тела, а второй член — знакомую нам кинетическую энергию движущегося тела:

7) точное выражение для кинетической энергии тела имеет вид

$$E_к = mc^2 - m_0c^2 = m_0c^2\left(\frac{1}{\sqrt{1-v^2/c^2}} - 1\right); \qquad (12)$$

8) импульс тела вычисляется по формуле

$$\mathbf{p} = m\mathbf{v} = \frac{m_0\mathbf{v}}{\sqrt{1-v^2/c^2}}. \qquad (13)$$

При $v \ll c$ $p \approx m_0v$. Компоненты релятивистского импульса по координатным осям в отличие от классических зависят не только от компонент скорости (v_x, v_y, v_z), но и от самой скорости v, например $p_x = \frac{m_0v_x}{\sqrt{1-v^2/c^2}}$ и т. д.

В этих формулах c означает *постоянную величину*, численно равную скорости распространения света в вакууме. Это есть также предельная скорость, при которой масса, импульс и энергия движущегося тела становятся бесконечно большими. Поэтому, пользуясь соотношениями, в которые входит c, следует различать, означает ли c в данном соотношении *универсальную* постоянную или же скорость распространения света в данной среде. Так, например, при расчете дифракционной и интерференционной картины c есть скорость света, тогда как в соотношении

$$E_0 = m_0c^2,$$

(где m_0 и E_0 — масса и энергия *покоющейся* частицы) c может означать только некоторую фундаментальную константу природы.

Заметим, что дифференциал полной энергии частицы (10) равен $v\mathrm{d}p$, но не $p\mathrm{d}v$:

$$v\,\mathrm{d}p = v\,\mathrm{d}\left(\frac{m_0v}{\sqrt{1-v^2/c^2}}\right) = \mathrm{d}\left(\frac{m_0c^2}{\sqrt{1-v^2/c^2}}\right) = c^2\,\mathrm{d}m,$$

тогда как

$$p\,\mathrm{d}v = \frac{m_0v}{\sqrt{1-v^2/c^2}}\,\mathrm{d}v = \mathrm{d}\left(-m_0c^2\sqrt{1-v^2/c^2}\right).$$

Очевидно, что величина B, определяемая из соотношения

$$v\,\mathrm{d}p + p\,\mathrm{d}v = \mathrm{d}(pv) = \mathrm{d}(mv^2); \quad B = mv^2, \qquad (14)$$

может также рассматриваться как некоторая характеристика состояния частицы. Если воспользоваться формулой де Бройля $\lambda = h/mv$ (см. ч. IV, § 10), то

$$B = mv^2 = h\frac{v}{\lambda} = h\nu, \qquad (15)$$

где $v/\lambda = \nu$ имеет размерность частоты колебаний. С другой стороны, энергия одной из важнейших элементарных частиц — фотона — выражается через массу и частоту колебаний аналогичным образом (см. ч. IV, § 1):

$$E = m_ф c^2 = h\nu. \qquad (16)$$

Связь между формулами (15) и (16) будет обсуждаться в § 5.

Соотношения (5)—(13) относятся к «специальной» теории относительности (СТО), в которой предполагается, что физические явления изучаются по отношению к инерциальным системам отсчета. А. Эйнштейн разработал также «общую» теорию относительности (ОТО), применимую и к неинерционным системам отсчета; главной частью этой теории является теория тяготения, изложение которой приводится только в курсах теоретической физики.

§ 3. ЗАКОНЫ СОХРАНЕНИЯ, ПРЕВРАЩЕНИЯ И ПЕРЕРАСПРЕДЕЛЕНИЯ ЭНЕРГИИ В ФИЗИЧЕСКИХ СИСТЕМАХ

Наиболее общая формулировка закона сохранения энергии может быть дана в следующем виде:

если система изолирована от всяких внешних воздействий, то сумма всех видов энергии в системе с течением времени не изменяется;

возможны только превращения одних видов энергии в другие с соблюдением строгих соотношений между их количествами.

Однако применение этого закона требует предварительного обсуждения двух вопросов:

1) каким образом можно в реальных условиях установить и контролировать отсутствие внешних воздействий;

2) каким образом находятся расчетные формулы для различных видов энергии. Очевидно, прежде чем утверждать существование закона сохранения энергии, необходимо дать однозначные и независимые друг от друга определения для «различных видов энергии».

Обсуждение этих вопросов имеет большое теоретическое значение. Можно согласиться с тем, что наиболее определенный (однозначный) метод контроля за внешними воздействиями — это наблюдение за состоянием самой системы. Очевидно, отсутствие внешних воздействий можно было бы установить по постоянству суммы всех видов энергии; в этом случае решение первого вопроса сводится к решению второго.

Таким образом, для понимания сущности закона сохранения энергии необходимо обсуждение вопроса об *определении различных видов энергии*. Понятие энергии впервые появляется при исследовании второго закона механики, когда доказывается, что элементарная работа всех сил, действующих на данное тело, равна

$$\mathrm{d}A = \mathrm{d}\left(\frac{mv^2}{2}\right).$$

Величина $mv^2/2$, названная *кинетической энергией*, является первой формой энергии, установленной в физике. Ее следовало бы полагать *эталонным* видом энергии; однако все остальные виды энергии обнаруживаются не по изменению кинетической энергии, а по-прежнему через работу сил. Так, например, энергия пружины определяется по работе, затрачиваемой на ее деформацию; энергия заряженного конденсатора $W = CU^2/2$ вычисляется по работе сил, преодолевающих куло-

новское отталкивание в процессе зарядки тела; энергия магнитного поля катушки с током $W = LI^2/2$ рассчитывается по работе сил, преодолевающих э. д. с. самоиндукции при возрастании тока от нуля до I и т. п. Некоторые виды энергии определяются через кинетическую энергию поступательного движения (например, энергия вращающегося тела $W = J\omega^2/2$ или внутренняя энергия идеального газа $W = \Sigma m_i v_i^2/2$ и т. п.), но так как кинетическая энергия сама определяется через работу сил, то и эти виды энергии не составляют исключения.

Таким образом, рассмотрев определения всех видов энергии и выводы формул для их расчета, можно убедиться в том, что исходным и весьма общим положением во всех случаях является *равенство изменения энергии и работы сил*:

$$\mathrm{d}W = \mathrm{d}A.$$

Интересно отметить, что даже самое общее определение энергии физической системы также дается через вычисление работы: *энергией физической системы называется величина, изменение которой равно работе внешних сил, приложенных к системе.*

При формулировке закона сохранения энергии используются два весьма важных утверждения:

1) работа $\mathrm{d}A$ в каждом отдельном случае является изменением некоторой потенциальной энергии, которая содержится в полях, возбуждающих силы взаимодействия между телами. Предполагается, что положительная работа сил равна уменьшению потенциальной энергии системы, т. е. $\mathrm{d}A = -\mathrm{d}U$;

2) потенциальные виды энергии могут превращаться в кинетические и обратно с точным соблюдением определенных соотношений между их изменениями.

Важно заметить, что второе утверждение вовсе не является обязательным следствием из опытных данных; можно допустить также, что кинетические и потенциальные виды энергии соответствуют противоположным формам движений в природе и поэтому их одновременные изменения означают не «превращения» между ними, а *взаимосвязанные и сопровождающие друг друга процессы,* в которых соблюдаются строго определенные соотношения между характеризующими их величинами. Благодаря этим соотношениям можно для общей характеристики состояния изучаемой физической системы суммировать все виды энергии независимо от того, существуют ли между ними взаимные превращения или же такие превращения не происходят. Полученная полная энергия для изолированной системы не будет изменяться со временем. Таким образом, физическим основанием для закона сохранения энергии является только существование строго определенных количественных соотношений между изменениями тех величин, которые мы на основании формулы $\mathrm{d}A = -\mathrm{d}U = \mathrm{d}W$ условились называть *энергиями*. Если бы эти соотношения строго не соблюдались, т. е. если данным значениям $\mathrm{d}A$ или $-\mathrm{d}U$ не соответствовали бы всегда и во всех условиях определенные значения $\mathrm{d}W_i$, то суммиро-

вание всех видов энергии системы не дало бы величины, которая для изолированных систем оставалась бы постоянной.

Очень часто *полную* энергию физической системы (W) делят на две части, из которых одна (E) сосредоточена в частицах вещества и связана с их массами соотношением $E_i = m_i c^2$, а вторая (U) распределена в пространстве, окружающем эти частицы, и связана с полями, осуществляющими взаимодействие между ними:

$$W = E + U.$$

В релятивистской физике (см. формулу (11)) энергию E частицы можно представить в виде суммы двух видов энергии: 1) $E_0 = m_0 c^2$, которая не связана с ее движением и существует в любом состоянии, включая и состояние покоя, и 2) $E_{\text{кин}} = m_0 v^2/2 + \ldots$, которая определяется скоростью движения частицы. Вопрос о том, следует ли полагать, что E_0 есть энергия, связанная с *внутренней структурой* частицы и поэтому может изменяться при внешних воздействиях на эту частицу, обычно не обсуждается. Полагают, что не следует вводить физические представления (например, о наличии у элементарных частиц внутренней структуры), которые не могут быть непосредственно доказаны. Однако современная физика уже подошла к экспериментальному изучению внутренней структуры некоторых частиц вещества (протонов и нейтронов), поэтому вопрос о «природе» энергии E_0 (и возможность ее изменения под действием внешних факторов) не должен игнорироваться.

Рассмотрим теперь величину энергии поля U. Рассчитаем сначала энергию, которая содержится, например, в электрическом поле частицы с зарядом e, имеющей форму шара радиуса R (см. ч. III, § 7):

$$U = \int_V \frac{\varepsilon_0 E^2}{2}\, dV, \tag{17}$$

где E — напряженность поля в пределах элементарного объема dV. Интегрирование должно быть произведено по всему объему пространства, за исключением собственного объема самой частицы. Положив $dV = 4\pi r^2 dr$, $E = e/(4\pi\varepsilon_0 r^2)$ и интегрируя от $r = R$ до $r = \infty$, получим

$$U = \frac{e^2}{8\pi\varepsilon_0 R} = k\frac{e^2}{2R}.$$

Эта энергия в основном сосредоточена вблизи частицы; например, в объеме пространства, заключенном между сферами, имеющими $r = R$ и $r = 2R$, содержится столько же энергии, сколько имеется во всем остальном пространстве от $r = 2R$ до $r = \infty$. Ввиду этого величина энергии поля сильно зависит от размера частицы — радиуса R.

В системе, содержащей множество частиц (имеющих различные по величине и знаку электрические заряды), полная энергия поля должна рассчитываться по той же формуле (17). Расчет осложняется тем, что напряженность в каждой точке поля определяется сложением векторов напряженностей, создаваемых отдельными зарядами. Очевидно, величина U будет функцией от радиусов частиц и от расстоя-

ний между ними. Например, для простейшей системы — атома водорода — энергия поля будет определяться радиусами протона R_p, электрона R_e и радиусом орбиты r. Если бы при изменении r (а следовательно, и при изменении интенсивности взаимодействия между частицами) радиусы частиц оставались бы постоянными, то энергию поля можно было бы представить в виде двух членов:

$$U_{\text{поля}} = U_0(R_p,\ R_e) + U(r), \tag{18}$$

где U_0 было бы постоянной величиной.

При этом предположении $U(r)$ является *переменной* частью энергии поля системы; ее обычно и называют потенциальной энергией системы. Расчет показывает, что для системы из двух одинаковых зарядов, находящихся на расстоянии r друг от друга, эта энергия равна

$$U(r) = \frac{e^2}{4\pi\varepsilon_0 r},$$

причем для одноименных зарядов эта энергия будет положительной, а для разноименных (притягивающихся) — отрицательной. Однако утверждение, что размеры частиц при изменении интенсивности взаимодействия остаются постоянными, является *предположением*, которое не имеет обоснования (ни теоретического, ни экспериментального). Если избегать произвольных предположений, то следует полагать, что в формуле (18) обе величины U_0 и U являются переменными и поэтому потенциальная энергия системы как переменная часть энергии поля уже не будет определяться только формулой $U(r)$.

Иногда для определения радиуса частицы (который, разумеется, является условной величиной, так как у нас нет оснований представлять частицу в виде шарика с резко ограниченной поверхностью) приравнивают энергию частицы mc^2 энергии связанного с ним электрического поля:

$$mc^2 = \frac{e^2}{8\pi\varepsilon_0 R};\quad R = \frac{e^2}{8\pi\varepsilon_0 mc^2}.$$

Однако из этого равенства следует, что изменение энергии (массы) частицы (при возрастании или убывании скорости) должно сопровождаться изменением размеров частицы, что не может быть контролируемо экспериментально, следовательно, это равенство является произвольным предположением

Применим теперь изложенные выше рассуждения к простейшей физической системе — атому водорода. Полная энергия этого атома может быть представлена в виде

$$W = E + U = [E_0 + E_к(v)] + [U_0 + U_п(r)], \tag{19}$$

где v — скорость электрона по орбите радиуса r. Допустим, что электрон переходит из удаленной орбиты на ближнюю и атом испускает один фотон с энергией $h\nu$. Если не вводить никаких *произвольных предположений*, то следует полагать, что изменение полной энергии

атома ΔW, равное энергии фотона, может сопровождаться изменениями всех видов энергии атома, т. е. что испускание фотона изменяет все величины, характеризующие состояние атома. Следовательно,

$$h\nu = \Delta W = \Delta E_0 + \Delta U_0 + \Delta E_к + \Delta U_п. \qquad (20)$$

Заметим, что при этом переходе кинетическая энергия электрона увеличивается (см. ч. IV, § 9), поэтому необходимо найти «источник» не только для энергии $h\nu$, но и для $\Delta E_к$. Обычно (и необоснованно) полагают, что $\Delta E_0 = 0$; $\Delta U = 0$ и тогда $\Delta U_п = h\nu + \Delta E_к$. Однако соотношение (20) может соблюдаться и в том случае, если, например, допустить, что $h\nu = \Delta E_0$; $\Delta E_к = (\Delta U_0 + \Delta U_п)$. Возможны и другие варианты интерпретации соотношения (20), в частности и такие, в которых учитывалась бы и энергия, содержащаяся в магнитном поле атома.

Заметим также, что представление о «превращениях энергии», как о некотором физическом процессе, вообще является весьма неопределенным. Мы можем отчетливо представить себе, каким образом происходит превращение упорядоченного движения тел (механической энергии) в беспорядочное движение их молекул (в тепловую энергию) при трении, сжатии и расширении газа в цилиндре с поршнем и т. п. Однако, по существу, в этих процессах никакого превращения энергии нет: происходит только перераспределение механической (кинетической) энергии между частицами рассматриваемой системы. В противоположность этому никакого *отчетливого* представления о том, что означает и каким образом происходит превращение энергии поля в энергию частиц (и обратно), у нас нет.

Резюмируя, можно отметить, что вопрос о взаимных превращениях между «энергией поля» и «энергией частиц вещества» не является окончательно решенным. Возможно, что изменения этих видов энергии лишь *сопровождают* друг друга, сохраняя между собой строго определенные соотношения, причем электрические (и магнитные) поля не обязательно должны быть «резервуарами», поглощающими или выделяющими энергию, а могут играть лишь роль физических факторов, управляющих поведением частиц вещества, способствующих взаимным превращениям между (наблюдаемой и измеряемой) кинетической энергией и скрытой от нас внутренней энергией, сосредоточенной в объеме самих частиц. Такие представления нисколько не умаляют универсального значения всеобщего закона сохранения энергии, но несколько изменяют его формулировку.

Рассмотрим несколько фундаментальных положений классической физики. Допустим, что имеется некоторая физическая система, состояние и свойства которой определяются значениями n физических величин: $x_1, x_2, \ldots, x_n$.

В классической физике предполагается, что:

1) все физические величины, характеризующие изменения в системе, представляют собой *непрерывные функции от времени* $x_i = x_i(t)$. С этим предположением связаны два важных утверждения;

2) в любой точно фиксируемый момент времени t_0 каждая из переменных величин x имеет только одно значение $x = x_0$ (все другие зна-

чения, даже бесконечно близкие к x_0, в этот момент времени *отсутствуют*);

3) время пребывания системы в состоянии с определенными значениями величин x, равно *нулю*.

Далее следует отметить еще два фундаментальных предположения:

4) состояние системы в начальный момент времени и внешние условия, в которых находится система, однозначно предопределяют непрерывную последовательность состояний, через которые проходит система с течением времени;

5) внешнее воздействие на систему и изменение *наблюдаемого* состояния системы происходят одновременно, без какого-либо запаздывания.

Первые предположения кажутся очевидными, однако, строгий анализ показывает, что между утверждениями (2) и (3) существует логическая некорректность: нельзя утверждать, что система *находится* в состоянии с определенными значениями величин x, если время существования в этом состоянии в точности равно нулю. В квантовой же физике утверждается, что в данный момент времени система или частица может иметь непрерывный *спектр* возможных состояний. Такие спектры описываются некоторой функцией распределения $\psi(x)$, так что вероятность реализации состояний, характеризуемых узким интервалом значений x, $x + \mathrm{d}x$, принимается равной

$$\mathrm{d}w = \psi(x)\,\mathrm{d}x.$$

В частности (см ч. IV, § 10—12), согласно соотношениям Гейзенберга, однозначное определение состояния любой элементарной частицы (ее координат и импульса) невозможно; можно только вычислить вероятности того, что частица находится в данном элементарном объеме пространства и имеет скорости, лежащие в заданных пределах.

Таким образом, от весьма ограничивающего предположения (2) следует перейти к более общему предположению, согласно которому:

в каждый определенный момент времени переменная величина x, описывающая состояние системы или частицы, *имеет некоторый непрерывный спектр значений,* характеризуемый функцией распределения $\psi(x)$. Вероятность реализации строго определенного значения x должна быть приравнена нулю, так как время существования этого значения также равно нулю.

Важно подчеркнуть, что применение вероятностных представлений в квантовой физике обусловлено не недостатками методики измерений, а самой природой элементарных частиц, наличием у них корпускулярных и волновых свойств. Сочетанием этих свойств объясняется существование спектра значений у каждой переменной физической величины, характеризующей состояние как отдельной частицы, так и системы, составленной из них.

Допустим, что функция распределения вероятностей в спектре данной физической величины x имеет один максимум при $x = x_0$; это значение будет *наивероятным* для данного момента времени. Если функция $\psi(x)$ имеет очень острый максимум при $x = x_0$, то измерительная аппаратура не сможет обнаружить существование спектра

значений x и поэтому $x = x_0$ будет единственным результатом измерений. Известно, что существование спектра значений обнаруживается при больших скоростях физических процессов (столкновения и др.), поэтому можно утверждать, что вид функции $\psi(x)$ зависит от скорости изменения величины x. Если эти скорости малы, то результатами измерений будут только наивероятные значения физических величин; следовательно, можно утверждать, что *классическая физика*, изучающая сравнительно медленные процессы, *содержит в своих законах и уравнениях движения только наивероятные значения переменных величин*.

Классической иллюстрацией этого утверждения может служить любая термодинамическая система, подвергаемая монотонному внешнему воздействию: сжатию, нагреванию, охлаждению и т. п. При помощи однозначных законов термодинамики можно рассчитать «гладкое» изменение со временем параметров системы (их средних, основных, наивероятных значений); это изменение может быть отмечено грубой измерительной аппаратурой. Более чувствительные приборы будут отмечать флуктуационные отклонения параметров от их «основных» значений. Для этих отклонений могут быть рассчитаны только вероятности реализаций; при этом предполагается, что в данный момент времени каждый переменный параметр имеет некоторый спектр значений, причем различные участки этого спектра имеют различные вероятности реализации.

Квантовой иллюстрацией может служить излучение атомом световой волны. Измерения показали, что каждая линия излучения разреженного газа всегда имеет некоторую «естественную ширину», т. е. очень острый спектр значений частот. Основная (наивероятная) частота линии $\nu = \nu_0$ может быть вычислена при помощи однозначного закона (правила Н. Бора) — по величине энергии уровней:

$$\nu_0 = \frac{E_1 - E_2}{h}.$$

Наличие спектра частот объясняется в квантовой теории (см. ч. IV, § 11) тем, что энергия возбужденных уровней атома несколько «размазаны», т. е. имеют некоторый «спектр значений», ширина которого определяется временем пребывания атома на этом уровне, т. е. быстротой изменения энергии атома в процессе излучения. Различные участки в спектре энергии уровня имеют различные вероятности реализации, что и определяет «форму» (распределение интенсивности между частотами) спектральной линии, т. е. вид функции $\psi(\nu)$.

Существование спектра возможных значений у каждой переменной физической величины (следовательно, спектра возможных состояний системы, находящейся в данный момент времени в определенных внешних условиях и имеющей заданное значение энергии) приводит к выводу, что *однозначное предопределение* последовательности состояний, через которые проходит система или частица с течением времени, *невозможно*. В квантовой физике, рассматривающей явления атомного масштаба, утверждается, что для элементарных частиц вещества имеет смысл рассчитывать только *вероятности* реализации различных про-

цессов. Однако это утверждение имеет более общее значение, т. е. применимо к любым системам и процессам. Только в макрофизических явлениях, ввиду того что функции распределения вероятностей реализации имеют очень острые максимумы, течение физических процессов можно с удовлетворительной точностью описывать одними только наивероятными значениями физических величин, т. е. однозначным образом.

Для анализа четвертого предположения рассмотрим термодинамическую систему, находящуюся в равновесном состоянии. Допустим, что в течение очень короткого времени Δt_1 ей сообщается некоторая энергия ΔW, после чего поступление энергии прекращается. Вследствие процесса перераспределения энергии, система по истечении времени релаксации Δt_2 придет к новому равновесному состоянию. Соотношения между Δt_1 и Δt_2 имеют следующее значение. Если $\Delta t_1 > \Delta t_2$, то система будет «успевать» за внешним воздействием; по окончании поступления энергии ΔW система окажется в новом равновесном состоянии. Если же $\Delta t_2 \gg \Delta t_1$, то можно констатировать, что переход в новое равновесное состояние в основном осуществляется уже без наличия внешнего воздействия, благодаря одним только внутренним процессам перераспределения энергии в системе (теплопроводности, диффузии и т. п.). В этом случае будет наблюдаться некоторое «запаздывание» между внешним воздействием и внутренними релаксационными процессами. Полное отсутствие такого «запаздывания» в процессе поступления энергии ΔW было бы возможно, если бы внутренние процессы перераспределения энергии протекали бы с *бесконечно большими скоростями*; тогда переход из начального равновесного состояния в конечное равновесное состояние состоял бы из непрерывной цепи промежуточных *равновесных* состояний. Заметим также, что у макрофизических систем никаких ограничений на величину сообщаемой энергии ΔW не накладывается (разумеется, если только сообщение этой энергии не приведет к разрушению системы и качественному изменению ее свойств).

При переходе к микрофизическим системам прежде всего необходимо отметить, что они имеют дискретный спектр стационарных состояний, который характеризуется определенными значениями энергии: E_1, E_2, ... Вследствие этого атомы или ядра могут поглощать и испускать только определенные «порции энергии» $\Delta W = E_n - E_k$. Таким образом, взаимодействие между телами, оказывающими «внешнее воздействие», и самой микрофизической системой не может заключаться только в простой передаче любого количества энергии; оно содержит в себе новый элемент — *дозировку* поступающей энергии в соответствии с физическими свойствами самой системы. В частности, если поступающая извне энергия меньше разности между энергией данного и ближайшего стационарного состояний системы, то поглощения энергии не произойдет.

Допустим, что время, в течение которого на микрофизическую систему (атом) оказывается внешнее воздействие (сообщение энергии), равно Δt_1. Если по истечении этого времени система оказалась в новом стационарном состоянии, то можно считать, что время релак-

сации Δt_2 равно Δt_1. Однако для микрофизических систем соотношение $\Delta t_2 \gg \Delta t_1$ не исключено. Если в системе протекают процессы, аналогичные релаксационным процессам в термодинамических системах, то разность $\Delta t_2 - \Delta t_1$ будет зависеть от скорости этих внутренних процессов. Очевидно, $\Delta t_2 = \Delta t_1$ в том случае, если эти релаксационные процессы происходят с бесконечно большой скоростью.

Отдельные (свободные) элементарные частицы по своим основным свойствам могут быть уподоблены либо макрофизическим, либо микрофизическим системам. В первом случае частица может иметь непрерывное множество устойчивых (равновесных) состояний, соответствующих непрерывному спектру значений ее энергии; такой частице можно сообщить любую порцию энергии ΔW. Во втором случае частица будет иметь дискретный (возможно, очень «густой») спектр устойчивых (стационарных) состояний с энергиями E_1, E_2, ... Такая частица может поглощать и испускать энергию, равную $E_п - E_к$. Для обоих представлений важное значение имеет отсутствие или наличие некоторого «запаздывания» между внешним воздействием на частицу и установлением нового устойчивого состояния.

Рассмотрим движение элементарной частицы с зарядом e и массой m в неоднородном электрическом поле. В классической физике предполагается, что в данной точке поля с напряженностью $\mathbf{E}$ ускорение частицы в точности равно $a = \frac{e}{m} E$. Однако *время пребывания частицы в каждой точке поля равно нулю,* поэтому указанное предположение допустимо, если только действие поля на частицу происходит *мгновенно*, без какого-либо запаздывания, т. е. с бесконечно большой скоростью. Поясним это утверждение.

Допустим, что заряженная частица (например, электрон) пробежала в ускоряющем электрическом поле с напряженностью E некоторый участок траектории Δx с разностью потенциалов $\Delta\varphi = E\Delta x$. Обозначим наблюдаемое при этом изменение кинетической энергии частицы через ΔW, а время пребывания — через $\Delta t = \frac{\Delta x}{v}$, где v — средняя скорость на этом участке. Измерения показали, что при этом соблюдаются следующие соотношения релятивистской физики:

$$e\,\mathrm{d}\varphi = \mathrm{d}\,(mc^2) \quad \text{или} \quad eE\,\mathrm{d}x = \mathrm{d}\left(\frac{m_0c^2}{\sqrt{1 - v^2/c^2}}\right), \tag{21}$$

откуда следует, что

$$\frac{\mathrm{d}v}{\mathrm{d}t} = \frac{eE}{m_0}\left(1 - \frac{v^2}{c^2}\right)^{3/2}; \quad \mathrm{d}\left(\frac{m_0v^2}{2}\right) = \left(1 - \frac{v^2}{c^2}\right)^{3/2} e\,\mathrm{d}\varphi. \tag{22}$$

Формула (21) интерпретируется как равенство между работой поля $\mathrm{d}A = e\mathrm{d}\varphi$ и изменением энергии частицы $\mathrm{d}W = \mathrm{d}\,(mc^2)$. Одновременно предполагается, что источником появившейся кинетической энергии частицы является энергия, содержащаяся в ускоряющем поле. Следовательно, на участке Δx должны происходить два процесса: 1) переход некоторого количества энергии поля (которая *непрерыв-*

ным образом распределена в окружающем пространстве) внутрь объема частицы и 2) преобразование этой энергии в кинетическую энергию частицы. Оба эти процесса, как и любые физические процессы в природе, должны протекать с *конечными* скоростями, поэтому время пребывания частицы на участке Δx должно быть достаточным для того, чтобы указанный процесс концентрации и преобразования энергии ΔW мог закончиться. Это условие может соблюдаться при малых скоростях частицы. Однако при очень больших скоростях v время Δt может оказаться значительно меньше того времени $\Delta\tau$, которое необходимо для передачи энергии ΔW от поля к частице и преобразования ее в кинетическую форму. Поэтому важно отметить еще одно предположение, которое используется при обычной трактовке соотношений (21): предполагается, что воздействие поля на частицу происходит с бесконечно большой скоростью, поэтому работа поля на участке Δx равна $e\Delta\varphi$ при *любой скорости* движения частицы на этом участке.

Однако предположение о том, что преобразование «энергии поля» в «энергию частицы» происходит с бесконечно большой скоростью, не является обязательным при трактовке экспериментальных соотношений (21). Можно, например, допустить, что воздействие поля на частицу в пределах участка Δx *зависит от скорости*, с которой частица пробегает этот участок; в частности, можно полагать, что эта работа равна не $e\Delta\varphi$, а должна приравниваться $e\Delta\varphi\,(1-v^2/c^2)^{3/2}$. Тогда, рассматривая формулы (21), (22) как выражение универсального равенства между работой сил и изменением кинетической энергии частицы, можно утверждать, что ΔW равно не $\Delta\,(mc^2)$, а $\Delta\,(m_0v^2/2)$.

Итак, при любой трактовке экспериментальных соотношений (21), (22) следует полагать, что работа поля на участке Δx всегда равна изменению энергии частицы. Однако если дополнительно полагать, что эта работа равна $e\Delta\varphi$ *независимо* от скорости частицы, то тогда энергией частицы следует полагать произведение mc^2. В этом случае при $v\to c$ масса и энергия частицы будут стремиться к бесконечно большим значениям. Если же допустить, что работа поля равна $e\Delta\varphi\,(1-v^2/c^2)^{3/2}$, то можно сохранить классическое выражение для кинетической энергии $m_0v^2/2$; тогда масса может полагаться постоянной, а энергия не будет достигать бесконечно больших значений при $v\to c$. При этих предположениях скорость света в вакууме есть та предельная скорость частицы, при которой поле теряет способность изменять ее кинетическую энергию.

Однако не обязательным оказывается также и предположение, что «источником» кинетической энергии частицы является энергия, содержащаяся в ускоряющем поле. В классической и релятивистской физике обычно игнорируется внутреннее состояние частиц; допускается, что частица может изменять свою скорость и кинетическую энергию без каких-либо *обязательных* внутренних изменений. Такой подход основан на явном желании не вводить в физическую теорию предположений о процессах, существование которых невозможно непосредственно обнаружить, а их течение — изучить при помощи измерительной аппаратуры. Однако общие представления о природе, сложившиеся к настоящему времени, свидетельствуют о том, что

всякое физическое воздействие на тела всегда сопровождается определенными изменениями в их внутреннем состоянии.

Этот вывод следует не только из наблюдений за макроскопическими телами; им руководствуются и при изучении микрофизических явлений, в которых участвуют отдельные сложные молекулы или атомы. Поэтому если элементарную частицу не полагать внутренне неизменяемой и представлять ее в виде физического тела, имеющего какое-то «внутреннее устройство», то тогда открывается возможность другой трактовки явлений, в которых поля действуют на частицы. Можно, например, полагать, что изменение кинетической энергии частицы происходит не за счет энергии поля, а вследствие изменений во внутреннем состоянии частицы. Воздействие поля может заключаться только в возбуждении этого процесса и в определении его интенсивности. Тогда не будет необходимости в весьма искусственных представлениях о возможности перехода «размазанной» в пространстве энергии поля в малый объем частицы при ускоренном движении и об обратном переходе кинетической энергии, концентрированной в объеме частицы, в распределенную энергию поля при замедленном движении. Следует также заметить, что внутреннее состояние частицы может зависеть от напряженности поля, действующего на частицу, независимо от того, покоится или движется частица в этом поле. Возможно также, что внутреннее состояние частицы определяет их некоторые наблюдаемые свойства. Тогда изменение «массы покоя» протона и нейтрона при переходе внутрь атомного ядра можно трактовать как результат воздействия сильного ядерного поля на их внутреннее состояние.

При более подробном рассмотрении движения заряженных частиц в ускоряющих или замедляющих полях необходимо интересоваться также и изменениями, происходящими в «собственном» электрическом и магнитном поле самих частиц. Особый интерес представляет *магнитное поле*; полагают, что собственное магнитное поле движущейся частицы также содержит в себе энергию (равную $\int_V \frac{\mu_0 H^2}{2} dV$; см. формулу (3.74), ч. III). Прежде всего заметим, что *внешнее* магнитное поле действует на заряженную частицу с силой Лоренца, которая всегда перпендикулярна скорости движения, поэтому *работы не совершает*. Следовательно, магнитное поле не может изменять кинетическую энергию частиц и возможность превращения энергии этого поля в энергию частиц не очевидна. Далее, при возрастании скорости частицы под действием электрического поля увеличиваются также напряженности магнитного поля вокруг частицы, т. е. возрастает энергия ее собственного магнитного поля. Таким образом, необходимо искать «источник» не только для кинетической энергии частицы, но и для энергии ее собственного магнитного поля. Этим источником не может быть работа электрического поля, так как она в точности равна $\Delta W_к$. Следовательно, предположение об участии в этих явлениях внутренней энергии частиц ($W_{внутр}$) представляется совершенно необ-

ходимым (в частности, возможно, что $\Delta W_к = -\Delta W_{внутр}$; $\Delta U_{маг} = -\Delta U_{элек}$).

В современной физике заметна тенденция к изучению внутренней структуры важнейших элементарных частиц (протона, нейтрона, электрона); дальнейший прогресс должен заключаться в изучении процессов, которые происходят в этой структуре при изменении энергии частицы. Предположение о том, что изменение кинетической энергии (или скорости) частицы возможно при полном отсутствии каких-либо внутренних изменений в структуре или внутреннем состоянии частицы, является одним из основных положений классической физики, требующим тщательного анализа. Более общие предположения о наличии таких «внутренних процессов», о конечной скорости их течения, о существовании дискретного спектра устойчивых (стационарных) состояний свободной частицы должны быть проверены в ходе дальнейшего прогресса теоретической и экспериментальной физики.

§ 4. О ВАРИАЦИОННЫХ ПРИНЦИПАХ ФИЗИКИ

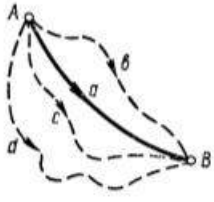
Изучение физических объектов состоит из нахождения:

1) состава и структуры объекта, физических свойств частиц, из которых состоит объект, и самого объекта в целом;

2) законов, управляющих поведением объекта в зависимости от данных внешних воздействий;

3) последовательности состояний, проходимых объектом с течением времени.

Процесс исследования заключается в том, чтобы по найденным сведениям определить неизвестное. Например, при экспериментальном изучении данного объекта можно найти его состав, структуру, свойства и последовательность пробегаемых состояний; после этого, анализируя полученные сведения, можно выявить те законы, которые управляют поведением объекта и определяют течение рассматриваемого процесса. В другом случае можно воспользоваться известными законами и на основании экспериментально наблюдаемого течения изучаемого процесса вынести некоторые заключения о составе и структуре объекта и т. п.



Рассмотрим следующую весьма общую задачу. Допустим, что на рисунке точки A и B условно изображают начальное и конечное состояния некоторой физической системы, а сплошная линия AB — процесс перехода, т. е. последовательность состояний, пробегаемых системой при переходе из A в B. Далее допустим, что пунктирные линии (b, c, d, ...) означают последовательности состояний, которые разрешаются общими законами сохранения (энергии, импульса, момента импульса и др.), но которые не реализуются в действительности (здесь необходимо отметить, что законы сохранения однозначно не предопределяют направление и интенсивность происходящих в природе процессов, т. е. являются не *управляющими* законами, а *соблюдающимися*

законами). Возникает интересная и важная проблема: найти способ выделения *реализующегося* процесса перехода (*a*) *от возможных* (*b*, *c*, *d*, ...), допускаемых законами сохранения, если конкретные законы, определяющие эти процессы, нам неизвестны. По существу, это означает, что мы хотим найти законы, управляющие процессом перехода. Очевидно, что при решении такой задачи надо исходить из каких-то новых положений, использовать какие-то новые предположения, новые свойства системы; одно только задание начального и конечного состояний системы и законов сохранения недостаточно. Решение подобных задач найдено на основе *вариационного принципа наименьшего действия Гамильтона*. Постулируется, что для каждой физической системы можно найти такую функцию состояния системы $L(x_1, x_2, \ldots)$, имеющую размерность энергии, при помощи которой реальные процессы, происходящие в этой системе, выделяются среди других (допускаемых свойствами системы и условиями, в которых она находится) *минимумом* интеграла действия:

$$S = \int_1^2 L(x_1,\ x_2,\ \ldots)\,\mathrm{d}t. \tag{23}$$

В вариационном исчислении доказано, что требование минимума (или экстремума) интеграла (23) выполняется, если для каждой величины x_i соблюдается соотношение

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial v_i} - \frac{\partial L}{\partial x_i} = 0, \tag{24}$$

где $v_i = \mathrm{d}x_i/\mathrm{d}t$. Эти соотношения и выражают искомые *конкретные* законы, управляющие физическими процессами, которые происходят в системе с заданным видом функции $L(x_i)$. Эта функция называется *лагранжианом* данной системы. В теоретической физике показано, что для многих физических систем имеется общее выражение для лагранжиана

$$L = T - U, \tag{25}$$

где T — сумма кинетических энергий всех частиц данной системы, а U — потенциальная энергия системы. Таким образом, согласно принципу наименьшего действия, существует весьма общее выражение (25) для лагранжиана; написав для данной физической системы конкретную форму этой функции состояния, можно по формулам (24) найти законы, действующие в рассматриваемой системе.

Допустим, что система — механическая, состоящая в простейшем случае из двух тел с кинетическими энергиями $\frac{m_1 v_1^2}{2} + \frac{m_2 v_2^2}{2}$ и потенциальной энергией гравитационного взаимодействия $U = -G\frac{m_1 m_2}{r}$ (где $r = x_1 + x_2$, а x_1 и x_2 — расстояния от тел до центра масс этой системы). Лагранжиан этой системы будет равен

$$L = T - U = \frac{m_1 v_1^2}{2} + \frac{m_2 v_2^2}{2} + G\frac{m_1 m_2}{x_1 + x_2}.$$

Воспользуемся соотношениями (24):

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial v_1} - \frac{\partial L}{\partial x_1} = \frac{\mathrm{d}}{\mathrm{d}t}(m_1 v_1 + 0) - G\frac{m_1 m_2}{(x_1 + x_2)^2} = 0.$$

Так как сила гравитационного взаимодействия между телами равна $F = G\frac{m_1 m_2}{r^2}$, то из этого соотношения следует

$$F_1 = \frac{\mathrm{d}}{\mathrm{d}t}(m_1 v_1).$$

Аналогичная формула будет получена и для второго тела. Таким образом, из принципа наименьшего действия и при выборе подходящего выражения для лагранжиана получается основной закон механики, однозначно определяющий течение механических процессов в системе.

В теоретической физике показано, что лагранжиан частицы (например, электрона), имеющей заряд e, массу m и находящейся в электрическом поле с потенциалом φ (определяющим напряженность E) и магнитном поле с вектор-потенциалом $\mathbf{A}$ (определяющим напряженность H), имеет вид

$$L = -m_0 c^2 \sqrt{1 - v^2/c^2} - e\varphi + \frac{e}{c}\mathbf{Av}. \qquad (26)$$

Эту формулу можно переписать в виде

$$L = (mv^2 - mc^2) - e\varphi + \frac{e}{c}\mathbf{Av} = \left(mv^2 + \frac{e}{c}\mathbf{Av}\right) - (mc^2 + e\varphi). \qquad (27)$$

Заметим, что в первой скобке содержатся величины, связанные со скоростью частицы (эти величины становятся равными нулю при $v = 0$), а во второй скобке — величины, определяющие энергию частицы, как движущейся, так и покоящейся. Так как скорость и импульс частицы определяют «волновое» поведение частицы (см. ч. IV, § 10, 11), а энергия — «корпускулярное», то можно утверждать, что лагранжиан частицы выражает корпускулярно-волновой дуализм элементарных частиц.

Выше (см. формулу (15)) было показано, что величина mv^2 может быть представлена через дебройлевскую длину волны $\lambda = h/mv$ или через некоторую частоту колебаний $\nu = v/\lambda$ (которую будем называть *дебройлевской*). Известны опыты, показывающие, что магнитное поле смещает интерференционную картину, образованную электронами после прохождения через щели, причем в пространстве, в котором перемещались электроны, напряженность магнитного поля была равна нулю, но вектор-потенциал $\mathbf{A} \neq 0$ (см., например, «Фейнмановские лекции по физике», 1965, т. 6, с. 23—26). Таким образом, обе величины mv^2 и $\frac{e}{c}\mathbf{Av}$, написанные в первой скобке, связаны с дебройлевскими волнами. Во второй скобке оказались величины, определяющие: 1) энергию частицы в электрическом поле $e\varphi$ и 2) энергию mc^2, не связанную с полями (напомним, что магнитное поле не может изменить энергию частицы, так как действует на них с силой, всегда перпендикулярной скорости движения; см. ч. III, § 19, 20). Если обе эти скобки по аналогии с (15) и (16) выразить через частоты колебаний («дебройлевских» — $\nu_{\text{деб}}$ и «энергетических» — $\nu_{\text{эн}}$), то

$$h\nu_{\text{деб}} = mv^2 + \frac{e}{c}\mathbf{Av}; \quad h\nu_{\text{эн}} = mc^2 + e\varphi. \qquad (28)$$

Тогда лагранжиан (27) частицы, движущейся в электрическом и магнитном полях, может быть записан в простом виде, явно выражаю-

щем волновые ($h\nu_{деб} = v/\lambda$) и корпускулярные ($h\nu_{эн}$) свойства этой частицы:

$$L = h\nu_{деб} - h\nu_{эн}. \tag{29}$$

Допустим, что имеется физическая система, внутри которой каждая отдельная частица находится в суммарном электрическом поле других частиц, поэтому для одной частицы можно написать:

$$h\nu_{эн} = mc^2 + \Delta U \tag{30}$$

(где ΔU учитывает действие электрического поля на данную частицу), а для системы из n взаимодействующих частиц

$$\sum_n h\nu_{эн} = \sum_n mc^2 + U. \tag{31}$$

Очевидно, что величину U можно отождествить с полной энергией электрического поля, существующего в системе, или же с ее переменной частью — потенциальной энергией системы. Таким образом, обычное деление полной энергии системы на кинетическую и потенциальную составляющие означает раздельный учет зависимости частоты колебаний от массы частиц и от действия силовых полей.

Выберем некоторый промежуток времени Δt и рассчитаем общее число «энергетических» и «дебройлевских» колебаний в системе за это время. Имеем:

$$N_1 = \sum \nu_{эн}\Delta t = \frac{\Delta t}{h}\left(\sum mc^2 + U\right),$$

$$N_2 = \sum \nu_{деб}\Delta t = \frac{\Delta t}{h}\sum mv^2.$$

Разность между ними

$$N_2 - N_1 = \frac{\Delta t}{h}\left(\sum mv^2 - \sum mc^2 - U\right) =$$

$$= \frac{\Delta t}{h}\left(-\sum m_0c^2\sqrt{1 - v^2/c^2} - U\right) = \frac{\Delta t}{h}(T - U),$$

где обозначено

$$T = -\sum m_0c^2\sqrt{1 - v^2/c^2}. \tag{32}$$

При малых скоростях движения переменная часть этой величины равна сумме кинетических энергий частиц в системе:

$$T_к \approx \sum \frac{m_0v^2}{2},$$

поэтому выражение

$$L = \sum (h\nu_{деб} - h\nu_{эн}) = T - U \tag{33}$$

совпадает с общей формулой для лагранжиана физических систем (25). Ввиду этого содержание принципа наименьшего действия Гамильтона сводится к утверждению, что реальные процессы, происходящие в любой физической системе, выделяются *минимальным* значением разности между числами «дебройлевских» и «энергетических» колебаний.

Если разность $N_2 - N_1 = \Delta N$ характеризует изменения, которые произошли в системе за время Δt, то отношение $\Delta N/\Delta t$ можно рассматривать как среднюю скорость изменения состояния системы.

В этом смысле лагранжиан является скоростью изменения состояния системы в данный момент времени (следует заметить, что для каждой функции состояния, введенной для описания данной системы, будет существовать своя «скорость изменения»; очевидно, что между этими скоростями, как и между самими функциями состояния, можно найти определенные соотношения).

В заключение заметим, что в микрофизических явлениях обнаруживается явная тенденция к уменьшению разности между частотами колебаний $\nu_{эн}$ и $\nu_{деб}$. Для отдельной частицы эта разность равна

$$\nu_{эн} - \nu_{деб} = m_0 \frac{c^2}{h} \sqrt{1 - v^2/c^2} \qquad (34)$$

и зависит от массы покоя и скорости движения. Приведем несколько примеров:

1) при свободном движении заряженной частицы в электрическом поле его скорость возрастает, поэтому разность частот (34) убывает;

2) при переходе атома из возбужденных состояний в нормальное скорости электронов, вращающихся по стационарным орбитам, возрастают;

3) при формировании физических систем (атомов, ядер и т. п.) из свободных частиц массы покоя частиц уменьшаются, а их скорости движения возрастают;

4) при аннигиляции частицы и античастицы образуются фотоны и нейтрино, у которых собственные массы равны нулю, а скорости имеют предельные значения.

§ 5. СТАТИСТИЧЕСКИЕ ЗАКОНЫ И ФЛУКТУАЦИИ

Статистическая физика предназначена для изучения процессов в системах из большого числа частиц, совершающих беспорядочное тепловое движение. Ее задачами являются: 1) нахождение связи между свойствами системы и течением процессов в ней (которые фиксируются при помощи физической аппаратуры) и свойствами частиц и характером их движения внутри системы; 2) нахождение и сопоставление средних значений случайных величин за большое время наблюдения; 3) расчет величины и частоты появления случайных (флуктуационных) отклонений данной измеряемой физической величины от ее среднего (равновесного) значения. Значительная часть статистической физики посвящена термодинамическим системам и является теоретической основой «феноменологической» термодинамики.

В статистической физике используется аппарат математической теории вероятностей к многократно повторяющимся физическим состояниям или явлениям. Типичным примером является рассмотренный ранее закон распада радиоактивных ядер. Полагая, что вероятности распада для каждого ядра в отдельности равны, можно принять число распадающихся ядер за время dt, пропорциональным dt и числу имеющихся ядер N. Заметим, что полученные и для других явлений экспоненциальные законы также имеют вероятностный смысл. Ввиду этого определение понятия «вероятность события» имеет в статистической физике важное значение.

Допустим, что произведено N однородных (по какому-нибудь признаку) испытаний, например: бросание игральной кости, многократное измерение какой-нибудь величины x (температуры, давления и т. п.) в данном месте в течение очень малых (но равных) промежутков времени Δt и т. д. Если значение $x = x_i$ наблюдалось n_i раз, то отношение n_i/N есть *частота появления* этого значения, а предел

$$w = \lim_{N\to\infty} \frac{n_i}{N} \tag{35}$$

(если он существует) называется *вероятностью появления* значения $x = x_i$. Для величин, непрерывно изменяющихся со временем, $n_i\Delta t = \tau_i$ — есть время существования значения x_i и поэтому вероятность реализации этого значения равна

$$w = \lim_{t\to\infty} \frac{\tau_i}{t}, \tag{36}$$

где $t = N\Delta t$ — полное время наблюдения. В некоторых случаях о вероятности событий можно судить до производства испытаний. Например, если игральная кость — идеальный кубик, то вероятности появления любой цифры будут равны $^1/_6$.

Для случайных величин, имеющих непрерывный спектр возможных значений (скорости молекул газа, частота колебаний излучения и др.), вводят *функции распределения*; вероятность того, что величина x может иметь значения, лежащие в пределах x, $x + \mathrm{d}x$, представляется в виде

$$\mathrm{d}w = f(x)\,\mathrm{d}x.$$

Если $f(x)$ имеет один максимум при $x = x_0$, то x_0 называют *наивероятным* значением x. Знание $f(x)$ позволяет вычислять средние значения переменных величин:

$$\langle x\rangle = \int xf(x)\,\mathrm{d}x$$

(по совокупности частиц, если x характеризует одну частицу, и по времени, если $x = x(t)$).

Функции распределения $f(x)$ имеют в статистической физике фундаментальные значения. В частности, у системы из N одинаковых частиц, имеющих дискретный спектр состояний с энергиями ε_1, ε_2, ..., равновесная функция распределения частиц по энергиям представляется в виде

$$f(\varepsilon_i) = \frac{\mathrm{e}^{-\mu\varepsilon_i}}{\sum \mathrm{e}^{-\mu\varepsilon_i}}, \tag{37}$$

где μ — постоянная величина (зависящая от температуры системы).

Переход системы из неравновесного состояния в равновесное можно обосновать различным образом. В частности допустим, что среднее время пребывания τ_i частицы в состоянии с энергией ε_i убывает с увеличением энергии и выражается формулой

$$\frac{1}{\tau_i} = \alpha + \beta\varepsilon_i, \tag{38}$$

где α и β — величины, зависящие от общего состояния системы, т. е. от вида функции распределения $f(\varepsilon_i)$. Если на уровне ε_i находится N_i

частиц, то N_i/τ_i будет означать число частиц, ежесекундно покидающих этот уровень. Обозначим через Z_i — число частиц, ежесекудно поступающих на этот уровень. Очевидно, что разность

$$\left| Z_i - \frac{N_i}{\tau_i} \right| \tag{39}$$

в равновесном состоянии должна равняться нулю. Величина Z_i зависит от общего состояния системы (от того, что в данный момент времени происходит на всех уровнях энергии) и поэтому будет изменяться сравнительно медленно, тогда как N_i/τ_i прямо пропорциональна N_i и будет изменяться гораздо быстрее, чем Z_i. Этого достаточно, чтобы разность (39) имела тенденцию к уменьшению до нуля, что означает переход системы к равновесному состоянию.

Заметим, что общее число энергетических переходов в системе (в единицу времени)

$$P = \sum \frac{N_i}{\tau_i} = N \sum \frac{f(\varepsilon_i)}{\tau_i} \tag{40}$$

является важной характеристикой состояния системы (имеющей смысл и для неравновесных состояний). В частности, равновесную функцию распределения частиц по уровням энергии (37) можно вывести из условия $P \to$ минимуму. Для идеального газа P есть число столкновений между молекулами в единицу времени и поэтому условие $P \to$ минимуму (см. формулу (2.76), ч. II), означает возрастание энтропии. Возможно, что это утверждение имеет смысл и для других систем, для которых число энергетических переходов может быть рассчитано.

Важной особенностью физических систем является наличие в них «борьбы» между хаосом теплового движения и односторонним действием релаксационных процессов, направленных к достижению равновесного состояния. В точном состоянии равновесия релаксационные процессы исчезают и система освобождается от их действия, поэтому случайные ситуации, возникающие в системе вследствие беспорядочного теплового движения частиц, могут легко вывести систему из равновесия. Тогда приходят в действие релаксационные процессы, вновь направляющие систему в равновесное состояние.

Флуктуациями называются беспорядочные отклонения физических систем от их равновесного состояния (или физических процессов — от их установившегося течения). Флуктуации существуют и в неравновесных состояниях и в неустановившихся процессах; при их отсутствии релаксация была бы «гладким» процессом и их можно было бы описывать однозначными функциями от времени. Наличие тепловых флуктуаций вызывает беспорядочные отклонения реальных процессов от такого «гладкого» течения.

Рассмотрим флуктуации в состоянии равновесия. Обозначим через x — значение флуктуирующей величины в некоторый момент времени, а через x_0 — его среднее значение за большое время наблюдения. Разность $x - x_0$ будет весьма быстро и беспорядочно изменяться и по величине и по знаку; ее среднее значение, найденное для достаточно большого времени, равно, очевидно, нулю. Для практических целей удобно оценивать флуктуации по среднему значению квадрата

этой разности: $|x - x_0|^2_{\text{ср}}$. Квадратный корень от этой величины обозначим через Δx. Отношение $\Delta x/x_0$ называют *относительной флуктуацией*.

Формулы для расчета флуктуаций различных физических величин выводятся в статистической физике; приведем некоторые из них:

1) флуктуация числа части N в объеме V (выделенном в газе, жидкости или твердом теле) при температуре T

$$(\Delta N)^2 = -\frac{kTN^2}{V^2}\left(\frac{\partial V}{\partial p}\right)_{T=\text{const}} \tag{41}$$

или относительная флуктуация

$$\frac{\Delta N}{N} = \sqrt{-\frac{kT}{V^2}\left(\frac{\partial V}{\partial p}\right)_{T=\text{const}}} \tag{42}$$

(k — постоянная Больцмана).

Если среда — идеальный газ, то $pV = NkT$ и поэтому

$$\left(\frac{\partial V}{\partial p}\right)_{T=\text{const}} = -\frac{V^2}{NkT};\quad \frac{\Delta N}{N} = \frac{1}{\sqrt{N}}. \tag{43}$$

В кубическом сантиметре воздуха при нормальных условиях содержится около $2{,}69 \cdot 10^{19}$ молекул и относительная флуктуация оказывается равной $2 \cdot 10^{-8}$, т. е. около двух миллионных долей процента. При малых плотностях относительная флуктуация может быть больше;

2) флуктуация температуры в данном объеме среды

$$(\Delta T)^2 = \frac{kT^2}{C_V};\quad \frac{\Delta T}{T} = \sqrt{\frac{k}{C_V}}. \tag{44}$$

У идеального газа теплоемкость при постоянном объеме (для всего газа, содержащего N частиц) выражается простой формулой $C_V = ikN/2$ и поэтому относительная флуктуация температуры обратно пропорциональна $\sqrt{N}$;

3) флуктуация энергии в данном объеме среды, содержащем N частиц,

$$(\Delta E)^2 = kT^2C_V; \tag{45}$$

4) флуктуация числа n электронов, вылетающих из катода к аноду в электронных лампах за определенное время t,

$$(\Delta n)^2 = n;\quad \frac{\Delta n}{n} = \frac{1}{\sqrt{n}}. \tag{46}$$

Так как $ne/t = I$ (e — заряд электрона) есть сила тока I, то относительная флуктуация анодного тока вычисляется по формуле

$$\frac{\Delta I}{I} = \sqrt{\frac{e}{It}}. \tag{47}$$

В радиотехнических цепях флуктуации токов обнаруживаются в виде так называемых «тепловых шумов»;

5) флуктуация напряжения U на концах проводника с сопротивлением R при температуре T (эти напряжения появляются вследствие случайных перераспределений электронного газа в объеме проводника). Расчет показывает, что $(\Delta U)^2$ пропорциональна kT и R;

6) беспорядочные колебания стрелок измерительных приборов (гальванометров и др.) вокруг нулевого положения, случайные откло-

нения тела, подвешенного на пружине, от равновесного состояния и т. д. определяются из равенства между энергией этого тела и энергией теплового движения молекул, приходящейся на одну степень свободы. Например, в колебательной системе из груза с массой m и пружины с упругостью $k_{\text{упр}}$, находящейся в идеальном одноатомном газе с температурой T, флуктуации кинетической и потенциальной энергии вследствие ударов молекул будут равны:

$$\frac{k_{\text{упр}}(\Delta x)^2}{2} = \frac{m(\Delta v)^2}{2} = \frac{1}{2}kT, \tag{48}$$

откуда следует, что $(\Delta x)^2$ обратно пропорциональна упругости пружины, а $(\Delta v)^2$ — массе груза; кроме того, обе эти флуктуации прямо пропорциональны температуре среды.

Флуктуационные отклонения указывают на предел, до которого имеет смысл повышать чувствительность измерительных приборов.

§ 6. О ДУАЛИЗМЕ В ФИЗИКЕ

Термин «дуализм» в физике в широком смысле означает:

1) существование противоположных свойств у физических объектов;

2) использование противоположных понятий при описании и объяснении физических явлений;

3) наличие противоположных (взаимоисключающих) утверждений в формулировке законов, управляющих физическими явлениями.

Наиболее фундаментальными проявлениями дуализма являются:

1) корпускулярно-волновой дуализм в свойствах элементарных частиц;

2) наличие в природе частиц и античастиц, противоположных электрических зарядов, различного знака лептонных и барионных чисел (см. ч. IV, § 23) и др.;

3) противоположные свойства у частиц вещества и у силовых полей, т. е. у «корпускулярной» и «полевой» материи;

4) использование понятий «энергия» и «работа»;

5) существование в физических системах сил отталкивания и сил притяжения, одновременное действие которых определяет свойства физических систем;

6) связь между количественными и качественными изменениями в свойствах физических систем;

7) однозначность и вероятность в законах физики;

8) дискретность и непрерывность в природе, связь между ними и т. д.

Сущность дуализма (т. е. содержание терминов «противоположные свойства», «понятия», «утверждения») может быть показана на примере сочетания корпускулярных и волновых свойств у элементарных частиц (фотонов, электронов и т. д.). В тексте (см. ч. IV, § 10—12) было показано, что:

1) корпускулярные и волновые свойства частиц *неотделимы* друг от друга. Каждая частица имеет оба эти свойства в единстве и взаим-

ной обусловленности, причем нет никакой возможности лишить частицу одного из этих свойств. По-видимому, не существуют частицы, обладающие только корпускулярными или только волновыми свойствами;

2) корпускулярные и волновые свойства *несводимы* друг к другу. Это означает, что волновые свойства частицы нельзя объяснить через корпускулярные, и наоборот;

3) корпускулярные и волновые свойства *неразрывно связаны* между собой.

Корпускулярно-волновой дуализм лежит в основе квантовой физики, описывающей микрофизические системы и процессы. Таким образом, один из важнейших разделов современной физики является дуалистическим по своему характеру и содержанию. Непрерывная волновая функция частиц и физических систем, с одной стороны, корпускулярные свойства этих же частиц и систем — с другой, существуют в квантовой физике в единстве и взаимной связи. Все попытки устранить этот дуализм успеха не имели. Поэтому можно утверждать, что дуализм в квантовой теории есть не временное, случайное, побочное явление, вызванное, например, трудностями описания микрофизических систем, а отражение господствующего в природе объективного дуализма.

Рассмотрим другое проявление дуализма в природе — наличие частиц и античастиц. Предварительно заметим, что физические свойства частиц можно условно разделить на две группы:

1) свойства, которые у различных частиц отличаются *только по величине*; к важнейшим из них относится инертная масса. Заметим, что масса не является аддитивным свойством (масса физической системы меньше суммы масс составных частиц, измеренных в свободном состоянии), зависит от состояния частицы (скорости движения) и от условий, в которых находится частица (масса нуклонов в поле ядерных сил отличается от их масс вне ядра);

2) свойства, отличающиеся *качественно*, например противоположные электрические заряды. Заметим, что заряды обладают *аддитивностью*, не зависят от скорости движения и от условий, в которых находятся заряженные частицы. Это означает, что заряды (а также и лептонные и барионные числа) являются более *фундаментальными* свойствами частиц, чем инертная масса.

Элементарные частицы могут сортироваться по набору присущих им фундаментальных свойств. В зависимости от характера и числа этих свойств определяется содержание таких понятий, как «одинаковые» или «различные» частицы. Очевидно, что *тождественность* частиц (или вообще физических объектов) есть предельный случай одинаковости, когда между объектами нет никакого различия: ни в наборе присущих им свойств, ни в их структуре, состоянии и поведении в различных условиях (такими тождественными объектами являются элементарные частицы определенного сорта, находящиеся в одинаковых условиях). *Противоположность* физических объектов следует рассматривать как предельный случай различия, когда это различие является *полным*, т. е. объекты не имеют никаких одинаковых свойств.

Заметим, что частицы и античастицы в этом смысле не являются противоположностями, так как они имеют кроме различных еще и одинаковые свойства (так, например, электрон и позитрон имеют различные заряды, но одинаковые по величине спины и массы покоя). Таким образом, частицы и античастицы являются *полярными*, но не противоположными объектами.

В связи с изложенным возникают следующие вопросы:

1) существуют ли в природе «противоположные объекты»;

2) возможно ли взаимодействие между ними, каковы особенности этого взаимодействия и значение в природе;

3) чем отличаются взаимодействия между одинаковыми, полярными и противоположными объектами.

Обсуждение этих вопросов имеет важное мировоззренческое значение; положительные результаты этого обсуждения позволят уточнить наши представления о том, как устроена окружающая нас природа. Такое обсуждение должно проводиться на основе определенной философской системы и затронет все разделы физики. В частности, можно полагать, что противоположными объектами в природе являются «вещество» и «поля». Под «веществом» обычно понимаются элементарные частицы и системы, составленные из них: атомные ядра, атомы, молекулы и т. д.; под «полем» понимаются различные силовые поля: гравитационные, электромагнитные, ядерные и т. д. Существуют два представления о полях. В одном из них предполагается, что поля *непрерывно* заполняют пространство вокруг частиц вещества и, будучи «особым образом» связаны с ними, определяют характер и интенсивность взаимодействия между ними. В другом представлении предполагается, что каждое поле состоит из «особых частиц поля», которые испускаются и поглощаются частицами вещества и тем самым вызывают силы взаимодействия между ними. Например, электромагнитное поле считается состоящим из фотонов («фотонный газ»); если их число в единице объема очень велико, то электромагнитное поле будет вести себя как непрерывная среда; если же это число мало и изучаются процессы, в которых участвуют отдельные фотоны, то понятие электромагнитного поля как непрерывной среды теряет смысл.

Здесь необходимо подчеркнуть, что существующие в настоящее время представления о веществе и полях не следует полагать окончательными. Развитие экспериментальной и теоретической физики может привести не только к уточнению, но и к радикальным изменениям наших представлений о природе и о сущности происходящих в ней явлений. Возможно, что в будущем восторжествуют *монистические* мировоззрения, согласно которым природа состоит: 1) либо только из частиц вещества, а поле есть лишь способ описания взаимодействия между ними; 2) либо только из различных полей, а частицы вещества есть лишь их «особые точки». Однако не исключено, что все известные опытные данные получат удовлетворительное объяснение и на основе *дуалистического* мировоззрения, в котором вещество и поля полагаются противоположными объектами, несводимыми и неотделимыми друг от друга, неразрывное взаимодействие которых является основой всех наблюдаемых нами явлений природы.

Дуализм обнаруживается и в одновременном существовании вероятностного и однозначного описания физических явлений. Классическое, строго детерминированное описание невозможно исключить из физики; оно необходимо для описания *наивероятного* течения физических явлений. С другой стороны, *всегда* существует разброс состояний изучаемых объектов (и физических величин, описывающих эти состояния), и этот разброс носит вероятностный характер. В настоящее время объективное существование вероятностных процессов в природе считается обоснованным теоретически и экспериментально; в квантовой физике (см. ч. IV, § 10, 11) вообще отрицается однозначность в поведении элементарных частиц и микросистем. Это означает не полное отрицание однозначности (детерминированности) в природе, а лишь ограничение области действия. Однозначность и вероятность являются дуалистическими понятиями; они *неотделимы* (вероятностный разброс существует вокруг наивероятных значений, входящих в однозначные законы), *несводимы* (невозможно ограничиться только одним способом описания физических явлений), а их взаимную связь можно заметить почти во всех разделах физики.

Дуализм у элементарных частиц имеет существенно важное значение в формировании свойств физических систем, образованных из этих частиц. Рассматривая известные микрофизические системы, можно заметить, что они образованы в конечном счете из различных частиц. Одинаковые частицы либо не взаимодействуют, либо же отталкиваются друг от друга и физической системы с качественно новыми свойствами не образуют. Так, например, протоны, нейтроны и электроны в отдельности не образуют физических систем, но, соединяясь вместе, образуют ядра и атомы различных веществ. Можно утверждать, что в совокупности одинаковых элементарных частиц всегда происходит простое (аддитивное) сложение их свойств. Только при взаимодействии частиц, обладающих противоположными свойствами, происходит особый (качественный) синтез этих свойств, благодаря чему физические системы приобретают новые свойства. Таким образом, можно утверждать, что *появление качественно новых свойств возможно только при взаимодействии существенно различных частиц*.

Объективный дуализм природы находит свое отражение и в важнейших физических понятиях. Типичным примером являются понятия *дискретности* и *непрерывности*. Они *несводимы* друг к другу; в противном случае можно было бы ограничиться использованием только одного из этих понятий. В истории физики известны попытки исключить дискретность или непрерывность из описания явлений, но они успеха не имели. Они *неотделимы* друг от друга и неразрывно *взаимосвязаны* во всех физических явлениях, так как в них обязательно участвуют частицы и поля, вносящие своими фундаментальными свойствами элементы дискретности и непрерывности.

В заключение заметим, что и сама физика как наука развивается на основе взаимодействия двух противоположных частей — теоретической и экспериментальной, которые неотделимы и взаимосвязаны, несводимы друг к другу и взаимодействуют, определяя направление и ход развития физических наук.

Аберрация линз 449
Абсолютно черное тело 519
Абсолютный нуль температуры 225
Автоколебания 87
Автоэлектронная эмиссия 293
Адиабатический процесс 145
Акустическая локация 118
Акустический спектр 114
Акустическое сопротивление 112
Акцепторы 317
Аморфные вещества 219
Ампер 329
Анизотропия кристаллов 229
Аннигиляция 608
Антистоксова область 545
Антиферромагнетики 375
Античастицы 607
Апериодический разряд 384
Апертурная диафрагма 451
Апостильб 407
Астигматизм 405
Атомная единица массы (а. е. м.) 598
Ахроматы 450

Барионы 609
Барометрическая формула 121
Бел 115
Биения 84

Вариконды 269
Варистор 320
Вебер 336
Вектор Пойнтинга 393
Возгонка 225
Волновая поверхность 91, 407
Вольт 253
Вязкость жидкостей 123, 203

Газовая постоянная 164
Гаусс 333
Гейри 355
Гистерезис 375
Градиент плотности 185
— скорости 123
— температуры 185
Градус Кельвина 225

Давление гидростатическое 120
— капиллярное (лапласовское) 208
— молекулярное 207
— света 402
Декремент затухания 78
Децибел 115
Диамагнитный эффект 370
Диод ламповый 303
— полупроводниковый 320
Диоптрия 445
Диполь электрический 242
Дисперсия вращательная 441
— света 412
— среды 103
Дисторсия 451
Дихроизм 438
Дозиметрия 536
Домены 262, 373
Доноры 317

Закон Ампера 329
— Архимеда 120
— Био—Савара—Лапласа 323
— Бугера 411
— Вавилова 546
— Вебера—Фехнера 114
— Видемана—Франца 281
— Вина (смещения) 525
— Вольта 297
— Гука 12
— Джоуля—Ленца 280, 283
— Дюлонга и Пти 228
— Ламберта 406
— Ленца 351
— Кеплера 29
— Кирхгофа 519
— Кулона 237
— Кюри 372
— Кюри—Вейса 374
— Малюса 438
— Мозли 531
— Ома (дифференциальный) 280
— Паскаля 119
— полного тока 325, 365
— Релея 414
— сохранения зарядов 234
— — энергии 56
— Стефана—Больцмана 524
— Фарадея (для электролиза) 313
— — (для электромагнитной индукции) 351
Запирающий слой 318
Звуковое давление 111
Зоны Френеля 425

Изобарический процесс 145
Изотермическая поверхность 185
Изотермический процесс 145
Изотопы 564
Изотропная среда 90
Изохорический процесс 145
Изоэнтропический процесс 146
Индуктивность 354
Индукция магнитного поля 330
Инерциальные системы отсчета 36
Интерференционные полосы 419
Интерферометры 417, 421

Кандела 406
Квазистационарные токи 355
Квантовые числа 485, 509
— генераторы 562
Когерентные волны 101, 416
Кома 450
Комбинационное рассеяние 515
Комптоновское рассеяние 470
Контактная разность потенциалов 294
Координационное число 226
Коэрцитивная сила 375
Коэффициент взаимной индукции 357
— внутреннего трения 186
— восстановления 64
— вязкости 123
— давления (температурный) 190
— дисперсии (света) 412
— диссоциации 311
— диффузии 185

— диффузного отражения 407
— затухания (колебаний) 77
— ионизации 312
— отражения звука, света 112, 410
— поверхностного натяжения 205
— поглощения (света) 411
— расширения 190, 230
— самоиндукции 354
Коэффициенты сжимаемости 119
— теплопроводности 185, 230
— упругости (пружины) 72
— усиления лампы 304
— Холла 340
Краевой угол 206
Критическая опалесценция 202
Критическое состояние вещества 192, 197
Кюри 566

Лазер 554
Ламинарное течение 121
Лептоны 609
Люкс 407
Люминесцентные лампы 549
Люминесцентный анализ 549
Люминесценции выход 546
Люминофоры 537

Магнетон (Бора, ядерный) 573
Магнитная восприимчивость 367
— индукция 330
— постоянная 330
— проницаемость 330
Магнитное сопротивление 380
Магнитный гистерезис 375
— момент контура 337
— поток 336
Магнитострикция 374
Мазер 562
Максвелл 336
Мезоны 609
Метастабильное состояние 137
Механический эквивалент света 404
Модуль объемной упругости 119
— упругости Юнга 12
Молизация ионов 312
Монокристалл 222

Намагниченность вещества 368
Невесомость 28
Нейтрино 593, 610
Нейтрон 563
Нит 406
Нуклид 564
Ньютон 23

Обертоны стоячей волны 110
Оптическая ось кристалла 432
— пирометрия 527
— разность хода лучей 519
— сила линзы 445
Оптически активные вещества 441
Оптическое излучение 393
Опыт Резерфорда 457
— Франка и Герца 459
— Штерна 166
— Эйнштейна и де-Гааза 372
Оси инерции 47

Параксиальные лучи 443
Параметрическое возбуждение 82, 387
Параметры состояния 134
Перпетуум мобиле второго рода 157
— — первого рода 156
Плазма 310
Пластинка в четверть волны 440
Пластичность 231
Плоскость поляризации 398
Плотность потока излучения 399
Поликристаллическая структура 222
Политропа 176
Порог болевого ощущения 114
— слышимости 114
Постоянная Больцмана 131
— Вина 525
— Планка 396
— распада 566
— Ридберга 461
— Стефана—Больцмана 524
Постулаты Бора 462
Потенциал поля 25, 253
Потенциальная энергия 55
— яма 61
Потенциальный барьер 60
Правила Кирхгофа 289
— Стокса 545
Преобразования Галилея 36
— Лоренца 601
Призма Николя 436
Принцип Гюйгенса 100, 408
— относительности 623
— — Галилея 34
— постоянства скорости света 623
— реактивного движения 31
— суперпозиции 100, 236, 323
Проводимость дырочная, электронная 316
Просветление оптических деталей 420
Пьезоэлектрический эффект 267

Равновесные состояния и процессы 136
Радиоактивные семейства 587, 588
Радиус кривизны поверхности 208
— — траектории 9
Разрешимость двух волн 429
Разряды в газах 309
Реверберация звука 116
Резонанс 81, 376, 386
Рекомбинация ионов 305, 312
Релаксация 136
Рентген 536
Ридберга температура 225

Светимость 402, 405
Световой поток 405
Световое ощущение 404
Сверхпроводимость 284
Свечение Вавилова—Черенкова 394
Свободные оси 47
Сегнетоэлектрики 268
Сила звука 111
— инерции 35
— Лоренца 328
— света 405
Смачивание 206
Спектр звуковых частот 113
Спектральная чувствительность 404
Спектральные линии 397
Спин 338
Спонтанная поляризация 262
Спонтанный переход 550
Стеклообразные вещества 219
Степени свободы молекул 130
Стигматические изображения 449
Стильб 406
Сторонние силы 285
Сублимация 225

Текучесть жидкостей 203
Тембр звука 114
Температура затвердевания 219
— Кюри 372
— размягчения 221
Теорема Остроградского—Гаусса 247
— о циркуляции 325
Тепловой насос 158
Термистор 320
Термоэлектродвижущая сила 298
Термоэлектронная эмиссия 300
Термоэлемент 298
Тесла 332
Ток конвекционный 277

— насыщения 302, 308
— поляризации 277
— проводимости 276
Токи молекулярные 327
— Фуко (вихревые) 361
Транзистор 319
Тройная точка вещества 224
Туннельный эффект 483
Турбулентное течение 121

Удельная ионизация 588
— теплота возгонки 255
— — испарения 217
Упругость 230
Уравнение Бернулли 121
— Ван-дер-Ваальса 195
— Клапейрона—Менделеева 164
— Майера 172
— Мещерского 33
Уравнения Пуассона 174
— Циолковского 32
— Шредингера 479
Уровни энергии 486
Условие Брюстера 436
— разрешимости двух волн 429

Фаза колебаний 68
Фазовая поверхность 91
— скорость 93
Фазовые переходы 135
Фарада 272
Ферриты 375
Ферромагнетики 372
Фигуры Лиссажу 85
Фон 115
Формула Вульфа-Брэгга 535
— де-Бройля 473
— Клапейрона—Клаузиуса 218, 221
— Коши 413
— Лапласа 208
— Эйнштейна (для фотоэффекта) 468
Фотон 396, 463
Фотосопротивления 320
Фотоэлектрический эффект 293, 467
Фотоэлектронная эмиссия 293
Фотоэлементы 440
Фронт волны 90, 407
Функция видности 404
— распределения молекул 165
— — энергии по спектру 399

Характеристический спектр 531

Центр инерции 66
— масс системы 66
Цикл Карно 178
Циклы (термодинамические) 149

Эквивалентная проводимость 313
Эквипотенциальные поверхности 25
Экстратоки 355
Электреты 269
Электродвижущая сила 286
— — индукции 345
— — самоиндукции 353
Электронвольт 253
Электронный газ 281
— захват 593
Электролитическая диссоциация 311
Электростатическая защита 259
Электростатическое поле 352
Электрострикция 266
Эмиссия электронов 292
Энергия заряженного проводника 275
— колебаний 74, 384
— потенциальная 55
— связи атомных ядер 568
— электрического поля 275
Энтропия 146
Эрстед 333
Эффект Джоуля—Томсона 200
— Доплера 112
— Керра 440
— Комптона 470
— Пельтье 299
— Холла 340

Яркость 405
Ядра зеркальные, магические 565

ОГЛАВЛЕНИЕ

Предисловие 3

Часть I

МЕХАНИКА

Глава 1. **Основные законы механики** 4

§ 1. Системы отсчета; линейные и угловые скорости и ускорения 4
§ 2. Законы механики Ньютона. Системы единиц 11
§ 3. Движение тела в поле тяготения Земли 24
§ 4. Движение тел с переменной массой 31
§ 5. Принцип относительности в механике 34

Глава 2. **Вращение твердого тела** 39

§ 6. Кинетическая энергия вращающегося тела; момент инерции тел 39
§ 7. Законы механики для вращательного движения 42
§ 8. Свободные оси. Балансировка роторов. Гироскопы 46

Глава 3. **Законы сохранения в системе взаимодействующих тел** 50

§ 9. Закон сохранения импульса (количества движения) 50
§ 10. Закон сохранения момента импульса (момента количества движения) 53
§ 11. Закон сохранения механической энергии 54
§ 12. Потенциальная энергия некоторых систем 57
§ 13. Столкновение шаров 61
§ 14. О центре масс системы 64

Глава 4. **Колебательное движение тел** 67

§ 15. Периодическое движение; гармонические колебания 67
§ 16. Скорость и ускорение колеблющегося тела. Дифференциальное уравнение гармонических колебаний 69
§ 17. Сила и энергия при гармонических колебаниях. Простейшие механические колебательные системы 71
§ 18. Собственные, свободные и вынужденные колебания. Параметрическое возбуждение колебаний 75
§ 19. Понятие о нелинейных колебаниях 82
§ 20. Сложение и разложение колебаний 83
§ 21. Автоколебания 87

Глава 5. **Волновые процессы; основы акустики** 90

§ 22. Образование и распространение волн в упругой среде 90
§ 23. Формула гармонической волны 92
§ 24. Поток энергии в волновых процессах 98
§ 25. Интерференция и дифракция волн. Распространение волн в средах с дисперсией 100
§ 26. Стоячие волны 104
§ 27. Звуковые колебания и волны 111
§ 28. Ультразвуки и их применения 116

Глава 6. **Механика жидкостей и газов** 119

§ 29. Гидростатическое давление. Барометрическая формула 119
§ 30. Ламинарное и турбулентное течения. Формула Бернулли 121

Часть II

МОЛЕКУЛЯРНАЯ ФИЗИКА И ТЕРМОДИНАМИКА

Глава 1. **Основы молекулярно-кинетической теории и термодинамики** . . . 128

§ 1. Тепловое движение 128
§ 2. Взаимодействие молекул; энергия связи молекул 132
§ 3. Агрегатные состояния и фазовые переходы. Равновесные и неравновесные состояния и процессы 134
§ 4. Внутренняя энергия термодинамической системы. Первый закон термодинамики 139
§ 5. Внешняя работа системы и теплообмен с окружающей средой. Теплоемкость тел 142
§ 6. Изопроцессы; адиабатический процесс. Энтропия системы. Замкнутые (круговые) процессы 145
§ 7. Второй закон термодинамики. Закон возрастания энтропии 151

Глава 2. **Идеальный газ** 160

§ 8. Основное уравнение кинетической теории газов. Средние скорости молекул 160
§ 9. Уравнение состояния идеального газа 163
§ 10. Изопроцессы в идеальном газе; теплоемкости газов 167
§ 11. Круговые процессы, совершаемые идеальным газом; цикл Карно. Энтропия идеального газа 176
§ 12. Диффузия, теплопроводность и внутреннее трение в газах; число столкновений и длина свободного пробега молекул 184

Глава 3. **Реальные газы, жидкости и твердые тела** 188

§ 13. Отступления от законов идеальных газов. Насыщенные пары. Критическое состояние 188
§ 14. Уравнение Ван-дер-Ваальса. Внутренняя энергия реального газа . . . 195
§ 15. Жидкости; их основные свойства. Молекулярное давление и поверхностное натяжение 201
§ 16. Капиллярные явления 207
§ 17. Испарение и кипение жидкостей; конденсация паров 211
§ 18. Кристаллические и аморфные тела. Кристаллизация, плавление и испарение твердых тел 218
§ 19. Типы кристаллических решеток. Тепловые и упругие свойства твердых тел 226

Часть III

ЭЛЕКТРИЧЕСТВО И МАГНЕТИЗМ

Глава 1. **Электростатика** 234

§ 1. Электрические заряды. Проводники и диэлектрики 234
§ 2. Электрическое поле; индукция и напряженность поля. Закон Кулона 235
§ 3. Электрическое поле системы зарядов; теорема Остроградского — Гаусса 242
§ 4. Работа перемещения заряда в электрическом поле; разность потенциалов. Энергия взаимодействия зарядов 252
§ 5. Проводники в электрическом поле 258
§ 6. Диэлектрики в электрическом поле 261
§ 7. Электроемкость. Энергия заряженного проводника. Плотность энергии поля 270

Глава 2. **Электрический ток** 276

§ 8. Электрический ток проводимости; плотность тока 276
§ 9. Электронная теория проводимости металлов. Законы Ома и Джоуля — Ленца в дифференциальной форме. Связь между электропроводностью и теплопроводностью металлов 279
§ 10. Работа электрического тока. Сопротивление проводников. Сверхпроводимость 282
§ 11. Электродвижущая сила; закон Ома и правила Кирхгофа 284
§ 12. Работа выхода электронов из металла. Контактная разность потенциалов. Термоэлектрические явления 291
§ 13. Термоэлектронная эмиссия. Электрический ток в вакууме. Электронная лампа 300
§ 14. Электрический ток в газах. Понятие о плазме 304
§ 15. Электрический ток в жидкостях. Законы Фарадея 311
§ 16. Полупроводники и их применения 313

Глава 3. **Электромагнетизм.** 321

§ 17. Магнитное поле 321
§ 18. Напряженность магнитного поля вокруг движущегося заряда и проводников с токами 322
§ 19. Действие магнитного поля на заряды и проводники с током; сила Лоренца и закон Ампера. Индукция магнитного поля. 328
§ 20. Действие магнитного поля на контур с током. Движение заряда в магнитном поле 334
§ 21. Работа перемещения проводника с током в магнитном поле 341
§ 22. Электромагнитная индукция; законы Фарадея и Ленца. Вращающийся виток в магнитном поле 343
§ 23. Явления самоиндукции и взаимоиндукции. Энергия магнитного поля. Вихревые токи 351
§ 24. Ток смещения и его магнитное поле; теория Максвелла 361
§ 25. Магнитное поле в веществе. Диамагнетизм; парамагнитные и ферромагнитные вещества 366
§ 26. Магнитное поле в ферромагнитных телах различной формы; магнитодвижущая сила и магнитное сопротивление 377

Глава 4. **Электромагнитные колебания и волны** 381

§ 27. Колебательный контур. Незатухающие и затухающие колебания . . . 381
§ 28. Вынужденные колебания; электрический резонанс 385
§ 29. Электромагнитные волны 388

Часть IV

ОПТИКА И ФИЗИКА АТОМА

Глава 1. **Основы волновой оптики** 396

§ 1. Волновая (электромагнитная) и корпускулярная (фотонная) теории света 396
§ 2. Фотометрические понятия и величины. Давление света 402
§ 3. Распространение, отражение, преломление и поглощение света. Дисперсия. Рассеяние света 407
§ 4. Интерференция света; когерентность световых лучей. Интерферометры 415
§ 5. Дифракция света; дифракционный спектр 423
§ 6. Распространение света в анизотропной среде; получение и применение поляризованного света 432
§ 7. Линзы; их оптическая сила. Аберрации. Диафрагмы в оптических приборах 443
§ 8. Понятие о голографии 453

Глава 2. **Физика атомов и молекул. Основы квантовой теории** 456

§ 9. Строение атома. Опыты Резерфорда, Франка и Герца. Элементарная теория атома водорода 456
§ 10. Корпускулярно-волновые свойства частиц. Фотоэлектрический эффект. Формула де Бройля. Дифракция электронов 466
§ 11. Основные положения квантовой теории. Уравнение Шредингера. Соотношение неопределенностей Гейзенберга 476
§ 12. Частицы и физические системы. Статистические распределения частиц по состояниям. Понятие температуры 490
§ 13. Взаимодействие атомов и молекул. Энергия связи атомов 501

Глава 3. **Излучение и поглощение энергии атомами и молекулами** 509

§ 14. Атомные и молекулярные спектры излучения; способы возбуждения 509
§ 15. Тепловое излучение 517
§ 16. Рентгеновские лучи. Характеристический рентгеновский спектр элементов 527
§ 17. Люминесцентное излучение 537
§ 18. Индуцированное излучение. Лазеры 550

Глава 4. **Атомные ядра и элементарные частицы** 563

§ 19. Основные свойства атомных ядер 563
§ 20. Модельные представления о структуре ядер 578
§ 21. Радиоактивные ядра и их излучения 586
§ 22. Ядерные реакции деления и синтеза. Термоядерные реакции 597
§ 23. Элементарные частицы 607

Заключение 616

§ 1. Физические измерения. Связь между способами измерения физических величин и формулировками законов физики 616
§ 2. О теории относительности Эйнштейна 620
§ 3. Законы сохранения, превращения и перераспределения энергии в физических системах 627
§ 4. О вариационных принципах физики 638
§ 5. Статистические законы и флуктуации 642
§ 6. О дуализме в физике 646

Предметный указатель 650

Рубен Георгиевич Геворкян

КУРС ФИЗИКИ

Редактор С. А Крылов. Переплет художника А В Пушкарного Художественный редактор В. И. Пономаренко. Технический редактор Р. С. Родичева. Корректор Г. И. Кострикова.

ИБ № 1641

Изд. № ФМ—595 Сдано в набор 24.10.78. Подп. в печать 28.02.79. Т-03261. Формат 60×90¹/₁₆. Бум. тип. № 3. Гарнитура литературная. Печать высокая. Объем 41 усл. печ. л. 43,39 уч.-изд. л. Тираж 60 000 экз. Зак. № 191. Цена 1 р. 40 к.

Издательство «Высшая школа», Москва, К-51, Неглинная ул., д. 29/14

Ордена Октябрьской Революции, ордена Трудового Красного Знамени Ленинградское производственно-техническое объединение «Печатный Двор» имени А М. Горького «Союзполиграфпрома» при Государственном комитете СССР по делам издательств, полиграфии и книжной торговли. 197136, Ленинград, П-136, Гатчинская, 26.