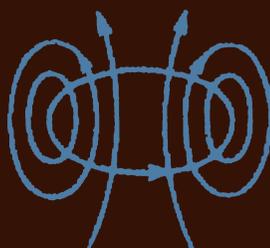


А. И. Ахиезер, И. А. Ахиезер

ЭЛЕКТРОМАГНЕТИЗМ и электромагнитные волны



А. И. Ахиезер, И. А. Ахиезер

ЭЛЕКТРОМАГНЕТИЗМ и электромагнитные волны

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



МОСКВА
«ВЫСШАЯ ШКОЛА»
1985

ББК 22.33
А95
УДК 538.3

Рецензенты:
кафедра общей физики Московского физико-технического ин-
ститута; проф. И. П. Ипатова (Ленинградский политехниче-
ский институт)

Ахиєзер А. И., Ахиєзер И. А.
А95 **Электромagnetизм и электромagnetитные волны:**
Учеб. пособие для вузов. — М.: Высш. шк., 1985. —
504 с., ил.
В пер.: 1 р. 40 к.

В пособии рассмотрены входящие в курс общей физики вопросы физики электромagnetитных явлений, включая оптику и основы теории относительности. От аналогичных книг пособие отличается широтой охвата и методом изложения — один из авторов, академик АН УССР А. И. Ахиєзер, известен своими фундаментальными теоретическими работами в области электродинамики. В первом и втором разделах книги выводятся уравнения для электромagnetитных полей в вакууме и веществе, изучаются электромagnetитные свойства вещества. Третий и четвертый разделы посвящены электромagnetитным волнам в вакууме и среде, основам нелинейной и геометрической оптики и др.

А $\frac{1704040000-149}{001(01)-85}$ 74—84

ББК 22.33
537

ПРЕДИСЛОВИЕ

Предлагаемая книга существенно отличается от других книг по электромагнетизму как по широте охвата, так и по характеру изложения. Нет сколько-нибудь важного, вошедшего в основной фонд современной физики вопроса, относящегося к электромагнетизму, который не освещался бы в этой книге. При этом, с одной стороны, на современном уровне отражены общие и принципиальные вопросы и, с другой стороны, уделено внимание специфике электрических, магнитных и оптических явлений в конкретных типах веществ и принципиальной стороне важнейших технических приложений.

Книга представляет собой единое целое, и все ее главы тесно переплетены друг с другом. Тем не менее она может быть разбита на четыре тесно связанные между собой части:

I. Основы электромагнетизма и физики пространства — времени.

II. Электрические и магнитные свойства вещества.

III. Электромагнитные волны и оптика.

IV. Взаимодействие электромагнитных волн с частицами и волнами.

Книга представляет собой часть курса общей физики и, по мнению авторов, не должна подменять ни курса теоретической физики, ни курса экспериментальной физики. Это значит, что, охватывая идейный фонд физики, мы старались избегать использования тонкостей теоретического аппарата и описания подробностей физических экспериментов. Поэтому в предлагаемой книге изложение ведется доступно и без излишних математических и экспериментальных деталей, всегда выясняется принципиальная, а не методическая сторона вопроса.

Хорошо известно, что понятия элементарного и неэлементарного в математике претерпели за последние годы большие изменения; в частности, выпускники средней школы обладают теперь математическими знаниями, которые раньше имели только студенты университетов к концу первого семестра обучения. Поэтому принятые во многих курсах общей физики примитивные приемы вывода формул авторы считают явно устаревшими. Это же относится и к физическим аспектам изложения. Теперь, когда уже школьники знакомы с основами квантовой физики и теории относительности, нам кажется неразумным ис-

кусственно строить курс электромагнетизма, избегая квантовых понятий. Разумеется, эти понятия должны поясняться и использоваться только в простейшем, общефизическом плане. Использование в книге ряда понятий квантовой физики позволило нам изложить элементы теории электропроводности металлов и полупроводников, разъяснить природу излучения Вавилова—Черенкова, эффекта Комптона, рождения электрон-позитронных пар и т. д. В отдельной главе излагаются основы специальной и общей теорий относительности. Заключение книги посвящено вопросам развития учения об электромагнетизме.

Книга предназначена для широкого контингента читателей: студентов и аспирантов физических факультетов университетов, педагогических и инженерно-физических вузов; преподавателей физики вузов и старших классов средних школ; научных работников и инженеров. Для чтения книги достаточно знания математики в объеме программы первых двух лет обучения на физических факультетах университетов, педагогических и инженерных вузов.

Глава «Физика пространства—времени» написана совместно с Ю. П. Степановским.

Мы хотим поблагодарить Ф. Г. Басса, Ю. Г. Гуревича и А. Н. Рязанова за полезные замечания.

Авторы

ВВЕДЕНИЕ

МЕСТО ЭЛЕКТРО- МАГНЕТИЗМА В СОВРЕМЕННОЙ ФИЗИЧЕСКОЙ КАРТИНЕ МИРА

В известной нам части Вселенной вся материя состоит в основном из электронов, протонов, нейтронов, нейтрино и фотонов. Их называют *элементарными частицами*. Всего элементарных частиц нам известно несколько сотен. Они отличаются чрезвычайным разнообразием свойств: достаточно сказать, что такие частицы, как фотон, лишены массы (имеется в виду масса покоя), в то время как самые тяжелые из известных частиц имеют массу ~ 10 ГэВ; такие частицы, как электрон или протон, могут существовать неограниченно долго, в то время как очень короткоживущие частицы, так называемые резонансы, распадаются самопроизвольно за времена порядка $\sim 10^{-23}$ с. Одно объединяет все частицы: все они подвержены действию некоторых сил, или, как говорят, *взаимодействий*. Хотя проявлений этих сил бесконечно много, но различных типов взаимодействий всего четыре: *гравитационное, сильное, электромагнитное и слабое*. Сейчас выясняется, что электромагнитное и слабое взаимодействия являются двумя проявлениями одного и того же взаимодействия, подобно тому, как электричество и магнетизм, которые первоначально представлялись двумя разными взаимодействиями, оказались проявлениями единых электромагнитных сил. В достигнутой же пока в эксперименте области энергий взаимодействующих частиц электромагнитная и слабая ветви единого электро-слабого взаимодействия так сильно отличаются по своим свойствам и проявлениям, что их можно считать независимыми. Можно думать, что и другие взаимодействия в конечном счете окажутся не независимыми, а сольются в единое фундаментальное взаимодействие с четырьмя различными проявлениями.

Не каждому типу взаимодействий подвержены все элементарные частицы. Это свойство присуще только взаимодействию гравитационному. В основе его лежит закон всемирного тяготения Ньютона, согласно которому *все тела притягиваются друг к другу с силой, пропорциональной произведению их масс и обратно пропорциональной квадрату расстояния между ними*.

Гравитационные силы — самые слабые из известных нам сил. Достаточно заметить, что сила гравитационного притяжения между двумя электронами в 10^{45} раз меньше силы их электрического отталки-

вания на том же расстоянии. Сильное взаимодействие характеризуется энергиями, в миллионы раз большими энергий, связанных с электромагнитными взаимодействиями.

В структуре нашего мира главным проявлением сильного взаимодействия является существование *атомных ядер*. Различных ядер известно несколько сотен, но все они состоят только из двух элементарных частиц — нейтронов и протонов, «склеенных» между собой сильным взаимодействием. Электрические силы тоже играют свою, хотя и меньшую, роль в строении ядер. Дело в том, что протоны несут электрический заряд и потому отталкиваются друг от друга; в результате ядра с очень большим числом протонов (или, как говорят, атомным номером) не могут быть устойчивы. В этом причина того, что в природе существует всего 92 химических элемента — ядра же тяжелее урана получают только искусственно.

Интенсивность сильного взаимодействия примерно в 100 раз превосходит интенсивность электромагнитного (не говоря уже о гравитационном и слабом взаимодействиях). И тем не менее не сильное взаимодействие, образно говоря, «правит самодержавно» миром. Это связано с тем, что сильному взаимодействию подвержены не все частицы (хотя и большинство известных частиц, объединяемое под названием *адронов*); не подвержены сильному взаимодействию электрон и фотон. Но, кроме того, для сильного взаимодействия характерен очень малый *радиус действия* — порядка 10^{-15} м, т. е. порядка размеров ядер. Поэтому на расстояниях, больших ядерных, сильное взаимодействие уже не играет роли.

На больших расстояниях возрастает роль электромагнитных сил, очень медленно ослабевающих с расстоянием. Именно они объединяют ядра и электроны в атомы, и они же объединяют атомы в молекулы. Их проявлением являются силы трения, силы сцепления между частицами в твердых телах, жидкостях и неидеальных газах, силы, управляющие плазмой, — одним словом, все (кроме силы тяжести) силы, с которыми приходится сталкиваться человеку вне стен ядерных лабораторий.

Из молекул строятся химические соединения, от таких простых, как вода или молекулярный кислород, до таких сложных, как лежащие в основе жизни белки и нуклеиновые кислоты. Мир химии, биологии и биохимии — это мир электромагнитных сил. Электромагнитную природу имеют свет, радиоволны, рентгеновские и ультрафиолетовые лучи. Можно смело сказать, что в тех масштабах, в которых проходит практически вся деятельность человечества, мир управляется электромагнитными силами. В этом особая роль и особая важность электромагнетизма.

Электромагнитное взаимодействие обусловлено существованием *электрических зарядов*, которые бывают двух знаков: заряды одного знака отталкиваются, а заряды разных знаков притягиваются друг к другу. Поэтому большие тела редко бывают электрически заряженными: действующие между зарядами силы приводят к сближению противоположных по знаку зарядов, так что тела в целом делаются электрически нейтральными.

Гравитационное взаимодействие не знает отрицательных масс, и, следовательно, невозможно возникновение «гравитационно нейтральных», негравитирующих систем. Поэтому, несмотря на то что гравитационные силы значительно слабее электрических, они играют доминирующую роль в случае систем с большой массой, в частности в масштабах Вселенной.

В космических масштабах мир управляется силами гравитационными. Правда, в случае больших масс или релятивистских скоростей (близких к скорости света) законы гравитации значительно сложнее закона Ньютона. Установление их относится к компетенции общей теории относительности Эйнштейна, которая хотя и не относится непосредственно к электромагнетизму, но является естественным развитием так называемой частной (или специальной) теории относительности, без которой невозможно понимание электромагнитных явлений и которая в значительной мере возникла из их анализа. Поэтому мы не оставляем общую теорию относительности полностью в стороне, а посвящаем ей часть главы «Физика пространства—времени».

Несколько особняком в строении мира стоит слабое взаимодействие. Дело в том, что мы не знаем объектов, построенных этим взаимодействием, — оно, образно говоря, «разрушает» (и рассеивает), но не «созидает». Слабое взаимодействие приводит к β -распаду ядер, к конечности времени жизни нейтрона (10^3 с) и к распадам относительно долгоживущих (10^{-6} — 10^{-10} с) элементарных частиц (некоторые более короткоживущие, $\sim 10^{-21}$ с, частицы распадаются в результате электромагнитного взаимодействия).

Подчеркнем еще раз, что существует широкая область явлений, где четыре фундаментальных взаимодействия не интерферируют между собой и могут рассматриваться как независимые. Исходя из этого, мы будем изучать в этой книге только одно, а именно электромагнитное, взаимодействие, причем только в той области, где оно не пересекается ни с сильным, ни со слабым взаимодействием.

Первые главы книги посвящены простейшим проявлениям электрических и магнитных свойств вещества. Далее излагаются электрические и магнитные свойства конкретных сред, таких, как диэлектрики, проводники, ферромагнетики, сверхпроводники. Мы стараемся по возможности оставаться в рамках классической физики, хотя, когда речь идет о строении конденсированных тел, этого оказывается недостаточно и нам приходится использовать элементы квантовой физики.

Изучив факты, для которых малосущественно единство электричества и магнетизма, мы переходим к единой максвелловской теории электромагнитных явлений, в которой излагаем общие законы электромагнетизма, а затем посвящаем отдельную главу физике пространства—времени, в которой излагаются специальная теория относительности и основные идеи общей теории относительности.

Далее мы рассматриваем электромагнитные волны различных типов в различных средах и физических системах, их распространение и способы их возбуждения. Книга завершается изложением основных фактов и закономерностей, касающихся взаимодействия заряженных частиц и фотонов с веществом.



I. ОСНОВЫ ЭЛЕКТРО- МАГНЕТИЗМА И ФИЗИКИ ПРОСТРАНСТВА — ВРЕМЕНИ

Глава 1. ЭЛЕКТРОСТАТИЧЕСКОЕ ПОЛЕ В ВАКУУМЕ

1.1. Закон Кулона

Электрическому взаимодействию подвержены не все тела, а лишь те, которые обладают электрическим зарядом; их называют заряженными. Существенно, что электрическое взаимодействие может приводить как к притяжению, так и к отталкиванию. Это значит, что электрические заряды бывают двух родов и их можно различать по знаку: тела, заряженные электричеством одинакового знака, отталкиваются, а тела, заряженные электричеством разных знаков, притягиваются.

Простейший вид сила электрического взаимодействия имеет в том случае, когда заряженные тела неподвижны и находятся в вакууме. Если при этом размеры тел малы по сравнению с расстояниями между ними (в таких случаях говорят о *материальных точках*), то силу взаимодействия определяют только заряды тел и расстояния между ними. Она направлена вдоль линии, соединяющей тела, и обратно пропорциональна квадрату расстояния.

Математически эту силу можно записать в виде $F = k_1 q_1 q_2 / r_{12}^2$, где q_1 и q_2 — заряды тел, r_{12} — расстояние между телами и k_1 — константа, зависящая от выбора единицы заряда. Этой формулой выражается закон Кулона — основной закон электрического взаимодействия. Он был открыт Кэвэндишем в 1771—1779 гг. и перерисован Кулоном в 1786—1789 гг.

Проще всего выбрать единицу заряда таким образом, чтобы константа k_1 была равна единице. Такой выбор делается в так называемой гауссовой системе единиц СГС, используемой главным образом в фундаментальных и теоретических исследованиях. Закон Кулона имеет в этой системе вид

$$F = q_1 q_2 / r_{12}^2.$$

Из этой формулы следует, что электрический заряд в системе СГС имеет размерность $[q] = \Delta^{3/2} T^{-1} M^{1/2}$, т. е. единицей заряда является $1 \text{ см}^{3/2} \text{ с}^{-1} \text{ г}^{1/2}$.

Эта единица (не имеющая специального названия) очень мала, и поэтому на практике пользуются большей единицей — *кулоном* (Кл), который численно в $3 \cdot 10^9$ раз больше гауссовой единицы заряда.

Кулон является единицей заряда в системе единиц, называемой СИ и широко используемой в настоящее время, особенно в технике. В отличие от гауссовой системы с тремя основными единицами — сантиметром (см), граммом (г) и секундой (с) для длины массы и времени, СИ является для электромагнетизма четырехразмерной с четырьмя основными единицами — метром (м), килограммом (кг), секундой (с) для длины, массы и времени и еще ампером (А) для силы тока. Кулон представляет собой производную единицу, равную ампер-секунде: $\text{Кл} = \text{А} \cdot \text{с}$. Далее (см. § 6.6) мы рассмотрим подробнее систему СГС, СИ, а также и другие возможные системы и дадим точные определения для всех встречающихся единиц, пока же запишем закон Кулона в СИ:

$$F = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r_{12}^2},$$

где $(4\pi\epsilon_0)^{-1}$ — константа, численно равная, по определению, 10^{-7} от квадрата скорости света c в вакууме. Так как $c = 3 \cdot 10^8$ м/с, то

$$(4\pi\epsilon_0)^{-1} = 9 \cdot 10^9 \text{ Н} \cdot \text{м}^2 \cdot \text{Кл}^{-2};$$

$$\epsilon_0 = 8,854 \cdot 10^{-12} \text{ Кл}^2 \text{ Н}^{-1} \text{ м}^{-2} = 8,854 \cdot 10^{-12} \text{ Ф} \cdot \text{м}^{-1}$$

(Ф обозначает единицу емкости — фарад).

Закон Кулона можно записать также в векторной форме. Если \mathbf{n}_{21} — единичный вектор, направленный от q_2 к q_1 , то сила \mathbf{F}_1 , действующая на заряд q_1 со стороны заряда q_2 , может быть представлена в виде

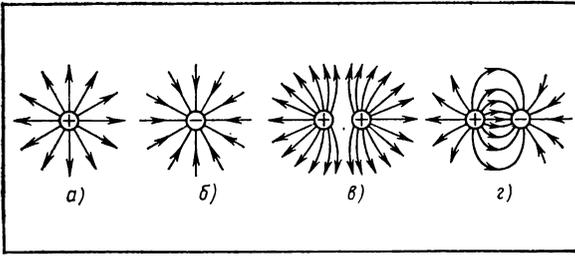
$$\mathbf{F}_1 = \mathbf{n}_{21} \frac{q_1 q_2}{4\pi\epsilon_0 r_{12}^2}.$$

Электрический заряд не может самопроизвольно появиться или исчезнуть. Иными словами,

суммарный электрический заряд замкнутой системы тел, равный сумме (алгебраической) зарядов тел, входящих в состав системы, не может измениться со временем.

Этот закон называют **законом сохранения электрического заряда**, он представляет собой один из фундаментальнейших законов природы. При электризации трением, например, трущиеся тела заряжаются всегда разноименными зарядами, равными по модулю.

Закон сохранения заряда действует и в мире элементарных частиц. Например, при столкновениях атомных ядер могут рождаться электроны, но вместе с каждым электроном рождается позитрон — элементарная частица, отличающаяся от электрона только знаком заряда. Так как все тела построены из атомов, состоящих из электронов и ядер, а ядра состоят из протонов и нейтронов, причем заряд протона только



1.1.

Силовые линии электрического поля в простейших случаях:

a — положительный, *б* — отрицательный точечные заряды, *в* — два положительных точечных заряда, *г* — положительный и отрицательный точечные заряды

знаком отличается от заряда электрона, а нейтрон не обладает зарядом, то из закона сохранения заряда следует, что заряд любого тела кратен заряду электрона или протона. Этот заряд имеет числовое значение

$$e = 1,60219 \cdot 10^{-19} \text{ Кл} = 4,80298 \cdot 10^{-10} \text{ СГС.}$$

Заряды всех известных заряженных элементарных частиц кратны или равны e , так что e выступает как элементарный заряд. Мы не знаем ни одной частицы, заряд которой был бы дробной частью e . Это одно из замечательных и не объясненных до сих пор явлений природы.

Так как между заряженными телами действуют электрические силы, то можно сказать, что

заряженное тело создает вокруг себя некоторое силовое поле. Это поле называют электрическим.

Чтобы характеризовать электрическое поле, создаваемое одним или несколькими заряженными телами, внесем в поле маленькое заряженное тело, несущее на себе малый электрический заряд $q_{\text{п}}$. Такое тело — его называют *пробным зарядом* — практически не искажает поля, в которое его внесли, и поэтому сила \mathbf{F} , действующая на пробный заряд, может характеризовать поле, создаваемое заряженным телом (или совокупностью заряженных тел):

$$\mathbf{F} = q_{\text{п}}\mathbf{E}.$$

Здесь вектор \mathbf{E} не зависит от $q_{\text{п}}$, а определяется только заряженными телами и местом, в которое внесен пробный заряд. Этот вектор, меняющийся, вообще говоря, от точки к точке, называют *напряженностью электрического поля*, создаваемого заряженными телами. Напряженность поля имеет размерность $[E] = [F]/[q] = \text{Н} \cdot \text{Кл}^{-1}$.

Перемещая пробный заряд из одного места в другое, можно установить структуру всего электрического поля. Наглядным и удобным оказывается графический метод описания поля — с помощью так называемых электрических силовых линий.

Силовая линия — это линия, касательная к которой в каждой точке имеет направление, совпадающее с направлением напряженности поля.

На рис. 1.1 изображены электрические силовые линии в четырех простейших случаях — одного точечного заряда (обоих знаков) и двух точечных зарядов, одноименных и разноименных.

Силовые линии всегда начинаются на положительных и кончаются на отрицательных зарядах или на бесконечности.

Если в какой-либо точке нет электрического заряда, то через эту точку проходит только одна силовая линия. Таким образом, силовые линии пересекаются только в тех точках, где находятся электрические заряды.

Электрическое поле обладает важным свойством — оно подчиняется принципу суперпозиции. Этот принцип состоит в том, что

напряженность поля E , создаваемая совокупностью заряженных тел, равна сумме напряженностей полей E_1, E_2, \dots , создаваемых каждым из тел в отдельности: $E = E_1 + E_2 + \dots$.

Принцип суперпозиции сводит определение напряженности поля заряженного тела сложной формы (или ряда тел) к вычислению суммы напряженностей полей, порождаемых каждым бесконечно малым элементом тела. Таким образом, в конечном счете нужно знать лишь напряженность поля, создаваемую точечным зарядом. Но ее можно найти с помощью закона Кулона: на расстоянии r от заряда q она составляет

$$E = \frac{nq}{4\pi\epsilon_0 r^2},$$

где n — единичный вектор, проведенный из точки, в которой находится заряд q , в точку, в которой находится пробный заряд.

Электрическое поле неподвижных и не изменяющихся со временем электрических зарядов (такое поле называют *электростатическим*) обладает еще одним важным свойством — оно является *потенциальным*. Это значит, что работа сил такого поля при перемещении пробного электрического заряда по произвольному замкнутому контуру равна нулю.

Чтобы убедиться в справедливости этого утверждения, рассмотрим сперва простейший случай — поле, создаваемое одним точечным зарядом q . Переместим в этом поле пробный заряд q_{II} из какой-либо точки с радиусом-вектором r в близкую к ней точку с радиусом-вектором $r + dr$, при этом мы совершим работу

$$dA = Fdr = q_{II}Edr = \frac{qq_{II}}{4\pi\epsilon_0 r^2} ndr.$$

Но $n = r/r$, поэтому $n/r^2 = -\text{grad}(1/r)$, где градиент — вектор с компонентами $(\partial/\partial x, \partial/\partial y, \partial/\partial z)$, так что $dA = -\frac{qq_{II}}{4\pi\epsilon_0} dr \text{ grad}(1/r)$. Переместим теперь заряд q по какому-либо пути на конечное расстояние, из точки a в точку b . Мы совершим работу

$$A_{a \rightarrow b} = -qq_{II} \int_{r=r_a}^{r=r_b} \text{grad} \frac{1}{r} dr = \frac{qq_{II}}{4\pi\epsilon_0} \left(\frac{1}{r_a} - \frac{1}{r_b} \right).$$

Работу электрических сил характеризуют *электростатическим потенциалом* φ — характеристикой самого поля, не зависящей от значения пробного заряда $q_{\text{п}}$, а именно: работу $A_{a \rightarrow b}$ записывают в виде $A_{a \rightarrow b} = q_{\text{п}} \{ \varphi(r_a) - \varphi(r_b) \}$. Очевидно, для точечного заряда q потенциал

$$\varphi(r) = \frac{q}{4\pi\epsilon_0 r},$$

где r — расстояние от заряда до точки наблюдения. Это соотношение — просто другая форма записи закона Кулона. Напряженность поля связана с потенциалом соотношением

$$\mathbf{E} = - \text{grad } \varphi.$$

Формула для работы $A_{a \rightarrow b}$ показывает, что работа электрических сил между двумя точками зависит от относительного положения этих точек, но не от пути, по которому перемещался заряд. В частности, если точки a и b совпадают, то $A = 0$. Иными словами, *работа электрических сил по любому замкнутому контуру L равна нулю*:

$$\oint_L \mathbf{E} d\mathbf{r} = 0.$$

Отсюда можно заключить, что *в электрическом поле не может быть замкнутых силовых линий*. Действительно, при перемещении пробного заряда вдоль замкнутой силовой линии мы получили бы отличную от нуля работу.

Линейный интеграл от какого-либо векторного поля, например $\mathbf{E}(\mathbf{r})$, вдоль замкнутого контура L называют *циркуляцией* векторного поля вдоль контура. Если контур охватывает бесконечно малую поверхность, то циркуляция пропорциональна площади Δs этой поверхности. Более точно: если ввести вектор площади Δs , ориентированный нормально к поверхности, то циркуляция равна скалярному произведению Δs на некоторый вектор, который называют *ротором* или *вихрем* поля и обозначают символом rot . Для произвольного контура L справедлива формула

$$\oint_L \mathbf{A} d\mathbf{l} = \int_{\Sigma} \text{rot } \mathbf{A} \cdot d\mathbf{s},$$

где $d\mathbf{l}$ — элемент контура, Σ — произвольная поверхность, опирающаяся на L , и $d\mathbf{s}$ — элемент ее площади, ориентированный определенным образом: направление $d\mathbf{s}$ связано с направлением обхода L правилом правого винта. Это соотношение называют *формулой Стокса*.

Явный вид ротора можно найти, если рассмотреть три бесконечно малые площадки в виде прямоугольников со сторонами, параллельными осям координат. Мы получим тогда

$$\left(\text{rot } \mathbf{A} \right)_z = \frac{\partial A_x}{\partial y} - \frac{\partial A_y}{\partial x}$$

и аналогичные соотношения для других проекций $\text{rot } \mathbf{A}$. В векторной форме ротор можно представить в виде векторного произведения оператора набла:

$$\nabla = \mathbf{i} \frac{\partial}{\partial x} + \mathbf{j} \frac{\partial}{\partial y} + \mathbf{k} \frac{\partial}{\partial z}$$

и вектора \mathbf{A} :

$$\text{rot } \mathbf{A} = [\nabla \mathbf{A}].$$

Так как циркуляция электростатического поля вдоль любого замкнутого контура равна нулю, то

$$\text{rot } \mathbf{E} = 0.$$

По этой причине электростатическое поле называют *безвихревым*.

Величину $U = qq_{\Pi}/(4\pi\epsilon_0 r)$ можно рассматривать не только как потенциальную энергию пробного заряда q_{Π} в поле, создаваемом зарядом q , но и как взаимную потенциальную энергию двух равноправных зарядов q и q_{Π} , находящихся на расстоянии r друг от друга. Пользуясь такой интерпретацией U , можно найти общую потенциальную энергию W системы точечных зарядов q_1, q_2, \dots , расположенных в точках $\mathbf{r}_1, \mathbf{r}_2, \dots$. Эта энергия равна сумме взаимных потенциальных энергий для каждой из пар зарядов, т. е.

$$W = \frac{1}{4\pi\epsilon_0} \left\{ \frac{q_1 q_2}{|\mathbf{r}_1 - \mathbf{r}_2|} + \frac{q_1 q_3}{|\mathbf{r}_1 - \mathbf{r}_3|} + \frac{q_2 q_3}{|\mathbf{r}_2 - \mathbf{r}_3|} + \dots \right\}.$$

Зная потенциал поля, создаваемого одним зарядом, можно найти потенциал поля любой системы зарядов. Для этого достаточно вспомнить, что согласно принципу суперпозиции напряженности полей разных зарядов складываются, поэтому складываются и потенциалы:

$$\varphi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \sum_i \frac{q_i}{|\mathbf{r} - \mathbf{r}_i|}.$$

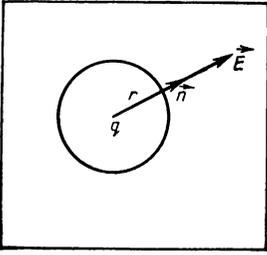
Здесь \mathbf{r}_i — координаты заряда q_i , а сумма берется по всем зарядам системы.

Если потенциал φ вдоль какого-либо направления остается неизменным, то поле \mathbf{E} не имеет составляющей вдоль этого направления. Поэтому *эквипотенциальные поверхности, т. е. поверхности $\varphi(\mathbf{r}) = \text{const}$, ортогональны к силовым линиям*. В частном случае одного точечного заряда эквипотенциальными поверхностями являются сферы, в центре которых находится заряд, а силовые линии направлены вдоль радиусов.

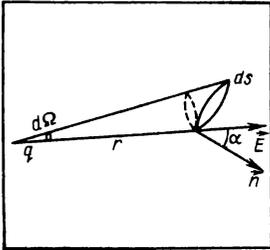
Единицу потенциала называют вольт (В). При перемещении заряда 1 Кл из одной точки поля в другую, разность потенциалов между которыми равна 1 В, нужно совершить работу, равную 1 Дж (1 Дж = 1 Кл · 1 В).

1.2. Теорема Гаусса и уравнение Пуассона

Поток напряженности электрического поля через какую-либо поверхность Σ определяют как



1.2.
К вычислению потока напряженности электрического поля точечного заряда через сферу



1.3.
К выводу теоремы Гаусса

$$N = \int_{\Sigma} \mathbf{E} ds,$$

где ds — элемент поверхности, т. е. вектор, направленный в каждой точке по нормали \mathbf{n} к поверхности и равный по модулю элементу площади ds .

Проще всего вычислить поток напряженности электрического поля точечного заряда q через сферу, в центре которой расположен заряд (рис. 1.2). В этом случае напряженность \mathbf{E} поля направлена в каждой точке по \mathbf{n} и равна $q/(4\pi\epsilon_0 r^2)$, где r — радиус сферы. Поскольку площадь поверхности сферы есть $4\pi r^2$, значение потока напряженности не зависит от радиуса сферы и равно q/ϵ_0 (рис. 1.2).

Легко убедиться, что формула $N = q/\epsilon_0$ справедлива не только для сферы, но и для любой другой замкнутой поверхности, окружающей точечный заряд. В самом деле, если ввести угол α между \mathbf{n} и \mathbf{E} (рис. 1.3), то $\cos\alpha \cdot ds = r^2 d\Omega$, где $d\Omega$ — элемент телесного угла. Далее, $E = q/(4\pi\epsilon_0 r^2)$, поэтому $dN = qd\Omega/(4\pi\epsilon_0)$, а поскольку всей замкнутой поверхности соответствует телесный угол 4π , то $N = q/\epsilon_0$.

Этот результат можно обобщить на произвольное число как угодно расположенных зарядов q_i ($i = 1, 2, 3, \dots$). Действительно, электрическое поле системы зарядов равно, согласно принципу суперпозиции, сумме полей, создаваемых каждым зарядом в отдельности. Поэтому и поток N равен сумме потоков поля, создаваемых каждым из зарядов. Мы приходим к важному выводу:

поток электрического поля через произвольную замкнутую поверхность всегда равен сумме зарядов, находящихся внутри нее, деленной на ϵ_0 :

$$N = \int_{\Sigma} \mathbf{E} ds = \frac{1}{\epsilon_0} \sum_i q_i.$$

Это общее положение называют теоремой Гаусса.

Используя теорему Гаусса, легко находить поля, создаваемые симметрично заряженными телами. Рассмотрим, например, шар со сферически симметричным распределением объемной плотности заряда $\rho(r)$. В этом случае из соображений симметрии ясно, что поле направлено по радиусам и может зависеть только от расстояния до центра шара (рис. 1.4). Поэтому в качестве замкнутой поверхности удобно выбрать сферу, центр которой совпадает с центром шара. Если радиус этой поверхности равен r , то поток напряженности электрического

поля через нее составит $4\pi r^2 E$. Приравнявая эту величину q_r/ϵ_0 , где q_r — заряд внутри поверхности радиуса r , получим $E = q_r/(4\pi\epsilon_0 r^2)$ и, следовательно, $\varphi = +q_r/(4\pi\epsilon_0 r)$.

Если $r > R$ (R — радиус шара), то под q_r следует понимать полный заряд шара q . Таким образом (для любого сферически симметричного распределения заряда внутри шара), напряженность поля и потенциал вне шара

$$E = q/(4\pi\epsilon_0 r^2), \quad \varphi = q/(4\pi\epsilon_0 r) \quad (r > R)$$

определяются теми же формулами, что и в случае точечного заряда.

Если весь заряд сосредоточен на поверхности шара ($r = R$), то, очевидно, напряженность поля внутри шара равна нулю. Если шар заряжен равномерно с объемной плотностью заряда ρ , то $q_r = \frac{4}{3}\pi r^2 \rho$, так что напряженность поля внутри шара пропорциональна r , а потенциал меняется как r^2 :

$$E = \frac{\rho}{3\epsilon_0} r, \quad \varphi = \frac{\rho}{6\epsilon_0} (R^2 - r^2) + \frac{q}{4\pi\epsilon_0 R}$$

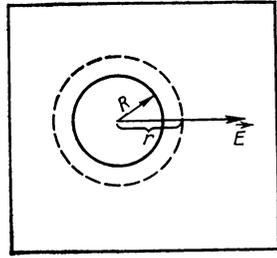
(постоянная в выражении для φ выбрана так, чтобы получить $\varphi = q/(4\pi\epsilon_0 R)$ на поверхности шара).

Определим напряженность поля, создаваемого равномерно заряженной плоскостью Σ , которую мы будем считать бесконечной. Из соображений симметрии следует, что поле в этом случае направлено нормально к плоскости и не зависит от координат вдоль плоскости. Выберем в качестве замкнутой поверхности поверхность параллелепипеда, две грани которого Σ_1 и Σ_2 параллельны Σ и находятся на равном расстоянии от нее (рис. 1.5). Поток напряженности поля отличен от нуля только через эти две грани, а так как поле на этих гранях повсюду одинаково и направлено перпендикулярно граням, то поток $N = 2ES$, где S — площадь грани. С другой стороны, согласно теореме Гаусса, $N = \epsilon_0^{-1}\sigma S$, где σ — поверхностная плотность заряда. Таким образом,

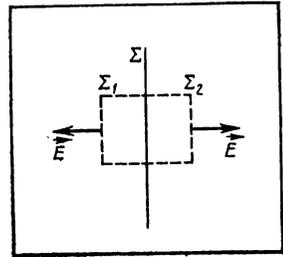
$$E = (2\epsilon_0)^{-1}\sigma.$$

Мы видим, что напряженность поля не зависит от расстояния до заряженной плоскости (этот результат связан, очевидно, с предположением о безграничности плоскости Σ). Для конечной плоскости значение напряженности близко к предельному $(2\epsilon_0)^{-1}\sigma$ близки ее центральной части и сильно отличается около краев. Потенциал безграничной заряженной плоскости

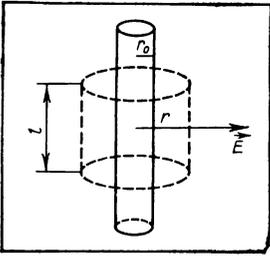
$$\varphi(x) = (2\epsilon_0)^{-1}\sigma x + \varphi_0,$$



1.4. Поле заряженного шара



1.5. Поле заряженной плоскости



1.6.
Поле заряженного цилиндра

где φ_0 — потенциал на самой плоскости и x — расстояние до нее.

Определим теперь напряженность поля бесконечно длинного кругового цилиндра радиуса r_0 (рис. 1.6). Напряженность поля в этом случае направлена по радиусам и зависит только от расстояния r до оси цилиндра. Выбирая в качестве поверхности коаксиальный цилиндр радиуса r и высоты l , имеем $N = 2\pi r l E$. С другой стороны, согласно теореме Гаусса, $N = l\tau/\epsilon_0$, где τ — линейная плотность заряда, поэтому

$$E = \frac{\tau}{2\pi\epsilon_0 r}, \quad \varphi = -\frac{\tau}{2\pi\epsilon_0} \ln \frac{r}{r_0} + \varphi_0,$$

где φ_0 — потенциал на поверхности цилиндра.

Теорема Гаусса может быть использована не только для того, чтобы находить напряженности полей при симметричном расположении зарядов. С ее помощью может быть выведено также важное дифференциальное соотношение, связывающее потенциал $\varphi(\mathbf{r})$ с объемной плотностью заряда в той же точке $\rho(\mathbf{r})$. Чтобы вывести это соотношение, окружим рассматриваемую точку P произвольной малой замкнутой поверхностью и найдем поток напряженности электрического поля ΔN через эту поверхность. Согласно теореме Гаусса, $N = \epsilon_0^{-1} \langle \rho \rangle \Delta V$, где ΔV — объем, ограниченный поверхностью, и $\langle \rho \rangle$ — средняя плотность заряда внутри объема ΔV . Стягивая поверхность к точке P , получим $\lim_{\Delta V \rightarrow 0} \Delta N / \Delta V = \rho / \epsilon_0$, где ρ — объемная плотность заряда в точке P . Величина $\lim_{\Delta V \rightarrow 0} \Delta N / \Delta V$ носит название *дивергенции электрического поля* в точке и обозначается символом $\text{div } \mathbf{E}$. Таким образом,

$$\text{div } \mathbf{E} = \rho / \epsilon_0.$$

Это важное соотношение, связывающее напряженность поля \mathbf{E} с объемной плотностью заряда, называют уравнением Пуассона.

Эквивалентное уравнение получается, если вспомнить, что $\mathbf{E} = -\text{grad } \varphi$. Дифференциальный оператор div grad называют лапласианом (или *оператором Лапласа*) и обозначают Δ . Таким образом,

$$\Delta \varphi = -\rho / \epsilon_0.$$

В отсутствие зарядов это уравнение принимает вид (уравнение Лапласа)

$$\Delta \varphi = 0.$$

Явный вид дивергенции поля (как и любого вектора) в декартовых координатах легко установить, выбирая в качестве ΔV прямоугольный параллелепипед со сторонами Δx , Δy , Δz , параллельными координатным осям. Получаем

$$\operatorname{div} \mathbf{E} = \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z}.$$

Вспомянув определение градиента, мы видим, что в декартовых координатах

$$\Delta \varphi = \partial^2 \varphi / \partial x^2 + \partial^2 \varphi / \partial y^2 + \partial^2 \varphi / \partial z^2.$$

Если какой-либо скаляр φ зависит только от длины вектора \mathbf{r} , то $\operatorname{grad} \varphi = \frac{\mathbf{r}}{r} \frac{d\varphi}{dr}$; выбирая в качестве ΔV в определении дивергенции шар радиуса r с центром в начале координат и устремляя $r \rightarrow 0$, имеем

$$\Delta \varphi = \operatorname{div} \operatorname{grad} \varphi = \frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d\varphi}{dr} \right).$$

1.3. Электрические диполь и квадруполь

Определим потенциал, создаваемый *диполем*, т. е. двумя смещенными друг относительно друга точечными электрическими зарядами, равными по значению и противоположными по знаку (рис. 1.7). Согласно принципу суперпозиции, потенциал в точке P

$$\varphi(\mathbf{r}) = \frac{q}{4\pi\epsilon_0} \left\{ \frac{1}{|\mathbf{r} - \mathbf{r}_+|} - \frac{1}{|\mathbf{r} - \mathbf{r}_-|} \right\},$$

где \mathbf{r}_+ , \mathbf{r}_- — координаты зарядов, модуль которых равен q . Наибольшее значение для приложений имеет тот случай, когда расстояние от диполя до точки наблюдения P велико по сравнению с длиной диполя: $l = |\mathbf{r}_+ - \mathbf{r}_-|$, $r \gg l$. В этом случае $|\mathbf{r} - \mathbf{r}_{\pm}| = r - \mathbf{r} \cdot \mathbf{r}_{\pm} / r$, поэтому

$$\varphi(\mathbf{r}) = \mathbf{r} \cdot \mathbf{d} / (4\pi\epsilon_0 r^3),$$

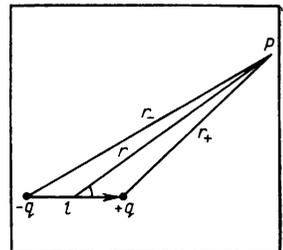
где $\mathbf{d} = q(\mathbf{r}_+ - \mathbf{r}_-)$. Вектор \mathbf{d} называют *электрическим моментом диполя*. Напряженность электрического поля диполя

$$\mathbf{E}(\mathbf{r}) = -\operatorname{grad} \varphi(\mathbf{r}) = \frac{3\mathbf{n}(\mathbf{n} \cdot \mathbf{d}) - \mathbf{d}}{4\pi\epsilon_0 r^3},$$

где $\mathbf{n} = \mathbf{r}/r$. По модулю

$$E(\mathbf{r}) = \frac{d}{4\pi\epsilon_0 r^3} \sqrt{1 + 3 \cos^2 \vartheta}$$

(ϑ — угол между \mathbf{n} и \mathbf{d}). Мы видим, что напряженность поля диполя убывает с расстоянием как r^{-3} (т. е. быстрее, чем убывающая как r^{-2} напряженность поля заряда) и не обладает сферической симметрией, а анизотропно; напряженность поля вдоль направления \mathbf{d} вдвое больше напряженности поля в перпендикулярном направлении (силовые линии поля диполя изображены на рис. 1.1, 2).



1.7.
К выводу потенциала диполя

Найдем энергию диполя W во внешнем электрическом поле \mathbf{E} . Если $\varphi(\mathbf{r})$ — потенциал этого поля, то $W = q[\varphi(\mathbf{r}_+) - \varphi(\mathbf{r}_-)]$. Пусть напряженность поля \mathbf{E} мало изменяется на длине l ; тогда $\varphi(\mathbf{r}_+) - \varphi(\mathbf{r}_-) = (\mathbf{r}_+ - \mathbf{r}_-) \text{grad } \varphi$ и, следовательно,

$$W = -d\mathbf{E}.$$

Если поле однородно и диполь *жесткий* (т. е. модуль его электрического момента не меняется со временем), то единственной величиной, которая может изменяться, является угол θ между \mathbf{d} и \mathbf{E} . Из механики известно, что $-dW/d\theta = K$, где K — момент силы; таким образом, $K = Ed \sin\theta$, или в векторной форме

$$\mathbf{K} = [d\mathbf{E}]_z.$$

Эта формула применима и в случае неоднородного поля; в неоднородном поле наряду с вращающим моментом \mathbf{K} на жесткий диполь будет действовать еще и сила $\mathbf{f} = \text{grad}(d\mathbf{E})$. Так как $d\mathbf{E} = d_x E_x + d_y E_y + d_z E_z$, то ее проекция на ось x равна $f_x = d_x \partial E_x / \partial x + d_y \partial E_y / \partial x + d_z \partial E_z / \partial x$. Но $E_y = -\partial\varphi / \partial y$, $E_z = -\partial\varphi / \partial z$, поэтому $\partial E_y / \partial x = -\partial^2\varphi / \partial x \partial y = \partial E_x / \partial y$ и $\partial E_z / \partial x = \partial E_x / \partial z$ и

$$f_x = d_x \frac{\partial E_x}{\partial x} + d_y \frac{\partial E_x}{\partial y} + d_z \frac{\partial E_x}{\partial z}.$$

Используя снова векторный дифференциальный оператор набла (см. § 1.1) $\nabla = \mathbf{i} \frac{\partial}{\partial x} + \mathbf{j} \frac{\partial}{\partial y} + \mathbf{k} \frac{\partial}{\partial z}$ ($\mathbf{i}, \mathbf{j}, \mathbf{k}$ — орты вдоль осей x, y, z), можно записать $f_x = (d\nabla)E_x$ и, следовательно,

$$\mathbf{f} = (d\nabla) \mathbf{E}.$$

К понятию диполя можно прийти, решая общую задачу о потенциале системы зарядов вдали от этой системы. Согласно принципу суперпозиции,

$$\varphi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \sum_a \frac{q_a}{|\mathbf{r} - \mathbf{r}_a|}$$

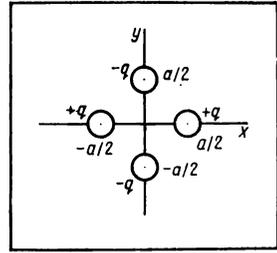
где \mathbf{r}_a — радиус-вектор заряда q_a . Разложим это выражение в ряд по степеням r_a ; получим $\varphi = \varphi^{(0)} + \varphi^{(1)} + \dots$, где член $\varphi^{(n)}$ пропорционален $r_a^{(n+1)}$. Очевидно, $\varphi^{(0)} = (4\pi\epsilon_0 r)^{-1} \sum_a q_a$ — это просто потенциал, создаваемый суммарным зарядом системы, помещенным в какую-либо ее точку (которую мы выбрали в качестве начала координат $r = 0$). Далее,

$$\varphi^{(1)}(\mathbf{r}) = \frac{r\mathbf{d}}{4\pi\epsilon_0 r^3}, \quad \mathbf{d} = \sum_a e_a \mathbf{r}_a$$

Если система состоит всего из двух равных (по величине) зарядов, противоположных по знаку, то эти формулы переходят в приведенные выше формулы для

потенциала диполя; эти формулы дают, таким образом, обобщение понятия дипольного момента.

Обратим внимание на то, что отдельные члены ряда $\varphi^{(0)}$, $\varphi^{(1)}$, $\varphi^{(2)}$, ... (но не их сумма φ) зависят от выбора начала координат $r = 0$: например, если сместить начало координат на вектор \mathbf{b} , то дипольный момент изменится на $\Delta \mathbf{d} = \mathbf{b} \sum_a q_a$. Такая неопределен-



1.8.
Пример квадруполья

ность, однако, не играет существенной роли. Дело в том, что пользоваться понятием диполя приходится главным образом для электрически нейтральной системы — только в этом случае поле диполя определяет поле системы зарядов, а не является лишь малой поправкой к полю суммарного заряда системы. Но как раз в этом случае $\sum_a q_a = 0$, так что значение \mathbf{d} не зависит от выбора начала координат.

Рассмотрим теперь следующий член $\varphi^{(2)}$ в разложении потенциала. Замечая, что $\frac{\partial^2}{\partial x_i \partial x_j} \frac{1}{r} = \frac{1}{r^5} (3x_i x_j - r^2 \delta_{ij})$, получим

$$\varphi^{(2)}(\mathbf{r}) = \frac{1}{8\pi\epsilon_0 r^3} \sum_{i,j} n_i n_j Q_{ij}; \quad Q_{ij} = \sum_a q_a (3x_{ai} x_{aj} - r_a^2 \delta_{ij}),$$

где $\mathbf{n} = \mathbf{r}/r$ и δ_{ij} — единичный тензор, или символ Кронекера ($\delta_{ij} = 1$, если $i = j$, и $\delta_{ij} = 0$, если $i \neq j$); совокупность величин Q_{ij} называют тензором квадрупольного момента. Так как тензор Q_{ij} симметричен ($Q_{ij} = Q_{ji}$) и $\sum_i Q_{ii} = 0$, то число независимых компонент

у него — пять. Потенциал квадруполья убывает с расстоянием как r^{-3} и, следовательно, поле — как r^{-4} . Таким образом, если у системы зарядов отличен от нуля дипольный момент (не говоря уже о суммарном заряде), то потенциал $\varphi^{(2)}$ определяет в выражении для φ лишь малую поправку. Если же $\sum_a q_a = 0$ и симметрия расположения зарядов

такова, что $d = 0$, то $\varphi^{(2)}$, или, как говорят, потенциал квадруполья будет главным членом в разложении φ по степеням r_a .

Рассмотрим в качестве примера систему из четырех зарядов q , $-q$, q , $-q$, расположенных в вершинах квадрата со стороной a (рис. 1.8). Дипольный момент такой системы равен нулю; квадрупольный момент имеет только два отличных от нуля компонента; $Q_{xx} = -Q_{yy} = 3qa^2$.

Найдем потенциал нашего квадруполья. Вводя полярный θ и азимутальный χ углы вектора \mathbf{n} ($n_x = \sin\theta \cos\chi$, $n_y = \sin\theta \sin\chi$, $n_z = \cos\theta$), найдем $\sum_{i,j} n_i n_j Q_{ij} = 3qa^2 \sin^2\theta \cos 2\chi$. Поэтому потенциал

$$\varphi^{(2)}(\mathbf{r}) = \frac{3qa^2}{8\pi\epsilon_0 r^3} \sin^2\theta \cos 2\chi.$$

Взяв градиент от $\varphi^{(2)}(\mathbf{r})$, можно найти напряженность поля квадруполья.

1.4. Проводник в электростатическом поле

По своему поведению в *электростатическом* (т. е. постоянном во времени электрическом) поле все тела разделяются на два класса: *проводники* и *изоляторы*, или диэлектрики.

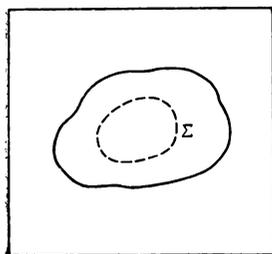
Проводник — это тело, в котором в электростатическом поле возникает направленное движение зарядов, т. е. электрический ток. **Изолятор** — тело, в котором достаточно слабое электростатическое поле тока не вызывает.

В очень сильных полях ток возникает и в диэлектрике (это явление называют *электрическим пробоем*).

Проще всего понять, проводником или диэлектриком является вещество, в случае газа. Нейтральный газ — диэлектрик, потому что составляющие его атомы или молекулы электрически нейтральны и поэтому не должны двигаться во внешнем электрическом поле. Наоборот, ионизованный газ — *плазма* — проводник, так как электроны и ионы в нем не связаны между собой и движутся независимо во внешнем поле. Аналогичная ситуация возникает в случае растворов: растворы, содержащие нейтральные молекулы растворенного вещества, — изоляторы, а растворы, содержащие ионы — так называемые растворы *электролитов*, — проводники. (Пример — слабый раствор поваренной соли, содержащий не молекулы NaCl , а ионы Na^+ и Cl^- .)

Сложнее обстоит дело с конденсированными телами. Почему одни кристаллы и жидкости проводят ток (например, металлы), а другие — нет? Этим вопросом мы займемся во второй части книги, а пока изучим те свойства этих двух классов тел, которые не зависят от деталей их структуры.

Легко понять, что если внести проводник во внешнее электрическое поле, то движение зарядов в нем вскоре должно прекратиться. В самом деле, в противном случае проводник, помещенный в электрическое поле, мог бы быть использован для построения вечного двигателя первого рода: движущиеся заряды совершали бы непрерывно работу (в частности, в проводнике непрерывно выделялась бы теплота), состояние же зарядов, создающих электростатическое поле, не менялось бы. Это значит, что электрическое поле внутри проводника должно обратиться в нуль.



1.9.

К доказательству равенства нулю объемной плотности заряда в проводнике

Итак, в состоянии равновесия электростатическое поле внутри проводника равно нулю. Отсюда следует, что нулю равна и объемная плотность заряда внутри проводника. Действительно, проведем в толще проводника произвольную замкнутую поверхность Σ (рис. 1.9). Так как напряженность поля на поверхности Σ равна нулю, то нулю равен и поток напряженности поля через нее. Отсюда, согласно теореме Гаусса, следует, что заряд внутри поверхности равен нулю, а так как поверхность Σ может быть вы-

брана произвольно, то объемная плотность заряда в проводнике равна нулю.

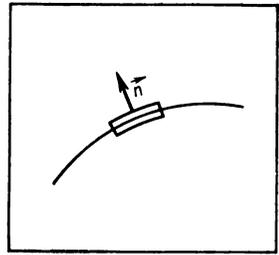
Поэтому если сообщить проводнику некоторый объемный заряд, то отдельные заряды (или, как говорят, *носители тока*) уйдут из объема проводника и окажутся на его поверхности. При этом они расположатся там таким образом, чтобы в нуль обращались напряженность поля внутри проводника и ее тангенциальная составляющая E_t на поверхности проводника (если бы последнее условие не выполнялось, то происходило бы непрерывное движение зарядов вдоль поверхности, что невозможно по той же причине, что и непрерывное движение зарядов внутри проводника). Мы видим, что

вне проводника вблизи его поверхности напряженность поля должна быть направлена нормально к поверхности проводника.

Ее значение E_n легко связать с поверхностной плотностью заряда σ . Используем для этого теорему Гаусса, выбирая в качестве поверхности Σ цилиндр с основаниями, бесконечно близкими к поверхности и параллельными ей, одно из которых находится внутри проводника, а другое — вне его (рис. 1.10). Отличен от нуля только поток напряженности электрического поля через внешнее основание цилиндра, равный $E_n \Delta S$ (ΔS — площадь основания), заряд же внутри цилиндра равен $\sigma \Delta S$, поэтому $E_n = \epsilon_0^{-1} \sigma$. Если $\sigma > 0$, то и $E_n > 0$, т. е. напряженность поля вблизи поверхности направлена по внутренней нормали n .

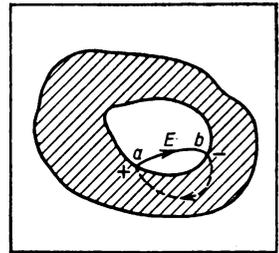
Так как напряженность поля и объемная плотность заряда внутри проводника равны нулю, то следует ожидать, что при удалении из проводника части вещества поле в возникшей полости по-прежнему остается равным нулю. Чтобы убедиться в этом, предположим противное, а именно: поле внутри полости отлично от нуля. Так как внутри вещества проводника поле равно нулю, то силовые линии внутри полости должны начинаться и кончаться на внутренней поверхности проводника, причем на концах силовых линий должны находиться поверхностные заряды противоположных знаков. Возьмем теперь какую-либо из таких силовых линий и замкнем ее произвольной линией, лежащей в проводниковой оболочке (рис. 1.11). Для циркуляции напряженности электрического поля вдоль такого замкнутого контура получим

$$\oint \mathbf{E} d\mathbf{r} = \int_a^b \mathbf{E} d\mathbf{r}$$
, а так как $\mathbf{E} \parallel d\mathbf{r}$, то этот интеграл положителен, тогда как циркуляция напряженности электростатического поля вдоль



1.10.

К выводу связи между E_n и поверхностной плотностью заряда



1.11.

Поле внутри полости, окруженной проводником

любого замкнутого контура должна быть равна нулю. Отсюда следует, что *напряженность поля внутри замкнутой полости, окруженной проводником, равна нулю.*

Это свойство поля тесно связано с теоремой Гаусса, которая, в свою очередь, является следствием закона Кулона. Если бы закон Кулона был несправедлив и сила электрического взаимодействия изменялась с расстоянием по закону $r^{-2+\nu}$, где $\nu \neq 0$, то напряженность поля в полости должна была бы отличаться от нуля. Экспериментальные измерения напряженности поля в полости позволили оценить верхнюю границу величины ν ; она оказалась меньше, чем 10^{-6} .

Этот факт лежит в основе *метода электростатической защиты*: для предохранения от действия посторонних электростатических полей прибор следует окружить слоем проводника, например замкнутой металлической оболочкой произвольной толщины. Эту оболочку не обязательно делать сплошной, ее можно взять и в виде сетки. Уже из размерных соображений ясно, что наружное возмущающее поле проникает внутрь объема, ограниченного металлической сеткой, на длину, сравнимую с размером отверстия, т. е. с размером ячейки сетки. При достаточной густоте сетки такое «провисание» поля очень мало.

Рассмотрим проводник, состоящий из двух сфер разного радиуса, соединенных тонкой проволокой. В отсутствие проволоки потенциалы сфер имели бы значения $\varphi_1 = q_1/(4\pi\epsilon_0 r_1)$ и $\varphi_2 = q_2/(4\pi\epsilon_0 r_2)$, где q_1, q_2 — заряды и r_1, r_2 — радиусы сфер, и напряженность поля на поверхностях сфер $E_1 = \varphi_1/r_1$, $E_2 = \varphi_2/r_2$. Если теперь соединить сферы тонкой проволокой, то это приведет в основном только к выравниванию потенциалов сфер (распределение зарядов на сферах почти не изменится). Но если $\varphi_1 = \varphi_2$, то $E_1/E_2 = r_2/r_1$, т. е. напряженности поля обратно пропорциональны радиусам сфер. Поверхностные плотности зарядов $\sigma_1 = q_1/(4\pi r_1^2)$, $\sigma_2 = q_2/(4\pi r_2^2)$ также обратно пропорциональны радиусам сфер: $\sigma_1/\sigma_2 = r_2/r_1$.

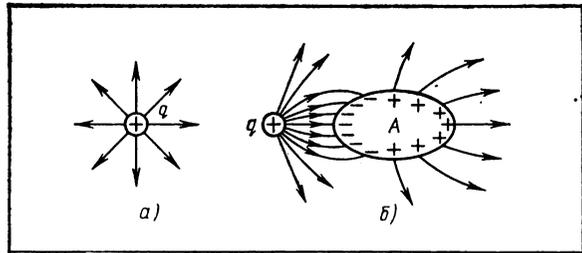
Полученный результат верен и в случае более общем, чем две соединенные проводником сферы разного радиуса. А именно: *чем меньше радиус кривизны проводника, тем больше напряженность поля вблизи его поверхности и тем больше поверхностная плотность заряда.* Поэтому особенно велики напряженности полей вблизи острых краев проводника.

1.5. Электростатическая индукция

Выясним, что произойдет, если незаряженный проводник внести в заданное электростатическое поле, созданное какими-либо внешними зарядами. Напряженность электростатического поля внутри проводника должна быть равной нулю. Но до внесения проводника поле в объеме, занимаемом впоследствии проводником, было отлично от нуля; теперь же оно должно обратиться в нуль. Следовательно, на проводнике, внесенном в электростатическое поле, должны возникнуть электрические заряды, поле которых, накладываясь на исходное поле внешних зарядов, должно приводить к полной компенсации поля внутри проводника. Это явление — возникновение зарядов на проводнике, внесенном в электростатическое поле, —

1.12.

Искажение поля при внесении в него проводника



называют *электростатической индукцией*. Из закона сохранения заряда следует, что суммарный заряд, индуцируемый на проводнике должен быть равен нулю. Таким образом,

при электростатической индукции возникают заряды обоих знаков, причем общие количества их одинаковы.

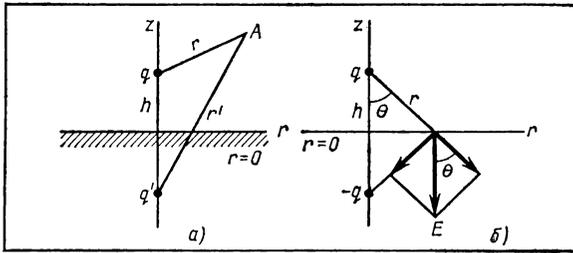
Пусть, например, в поле, создаваемое заряженным шаром, внесен проводник. До внесения проводника силовые линии поля имели вид прямых, исходящих из центра шара (рис. 1.12, а). После внесения проводника А (рис. 1.12, б) поле в его объеме должно исчезнуть. Это значит, что некоторые из силовых линий, исходящих из q , должны закончиться на поверхности А, некоторые же, после разрыва, должны начаться на этой поверхности. Но там, где начинаются и кончаются силовые линии, находятся электрические заряды. Таким образом, проводник, внесенный в электростатическое поле, должен заряжаться, причем на близлежащей к исходному заряду q части поверхности проводника появятся заряды, противоположные по знаку заряду q , а на более удаленной части поверхности А — заряды того же знака, что и заряд q .

Силовые линии искаженного поля подходят ортогонально к поверхности внесенного проводника, а плотность индуцированного поверхностного заряда σ связана с нормальной составляющей поля E_n известным уже нам соотношением $\sigma = \epsilon_0 E_n$. Далее, поскольку $\mathbf{E} = -\text{grad } \phi$, из равенства нулю напряженности поля вытекает постоянство потенциала:

потенциал имеет одно и то же значение во всех точках проводника как внутри, так и на его поверхности.

Это значит, что поверхность проводника является эквипотенциальной. Подчеркнем, что эти выводы справедливы независимо от того, вносится ли проводник во внешнее электростатическое поле или поле создается зарядом, сообщаемым проводнику.

Так как объемная плотность заряда в проводнике равна нулю, то потенциал поля, создаваемого заряженными проводниками, удовлетворяет уравнению Лапласа $\Delta\phi = 0$. Если ищется потенциал поля, создаваемого проводником, которому сообщен определенный заряд, то должно быть взято то решение этого уравнения, которое имеет постоянное значение на поверхности проводника. При этом, так как $E_n =$



1.13. Точечный заряд в присутствии плоской проводящей пластины

$= -\partial\varphi/\partial n$ и $E_n = \epsilon_0^{-1}\sigma$, суммарный заряд проводника Q должен быть связан с φ соотношением

$$Q = -\epsilon_0 \int \frac{\partial\varphi}{\partial n} ds,$$

где интегрирование совершается по поверхности проводника и ds — элемент этой поверхности ($\partial/\partial n = \mathbf{n}\nabla$ — производная по направлению нормали).

Если проводник находится во внешнем поле, то должно быть взято то решение уравнения Лапласа, которое вдали от проводника, где нет искажения исходного поля, совпадает с потенциалом этого поля.

Вычисление плотности поверхностных зарядов на проводнике, помещенном в электрическое поле, в общем случае представляет собой сложную задачу; лишь в некоторых простейших случаях решение можно получить сразу, не прибегая к уравнению Лапласа.

Определим сперва поле точечного заряда в присутствии плоской проводящей пластины. Пластину считаем бесконечной, и, кроме того, заземленной, поэтому индуцированный заряд на обратной ее стороне отсутствует (ушел в землю). Эта задача эквивалентна, очевидно, задаче о нахождении поля точечного заряда в присутствии проводника, имеющего плоскую границу и занимающего полупространство $z < 0$ (рис. 1.13, а). Если бы проводника не было, то заряд q создавал бы в точке A потенциал $\varphi_q = q/(4\pi\epsilon_0 r')$. Индуцированные на поверхности проводника заряды создают дополнительный потенциал

$$\varphi_\sigma(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{\sigma ds}{|\mathbf{r} - \mathbf{r}_\sigma|},$$

где σ — плотность поверхностного заряда и \mathbf{r}_σ — радиус-вектор точки поверхности. Таким образом, измененный потенциал $\varphi = \varphi_q + \varphi_\sigma$.

Попытаемся найти φ_σ , предполагая, что действие всей заряженной плоскости $z = 0$ при $z > 0$ эквивалентно действию некоторого одного точечного заряда q' , расположенного в нижней полуплоскости. Иными словами, мы предполагаем, что $\varphi_\sigma = q'/(4\pi\epsilon_0 r')$, где r' — расстояние от фиктивного заряда q' до точки A . Легко видеть, что условие $\varphi_q + \varphi_\sigma = \text{const}$ при $z = 0$ удовлетворится, если фиктивный заряд поместить в точку, являющуюся зеркальным отражением точки нахождения истинного заряда q относительно границы проводника, и положить $q' = -q$.

Спрашивается, насколько верно найденное нами решение задачи? Ответ дается теоремой единственности решения электростатической задачи:

если каким-либо образом найдена (или угадана!) функция $\varphi(r)$, удовлетворяющая уравнению Лапласа $\Delta\varphi = 0$, причем отвечающее ей поле $E = -\text{grad } \varphi$ удовлетворяет граничным (на поверхностях проводников) условиям электростатики $E_t = 0$ и обращается в нуль при $r \rightarrow \infty$, то потенциал поля однозначно определяется этой функцией $\varphi(r)$.

Так как все эти условия в нашем случае выполнены, то $q' = -q$. А так как электростатическая задача имеет единственное решение, то мы можем быть уверены в том, что найденное нами решение правильно. Такого рода метод решения электростатической задачи, в котором действие реальных поверхностно распределенных зарядов заменяется действием некоторых фиктивных точечных зарядов (в случае нескольких поверхностей их может быть и несколько), называют методом электрических изображений.

Найти теперь плотность поверхностного заряда не представляет труда. В самом деле, зная φ , легко найти проекцию E_n поля $E = -\text{grad } \varphi$ на направление внешней нормали к поверхности проводника (рис. 1.13, б). Поле при $z = 0$ направлено по внутренней нормали, причем $E = q\cos\theta/(2\pi\epsilon_0 r^2) = qh/(2\pi\epsilon_0 r^3)$, поэтому

$$\sigma = -qh/(2\pi r^3).$$

Полный поверхностный заряд равен $-q$.

Так как поверхностные заряды имеют знак, противоположный заряду q , то они притягивают к себе этот заряд. Иными словами, заряд q притягивается к проводнику. Сила F этого притяжения равна силе притяжения зарядов q и $-q$, т. е. $F = q^2/(16\pi\epsilon_0 h^2)$.

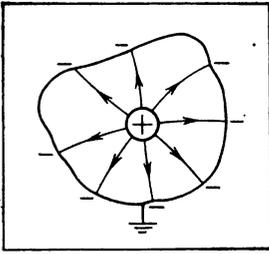
1.6. Электрическая емкость проводника

Если уединенному проводнику сообщить некоторый заряд, то он, как мы знаем, распределится по поверхности проводника, причем таким образом, что поверхность проводника станет эквипотенциальной. Будем считать потенциал, создаваемый зарядом на проводнике, равным нулю на бесконечном расстоянии от проводника. Тогда, если мы увеличим заряд проводника вдвое, вдвое увеличится и потенциал поля, создаваемого проводником, а следовательно, вдвое увеличится и потенциал проводника. Таким образом, заряд проводника q пропорционален его потенциалу φ :

$$q = C\varphi,$$

где C — коэффициент пропорциональности, зависящий от размеров и формы проводника. Его называют *электрической емкостью* или просто *емкостью* проводника. Единицу емкости называют фарад (Φ); $\Phi = \text{Кл} \cdot \text{В}^{-1}$.

В системе СГС единица емкости совпадает с единицей длины, причем $1 \Phi = 9 \cdot 10^{11}$ см. Легко видеть, что если проводник имеет форму шара, то его емкость в СГС равна R , где R — радиус шара.



1.14.
Конденсатор

Если приближать к заряженному проводнику другие проводники (незаряженные), то его потенциал уменьшается (при том же заряде на проводнике), так как на соседних проводниках индуцируются заряды, причем ближе к заряженному проводнику индуцируются заряды противоположного знака, а дальше — того же знака. Это уменьшение потенциала заряженного проводника при приближении к нему других незаряженных проводников можно интерпретировать как увеличение емкости заряженного проводника.

Особенно сильно увеличится емкость проводника, окруженного (рис. 1.14) заземленным замкнутым проводником (роль земли может играть просто массивный проводник, например шасси приемника). Действительно, в этом случае индуцированные на замкнутом проводнике заряды того же знака, что и заряд внутреннего проводника, уйдут в землю, а общий индуцированный заряд противоположного знака по модулю равен заряду внутреннего проводника, так как все силовые линии, исходящие из внутреннего проводника, закончатся на окружающем его проводнике.

Рассмотрим два проводника с зарядами q и $-q$. Пусть порождаемые ими потенциалы проводников равны φ_1 и φ_2 . Тогда разность этих потенциалов $\varphi_1 - \varphi_2$ пропорциональна заряду q .

Отношение заряда к разности потенциалов называют емкостью системы проводников, а саму систему — электрическим конденсатором.

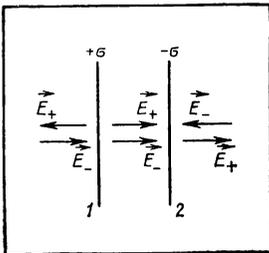
Емкость конденсатора

$$C = q/(\varphi_1 - \varphi_2)$$

определяется только геометрией проводников (обкладок конденсатора) и не зависит от q .

Простейшим является *плоский конденсатор*, обкладки которого представляют собой две параллельные проводящие пластинки. Считая их очень большими, легко найти емкость плоского конденсатора.

Напомним, что напряженность диполя заряженной плоскости равна $\sigma/(2\epsilon_0)$, где σ — поверхностная плотность заряда. Напряженности полей двух заряженных плоскостей с поверхностными плотностями зарядов σ и $-\sigma$ складываются между плоскостями и вычитаются вне них (рис. 1.15), так что внутри конденсатора $E = \sigma/\epsilon_0$, а снаружи $E = 0$. Этот вывод, конечно, справедлив только в случае очень больших пластин, т. е. пластин, размеры которых значительно больше расстояния между пластинами. Умножив напряженность поля E на расстояние d между



1.15.
Плоский конденсатор

пластинами, найдем разность потенциалов φ_1 и φ_2 обеих обкладок: $\varphi_1 - \varphi_2 = Ed = \sigma d / \epsilon_0$. Так как заряд обкладки $q = \sigma S$, где S — площадь обкладки, то емкость плоского конденсатора

$$C = \epsilon_0 S / d.$$

1.7. Энергия электрического поля

Заряжая проводник, мы одновременно сообщаем ему некоторый запас энергии. Действительно, на каждый новый заряд, переносимый на проводник, действует сила отталкивания со стороны уже находящихся на проводнике зарядов и для преодоления этой силы должна быть совершена работа.

Если на проводник приносится заряд dq' , то совершаемая работа равна, очевидно, энергии dW этого заряда в поле, созданном предыдущими элементами заряда, т. е. $dW = dq' \varphi'$, где φ' — потенциал проводника, созданный зарядами, внесенными до заряда dq' . Этот потенциал равен, согласно результатам предыдущего параграфа, $\varphi' = q' / C$, где q' — заряд проводника в момент внесения элемента заряда dq' и C — емкость проводника. Итак,

$$dW = \frac{q'}{C} dq' = d \frac{q'^2}{2C}$$

Поэтому если окончательный заряд проводника равен q , то полная работа, производимая при его накоплении,

$$W = q^2 / (2C)$$

Работа эта переходит в *электрическую энергию проводника*.

Замечая, что $q = C\varphi$, где φ — потенциал проводника, имеющего заряд q (потенциал на бесконечном расстоянии от проводника считается равным нулю), можно переписать выражение для энергии проводника в виде

$$W = 1/2 C \varphi^2 = 1/2 q \varphi$$

Обратим внимание на множитель $1/2$ в последней формуле: если бы потенциал φ создавался посторонними зарядами, а не зарядом q , то энергия заряда q в таком поле составила бы $q\varphi$.

Если имеется несколько проводников с зарядами q_1, q_2, \dots , то их электрическая энергия

$$W = 1/2 q_1 \varphi_1 + 1/2 q_2 \varphi_2 + \dots,$$

где $\varphi_1, \varphi_2, \dots$ — потенциалы проводников, создаваемые всеми зарядами q_1, q_2, \dots . В частности, энергия конденсатора

$$W = 1/2 q (\varphi_1 - \varphi_2),$$

где q — заряд обкладки и φ_1, φ_2 — потенциалы обеих обкладок. Это выражение можно переписать также в виде

$$W = \frac{1}{2} \frac{q^2}{C} = \frac{1}{2} C (\varphi_1 - \varphi_2)^2,$$

где C — емкость конденсатора.

Рассмотрим подробнее плоский конденсатор. Так как емкость его равна $\epsilon_0 S/d$, то энергия плоского конденсатора может быть представлена в виде

$$W = \frac{\epsilon_0}{2} \sigma^2 Sd,$$

где σ — поверхностная плотность зарядов. Эта формула имеет простой смысл. Так как $\sigma/(2\epsilon_0)$ представляет собой напряженность поля, создаваемого одной из пластин (обкладок), то каждая из пластин испытывает со стороны другой пластины силу притяжения

$$F = \frac{\epsilon_0}{2} \sigma^2 S.$$

Поэтому $W = Fd$, т. е. энергия W равна работе, которую нужно совершить, чтобы раздвинуть пластины на расстояние d . Вспоминая, что $E = \sigma/\epsilon_0$, можно представить энергию конденсатора в форме $W = 1/2 \sigma E V = 1/2 \epsilon_0 E^2 V$, где $V = Sd$ — объем конденсатора. Мы видим, что *электрическая энергия конденсатора пропорциональна объему, в котором находится электрическое поле*. Поэтому можно ввести понятие *плотности электрической энергии*

$$w = W/V = \epsilon_0 E^2/2,$$

которая определяется только напряженностью электрического поля.

В гауссовой системе плотность энергии определяется формулой

$$w = E^2/(8\pi).$$

Это выражение можно интерпретировать таким образом, что электрическая энергия сосредоточена не в зарядах, а в электрическом поле, и что в каждом элементе объема dV содержится энергия $dW = wdV$. Далее мы убедимся в справедливости именно такой интерпретации в самом общем случае как угодно изменяющихся в пространстве и во времени полей: формула $W = 1/2 q\varphi$, справедливая в электростатике, теряет смысл, формула же $W = \int wdV$ для энергии поля остается справедливой всегда (речь пока идет о вакууме).

На заряженный проводник действуют силы со стороны им же создаваемого поля. Так как элементы заряда находятся на поверхности проводника и отталкиваются друг от друга, эти силы приложены к поверхности, т. е. носят характер сил натяжения или внутреннего давления. К единице поверхности приложена сила $\mathbf{f} = 1/2 \sigma \mathbf{E}$, где σ — поверхностная плотность заряда. (Множитель $1/2$ входит сюда потому, что мы должны учитывать только то поле, которое создается всеми элементами поверхности за вычетом рассматриваемого элемента, а это поле равно половине полного поля.) Так как $\sigma = \epsilon_0 E$, то

$$\mathbf{f} = \frac{\epsilon_0}{2} E^2 \mathbf{n},$$

где \mathbf{n} — единичный вектор вдоль внешней нормали к поверхности проводника. Эту же формулу можно получить иначе, вспомнив общее соотношение $p = -\partial W/\partial V$ между внутренним давлением p и энергией тела W и подставив сюда вместо W энергию поля. Мы видим, что силы действуют так, как будто силовые линии представляют собой натянутые материальные упругие нити, к которым приложено упругое напряжение $\epsilon_0 E^2/2$.

Практически для определения сил, действующих на проводник, удобно поступить следующим образом. Введем геометрические параметры ξ_1, ξ_2, \dots , характеризующие форму и размеры проводника и играющие роль координат, определяющих относительное положение различных его элементов. При изменении $\delta\xi$ какого-либо из этих параметров возникнет сила f , которая совершит работу $\delta A = f\delta\xi$. Но работа при неизменном заряде проводника равна убыли энергии W , т. е. $\delta A = -\delta W$, поэтому $f = -\partial W/\partial\xi$ при $q = \text{const}$. Замечая, что в выражении $W = q^2/(2C)$ от параметров ξ зависит емкость C , получим окончательно

$$f = \frac{q^2}{2C^2} \frac{\partial C}{\partial\xi}.$$

Полагая здесь $\xi = V$, найдем избыточное давление Δp , действующее изнутри заряженного проводника. Умножив Δp на коэффициент всестороннего растяжения проводника, найдем изменение его объема. Деформация проводника (т. е. изменение его объема и формы) под действием электрических сил носит общее название *электрострикции*.

Отметим в заключение главы некоторые свойства электростатической энергии проводников.

Заряды, которые вносятся на проводники, располагаются на поверхностях проводников таким образом, чтобы энергия возникающего электростатического поля была минимальной (теорема Томсона). Внесение незаряженного проводника в поле фиксированной совокупности зарядов всегда приводит к уменьшению энергии поля.

Потенциальная энергия системы зарядов $W = (8\pi\epsilon_0)^{-1} \sum_{i,k} q_i q_k / r_{ik}$, где r_{ik} — расстояние между зарядами q_i и q_k , не имеет как функция координат зарядов минимума (и максимума). Действительно, для нахождения минимума (или максимума) потенциальной энергии, рассматриваемой как функция координат x, y, z какого-либо заряда, нужно решить уравнения $\partial W/\partial x = 0$, $\partial W/\partial y = 0$, $\partial W/\partial z = 0$ относительно x, y, z . Найденные из этих уравнений значения x, y, z должны в случае минимума удовлетворять, в частности, неравенствам

$$\frac{\partial^2 W}{\partial x^2} > 0, \quad \frac{\partial^2 W}{\partial y^2} > 0, \quad \frac{\partial^2 W}{\partial z^2} > 0.$$

Между тем каждая из функций $1/r_{ik}$ удовлетворяет при $r_{ik} \neq 0$ уравнению Лапласа $\Delta(1/r_{ik}) = 0$; следовательно, такому же уравнению удовлетворяет и функция W . Поэтому неравенства для вторых производных функции W не могут выполняться. Отсюда следует, что

статическая конфигурация электрических зарядов, находящихся только под действием электростатических сил, не может быть устойчивой (теорема Ирншоу).

Из теоремы Ирншоу следует, что если в атоме действуют в основном электрические силы, то он не может быть статической конфигурацией ядра и электронов. Иными словами, электроны в атоме должны обязательно находиться в состоянии движения. Иначе он был бы неустойчив.

Электроны в металлах можно (в некотором смысле, как будет показано в гл. 9) рассматривать как свободные заряды. Они не покидают самопроизвольно металл, так как их удерживают положительные заряды, — как говорят, ионный остов. Ионы в твердых и жидких металлах удалены друг от друга на расстоянии порядка 10^{-10} м и не разлетаются, несмотря на кулоновское отталкивание. В конечном счете силы, связывающие ионы, тоже сводятся к кулоновским; однако объяснить их можно лишь с помощью квантовой механики. Эти силы возникают из-за тождественности электронов и, грубо говоря, связаны с тем, что ионы «совместно владеют» одними и теми же электронами. (Аналогичную природу имеют силы, связывающие, например, два отрицательных иона Cl^- в молекулу Cl_2 . Мы остановимся на этом вопросе несколько подробнее в § 12.7.)

ОСНОВНЫЕ ФОРМУЛЫ

Закон Кулона в вакууме

$$F = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r_{12}^2}$$

Потенциал

$$\varphi = q/(4\pi\epsilon_0 r)$$

**Связь напряженности поля с потенциалом
Безвихревой характер электростатического поля**

$$\mathbf{E} = -\text{grad } \varphi$$

$$\text{rot } \mathbf{E} = 0, \quad \oint \mathbf{E} d\mathbf{l} = 0$$

Теорема Гаусса

$$N \equiv \int_{\Sigma} \mathbf{E} d\mathbf{s} = \frac{1}{\epsilon_0} \Sigma q_i$$

Уравнение Пуассона

$$\Delta\varphi = -\rho/\epsilon_0$$

Потенциал и поле диполя

$$\varphi_d = \frac{1}{4\pi\epsilon_0} \frac{\mathbf{r} \cdot \mathbf{d}}{r^3},$$

$$\mathbf{E}_d = \frac{3\mathbf{n}(\mathbf{n} \cdot \mathbf{d}) - \mathbf{d}}{4\pi\epsilon_0 r^3}$$

Плотность электрической энергии

$$\omega = \epsilon_0 E^2 / (8\pi)$$

Глава 2. ЭЛЕКТРОСТАТИКА ДИЭЛЕКТРИКОВ

2.1. Электрическая индукция и поляризация

В гл. 1 было показано, что смещение зарядов в молекулах диэлектрика, помещенного в электростатическое поле, приводит к уменьшению результирующего поля в диэлектрике. Это значит, что если внести заряд q в диэлектрик, то напряженность поля, создаваемая им на расстоянии r , уже не равна $q/(4\pi\epsilon_0 r^2)$, как было бы, если бы заряд находился в вакууме, а меньше. Именно: если диэлектрик однороден, изотропен и неограничен, то

$$E = \frac{q}{4\pi\epsilon_0\epsilon r^2}$$

где ϵ — характерная для каждого диэлектрика величина, большая единицы и показывающая, во сколько раз уменьшается напряженность поля по сравнению с его напряженностью в вакууме. Ее называют *диэлектрической проницаемостью* диэлектрика (иногда ее называют *относительной диэлектрической проницаемостью* и обозначают ϵ_r), а величину ϵ_0 называют *электрической постоянной*. Так, для воды $\epsilon = 81$, для спирта $\epsilon = 26$, для стекла $\epsilon = 5$, для керосина $\epsilon = 2$, для воздуха $\epsilon = 1,006$.

Уменьшение напряженности поля в диэлектрике, в свою очередь, означает, что сила электрического взаимодействия между двумя точечными зарядами q_1 и q_2 , помещенными в диэлектрик, при том же расстоянии между зарядами в ϵ раз меньше, чем в вакууме:

$$F = \frac{q_1 q_2}{4\pi\epsilon_0\epsilon r^2}$$

Электростатическое поле в диэлектриках, так же как и в вакууме, является безвихревым, т. е. $\text{rot } \mathbf{E} = 0$, и связано с потенциалом φ прежней формулой $\mathbf{E} = -\text{grad } \varphi$.

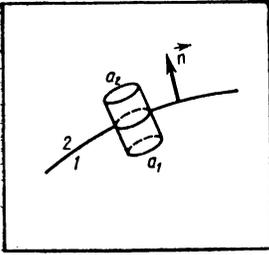
Уменьшение силы взаимодействия между зарядами в диэлектрике приводит к уменьшению потенциала в диэлектрике. Именно: заряд q создает в диэлектрике на расстоянии r потенциал

$$\varphi = \frac{q}{4\pi\epsilon_0\epsilon r}$$

Наконец, так как напряженность поля в диэлектрике в ϵ раз меньше напряженности поля в вакууме, то и поток напряженности электрического поля в диэлектрике в ϵ раз меньше, чем в вакууме. Поэтому, если ввести вектор

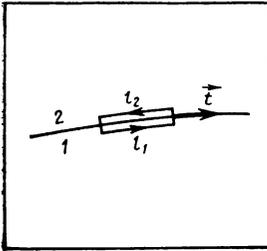
$$\mathbf{D} = \epsilon_0\epsilon\mathbf{E}$$

[его называют *электрическим смещением* (*электрической индукцией*)], то поток Ψ этого вектора через произвольную замкнутую поверхность S , как и в вакууме, равен суммарному стороннему электрическому заряду $\Sigma q^{(e)}$ внутри этой поверхности (теорема Гаусса):



2.1.

Непрерывность D_n на границе двух диэлектриков



2.2.

Непрерывность E_t на границе двух диэлектриков

$$\Psi = \int \mathbf{D} ds = \sum q^{(e)}.$$

В дифференциальной форме этот закон имеет вид $\operatorname{div} \mathbf{D} = \rho$, где ρ — объемная плотность сторонних зарядов.

Подчеркнем, что сторонние заряды $q^{(e)}$, фигурирующие в теореме Гаусса, не входят в состав молекул диэлектрика. В частности, если сторонних зарядов внутри замкнутой поверхности нет, то поток электрического смещения \mathbf{D} через эту поверхность равен нулю.

Отсюда следует, что на границе раздела двух диэлектриков должна быть непрерывной нормальная составляющая вектора \mathbf{D} . Действительно, рассмотрим цилиндр (рис. 2.1), основания которого a_1 и a_2 расположены в первом 1 и втором 2 диэлектриках непосредственно вблизи от границы раздела и параллельно ей. Тогда поток электрического смещения через верхнее основание равен $\mathbf{D}_2 \mathbf{n} \Delta S$, где \mathbf{n} — единичный вектор, нормальный к границе и идущий из первого во второй диэлектрик, и ΔS — площадь основания; поток через нижнее основание составит $-\mathbf{D}_1 \mathbf{n} \Delta S$. Если устремить высоту цилиндра к нулю, то поток электрического смещения через боковую поверхность цилиндра

будет стремиться к нулю, а так как внутри цилиндра нет сторонних зарядов, то мы получим $\mathbf{D}_2 \mathbf{n} \Delta S - \mathbf{D}_1 \mathbf{n} \Delta S = 0$, откуда

$$D_{1n} = D_{2n}$$

(нижний индекс n обозначает нормальные составляющие векторов).

Покажем теперь, что при переходе из одной среды в другую сохраняется, кроме того, тангенциальная составляющая электрического поля. Рассмотрим для этого небольшой контур, две стороны которого l_1 и l_2 расположены в первом 1 и втором 2 диэлектриках, непосредственно вблизи границы раздела и параллельно ей (рис. 2.2), и определим работу сил электростатического поля при перемещении пробного заряда q вдоль этого контура. Работа при перемещении заряда q вдоль стороны l_1 равна $qE_{1t}l_1$, где E_{1t} — составляющая поля \mathbf{E}_1 вдоль тангенциального направления; аналогично, работа при перемещении заряда вдоль стороны l_2 равна $-qE_{2t}l_2$. Если расстояние между l_1 и l_2 устремить к нулю, то будет равна нулю и работа сил поля при перемещении заряда вдоль остающихся двух сторон контура, перпендикулярных l_1 и l_2 ; при этом $l_2 \rightarrow l_1 = \Delta l$. Итак, суммарная работа равна $qE_{1t}\Delta l - qE_{2t}\Delta l$. Но поле является статическим, и поэтому эта работа должна быть равной нулю: $qE_{1t}\Delta l - qE_{2t}\Delta l = 0$, откуда и вы-

текает высказанное выше утверждение о непрерывности тангенциальной составляющей электрического поля при переходе из одной среды в другую:

$$E_{1t} = E_{2t}.$$

Тот факт, что нормальная составляющая электрической индукции и тангенциальная составляющая напряженности электрического поля непрерывны, позволяет выяснить, что происходит с силовой линией при переходе из одного диэлектрика в другой. Легко видеть, что силовая линия претерпевает преломление на границе раздела двух диэлектриков. Действительно, переписав граничные условия в виде

$$D_1 \cos \alpha_1 = D_2 \cos \alpha_2, \quad E_1 \sin \alpha_1 = E_2 \sin \alpha_2,$$

где α_1 и α_2 — углы, образуемые силовой линией с нормалью к границе раздела (рис. 2.3), и замечая, что $D_1 = \epsilon_0 \epsilon_1 E_1$ и $D_2 = \epsilon_0 \epsilon_2 E_2$ (ϵ_1 и ϵ_2 — диэлектрические проницаемости обоих диэлектриков), получим

$$\frac{\operatorname{tg} \alpha_1}{\operatorname{tg} \alpha_2} = \frac{\epsilon_1}{\epsilon_2}.$$

Таким образом, из граничных условий $D_{1n} = D_{2n}$ и $E_{1t} = E_{2t}$ вытекает необходимость преломления силовых линий, при этом

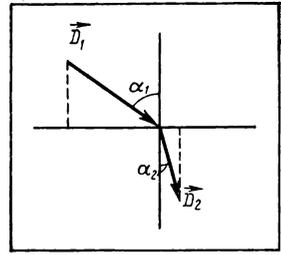
отношение тангенсов углов падения и преломления постоянно и равно отношению диэлектрических проницаемостей; кроме того, силовая линия по обе стороны границы раздела и нормаль к границе лежат в одной плоскости. Если с диэлектриком граничит проводник, то силовые линии на границе перпендикулярны поверхности проводника.

Если проводник заряжен, то нормальная составляющая вектора D_n связана с поверхностной плотностью заряда σ на проводнике соотношением $D_n = \sigma$ (нормаль предполагается внешней по отношению к проводнику). Отсюда следует, что на границе диэлектрика с проводником нормальная составляющая напряженности электрического поля $E_n = \sigma / (\epsilon_0 \epsilon)$. Тангенциальная составляющая E_t напряженности поля на поверхности проводника равна нулю: $E_t = 0$.

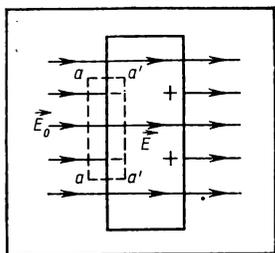
Уменьшение напряженности поля в диэлектрике по сравнению с вакуумом (при тех же сторонних зарядах) связано, как уже говорилось, со смещением зарядов в молекулах диэлектрика под действием поля.

Смещение разноименных зарядов в диэлектрике называют его поляризацией.

Рассмотрим, например, диэлектрическую пластину, помещенную в однородное поле напряженности E_0 , перпендикулярное пластине (рис. 2.4). В этом случае диэлектрик поляризуется так, что разноименные заряды соседних диполей внутри него скомпенсируют друг

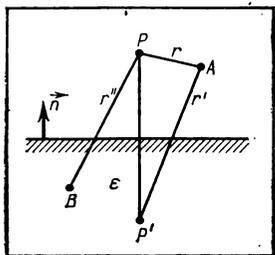


2.3. Преломление силовых линий на границе двух диэлектриков



2.4.

К выводу связи между электрическими индукцией и поляризацией



2.5.

Метод изображений в электростатике диэлектриков

друга и нескомпенсированными окажутся только отрицательные заряды крайних левых и положительные — крайних правых диполей. Диэлектрическая пластина превратится в макроскопический диполь с дипольным моментом, равным $q_r h$ и направленным вдоль E_0 (q_r — нескомпенсированный заряд, вышедший на границу пластины, и h — ее толщина). Ясно, что этот дипольный момент равен сумме дипольных моментов d всех молекул диэлектрика и пропорционален объему V пластины:

$$q_r h \frac{E_0}{E_0} = \Sigma d = PV.$$

Поляризованность (вектор поляризации) P характеризует дипольный момент единичного объема: $P = d/V$.

Нескомпенсированный заряд q_r и, следовательно, поляризацию легко связать с электрическим смещением D . Применим для этого теорему Гаусса для поля в вакууме. Выберем в качестве замкнутой поверхности параллелепипед $aaa'a'$ (рис. 2.4). Получим $(E_0 - E)S = q_r/\epsilon_0$, где E — напряженность поля в диэлектрике, S — площадь грани aa . Замечая, что $q_r h = PV = PhS$, найдем $P = \epsilon_0(E_0 - E)$.

Вспомним теперь, что на границе диэлектрика непрерывна нормальная составляющая вектора D , поэтому в рассматриваемом случае $\epsilon_0 E_0 = D$. Мы приходим к важному соотношению, лежащему в основе всей электродинамики диэлектриков:

$$D = \epsilon_0 E + P.$$

Учитывая, что $D = \epsilon \epsilon_0 E$, имеем $P = \epsilon_0(\epsilon - 1)E = k_a E$. Величину k_a называют *абсолютной диэлектрической восприимчивостью*, а величину $\alpha = k_a/(4\pi\epsilon_0)$ — *поляризуемостью*.

Определим, наконец, потенциал, создаваемый точечным зарядом в том случае, когда диэлектрик заполняет не все пространство, а только полупространство с плоской границей (рис. 2.5). Для решения этой задачи воспользуемся методом электрических изображений, т. е. предположим, что напряженность поля в вакууме (над диэлектриком) будет такой же, как и в отсутствие диэлектрика, но при наличии некоторого добавочного заряда q' , находящегося в точке P' — зеркальном изображении точки P (в которой находится реальный заряд q). Что же касается поля в диэлектрике, то предположим, что оно будет таким же, как поле, создаваемое некоторым фиктивным точечным зарядом q'' , находящимся в неограниченном диэлектрике и расположен-

ном в точке P (в той же точке, где находится реальный заряд). Итак мы предполагаем, что потенциал в вакууме, в произвольной точке A , определяется формулой

$$\varphi = \frac{1}{4\pi\epsilon_0} \left(\frac{q}{r} + \frac{q'}{r'} \right),$$

а потенциал в диэлектрике, в произвольной точке B , — формулой

$$\varphi_\epsilon = \frac{q''}{4\pi\epsilon_0\epsilon r''},$$

где r, r', r'' — расстояния от точек A и B до точек P и P' .

Попытаемся теперь определить константы q' и q'' таким образом, чтобы удовлетворялись граничные условия на поверхности диэлектрика, т. е. чтобы были непрерывны тангенциальная составляющая напряженности поля и нормальная составляющая электрического смещения. Если это удастся сделать, то мы можем быть уверены в том, что наше предположение является правильным.

Непрерывность тангенциальной составляющей напряженности поля эквивалентна непрерывности потенциала. Это значит, что на границе диэлектрика, когда точки A и B совпадают и $r = r' = r''$, должны быть равны потенциалы φ и φ_ϵ , откуда следует, что $q + q' = q''/\epsilon$. Найдем теперь нормальные составляющие вектора электрического смещения. Напряженности электрических полей в вакууме и в диэлектрике составляют соответственно

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0} \left(\frac{q\mathbf{r}}{r^3} + \frac{q'\mathbf{r}'}{r'^3} \right); \quad \mathbf{E}_\epsilon = \frac{1}{4\pi\epsilon_0\epsilon} \frac{q''\mathbf{r}''}{r''^3}$$

(радиусы-векторы \mathbf{r} и \mathbf{r}'' имеют начало в точке P , а радиус-вектор \mathbf{r}' — в точке P'). Поэтому электрическое смещение в вакууме и в диэлектрике определяется формулами

$$\mathbf{D} = \frac{1}{4\pi} \left(\frac{q\mathbf{r}}{r^3} + \frac{q'\mathbf{r}'}{r'^3} \right); \quad \mathbf{D}_\epsilon = \frac{q''\mathbf{r}''}{r''^3}.$$

На границе диэлектрика $r = r' = r''$ и, кроме того, $\mathbf{r}\mathbf{n} = \mathbf{r}''\mathbf{n} = -\mathbf{r}'\mathbf{n}$ (\mathbf{n} — единичный вектор, нормальный к границе). Поэтому равенство $\mathbf{D}\mathbf{n} = \mathbf{D}_\epsilon \mathbf{n}$, которое должно выполняться на границе, дает $q - q' = q''$.

Из этого соотношения и уравнения $q + q' = q''/\epsilon$ найдем искомые величины q' и q'' :

$$q' = -\frac{\epsilon - 1}{\epsilon + 1} q, \quad q'' = \frac{2\epsilon}{\epsilon + 1} q.$$

Заметим, что заряд q притягивается к диэлектрику с силой

$$F = \frac{qq'}{16\pi\epsilon_0 h^2} = -\frac{\epsilon - 1}{\epsilon + 1} \frac{q^2}{16\pi\epsilon_0 h^2}.$$

2.2. Диэлектрик в однородном поле

Одинаковое (в ϵ раз) уменьшение напряженности поля во всех точках диэлектрика имеет место только в диэлектрике безграничном, однородном и изотропном. Если диэлектрик ограничен (но однороден и изотропен), то изменение поля не сводится к масштабному уменьшению; напротив, при этом происходит, как правило, существенное изменение пространственной структуры поля.

Нахождение напряженности поля зарядов в присутствии ограниченного диэлектрика представляет собой сложную задачу математической физики, аналитическое решение которой известно лишь в нескольких простейших случаях. Всякий раз при этом используется известный уже нам принцип, заключающийся в том, что каждая задача электростатики допускает единственное решение и поэтому, если с помощью суперпозиции известных электростатических полей удастся удовлетворить граничным условиям на поверхности диэлектрика, можно быть уверенным в том, что эта суперпозиция представляет собой истинное поле.

Рассмотрим, например, задачу о модификации однородного электростатического поля напряженности \mathbf{E}_0 , в которое вносится диэлектрический шар радиуса a с диэлектрической проницаемостью ϵ . При внесении диэлектрика в однородное поле поле внутри диэлектрика становится, вообще говоря, неоднородным. Но мы сделаем предположение, которое далее будет оправдано, что поле внутри диэлектрического шара, внесенного в однородное поле, остается однородным. Что касается поля вне шара, то его уже нельзя считать однородным. Мы предположим, что влияние шара (в пространстве вне шара) таково, как если бы он представлял собой некоторый диполь, расположенный в центре шара. Иными словами, если обозначить через \mathbf{d} дипольный момент этого эквивалентного диполя, то мы будем предполагать, что напряженность поля вне шара имеет вид

$$\mathbf{E}^{(e)} = \mathbf{E}_0 + \frac{3(\mathbf{dn})\mathbf{n} - \mathbf{d}}{4\pi\epsilon_0 r^3},$$

где \mathbf{r} — радиус-вектор, проведенный из центра шара в точку наблюдения, и \mathbf{n} — единичный вектор в этом направлении. Поле внутри шара будем считать однородным и запишем его напряженность в форме

$$\mathbf{E}^{(i)} = \mathbf{E}_0 - \nu \mathbf{P} / (4\pi\epsilon_0),$$

где $\mathbf{P} = \mathbf{d}/V$ — поляризованность шара и ν — некоторый числовой коэффициент.

Покажем, что величины \mathbf{d} и ν можно выбрать таким образом, чтобы удовлетворялись граничные условия на поверхности шара. Они заключаются в непрерывности тангенциальной составляющей поля $\mathbf{E}^{(e)}\mathbf{t} = \mathbf{E}^{(i)}\mathbf{t}$ (\mathbf{t} — единичный вектор, касательный к поверхности шара) и нормальной составляющей вектора электрического смещения $\mathbf{D}^{(e)}\mathbf{n} = \mathbf{D}^{(i)}\mathbf{n}$. Замечая, что $\mathbf{D}^{(e)} = \epsilon_0\mathbf{E}^{(e)}$, $\mathbf{D}^{(i)} = \epsilon_0\epsilon\mathbf{E}^{(i)}$, перепишем послед-

нее условие в виде $\mathbf{E}^{(e)} \mathbf{n} = \epsilon \mathbf{E}^{(i)} \mathbf{n}$. Но $\mathbf{n} \mathbf{t} = 0$, поэтому при $r = a$ имеем

$$\mathbf{E}^{(e)} \mathbf{t} = E_0 \mathbf{t} - \frac{d\mathbf{t}}{4\pi\epsilon_0 a^3}, \quad \mathbf{E}^{(i)} \mathbf{t} = E_0 \mathbf{t} - \gamma \mathbf{P} \mathbf{t} / (4\pi\epsilon_0),$$

откуда, учитывая, что $\mathbf{d} = \frac{4}{3}\pi a^3 \mathbf{P}$, получим $\gamma = \frac{4}{3}\pi$. Условие непрерывности нормальной составляющей вектора индукции дает

$$E_0 \mathbf{n} + \frac{d\mathbf{n}}{2\pi\epsilon_0 a^3} = \epsilon \left(E_0 \mathbf{n} - \frac{1}{3\epsilon_0} \mathbf{P} \mathbf{n} \right)$$

или $E_0 \mathbf{n} + 2\mathbf{P} \mathbf{n} / (3\epsilon_0) = \epsilon (E_0 \mathbf{n} - \frac{1}{3\epsilon_0} \mathbf{P} \mathbf{n})$, откуда $(2 + \epsilon) \mathbf{P} \mathbf{n} / (3\epsilon_0) = (\epsilon - 1) E_0 \mathbf{n}$. Этому условию можно удовлетворить, выбрав \mathbf{P} равным

$$\mathbf{P} = 3\epsilon_0 \frac{\epsilon - 1}{\epsilon + 2} \mathbf{E}_0.$$

При таком значении \mathbf{P} напряженность поля внутри диэлектрического шара

$$\mathbf{E}^{(i)} = \frac{3}{\epsilon + 2} \mathbf{E}_0.$$

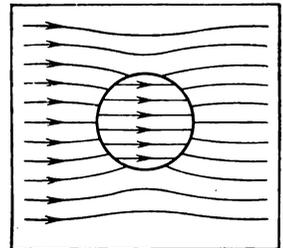
Мы видим, что поле внутри шара, внесенного в однородное поле, действительно остается однородным. Так как $\epsilon > 1$, то внутреннее поле $E^{(i)}$ меньше поля E_0 , бывшего до внесения диэлектрика, но не в ϵ , а в $(\epsilon + 2)/3$ раз (заметим, что всегда $[(\epsilon + 2)/3] < \epsilon$). Поле вне шара неоднородно, но вдали от него приближается к однородному и его напряженность не отличается от напряженности поля E_0 , существовавшей до внесения шара (рис. 2.6).

Заметим, что если в формуле, связывающей $\mathbf{E}^{(i)}$ и \mathbf{E}_0 , заменить ϵ на $1/\epsilon$, то полученная формула определит напряженность в шарообразной полости, сделанной в безграничном диэлектрике с диэлектрической проницаемостью ϵ , в котором в отсутствие полости поле однородно и имеет напряженность E_0 .

Таким образом, если в однородное поле вносится ограниченный диэлектрик, имеющий форму шара, то поле в нем тоже однородно (хотя его напряженность изменяется). Возникает вопрос: при каких еще формах ограниченного диэлектрика соблюдается это свойство? Оказывается, что в общем случае диэлектрик может иметь форму эллипсоида; тогда при внесении его в однородное электростатическое поле напряженностью E_0 в диэлектрике также установится однородное поле напряженностью $E^{(i)}$, причем

$$\mathbf{E}^{(i)} = \mathbf{E}_0 - N \mathbf{P} / \epsilon_0^{-1},$$

где $\mathbf{P} = (\epsilon - 1) \epsilon_0 \mathbf{E}^{(i)}$ и N — числовой множитель (точнее, тензор), зависящий от формы эллипсоида и его расположения по отношению



2.6. Силовые линии поля, в которое помещен шар из диэлектрика

к внешнему полю. Его называют *коэффициентом деполаризации* или *размагничивающим фактором* (последнее название связано с аналогичной ситуацией в случае магнетика, внесенного в магнитное поле, см. гл. 12). Для шара $N = 1/3$. Для бесконечного цилиндра, ось которого направлена вдоль внешнего поля, $N = 1/2$.

2.3. Энергия поля в диэлектрике

Если поместить какой-либо заряд в неограниченный однородный диэлектрик, то создаваемый им потенциал в ϵ раз меньше потенциала, создаваемого тем же зарядом в вакууме. Поэтому емкость конденсатора, если заполнить пространство между его обкладками однородным диэлектриком, возрастет в ϵ раз. В частности, емкость плоского конденсатора, заполненного диэлектриком, $C = \epsilon_0 \epsilon S/d$.

Что касается электрической энергии, запасенной в конденсаторе, то она определяется прежней формулой $W = \frac{1}{2} C (\varphi_1 - \varphi_2)^2 = q^2 / (2C)$. Так как $\varphi_1 - \varphi_2 = Ed$, то $W = \epsilon_0 \epsilon E^2 V / 2$, где V — объем поля в конденсаторе. Это формула показывает, что в 1 м^3 диэлектрика запасена энергия

$$w = \epsilon_0 \epsilon E^2 / 2 = \mathbf{ED} / 2.$$

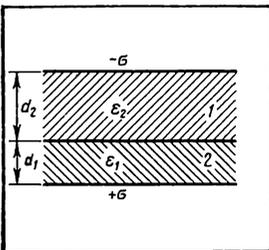
Она представляет собой сумму собственно энергии электрического поля и той части энергии диэлектрика, которая связана с наличием поля. Величину w можно интерпретировать как плотность внутренней энергии, если конденсатор теплоизолирован. Если же диэлектрик находится при постоянной температуре, то w играет роль плотности свободной энергии. При не зависящей от температуры ϵ свободная энергия совпадает с внутренней энергией. Величину w называют *плотностью электрической энергии* в диэлектрике.

Рассмотрим теперь конденсаторы, заполненные неоднородным диэлектриком.

На рис. 2.7 изображен плоский слоистый конденсатор, заполненный двумя диэлектриками с диэлектрическими проницаемостями ϵ_1 и ϵ_2 ; граница между диэлектриками плоская, параллельная обкладкам. При определении емкости такого конденсатора следует исходить из того, что вектор электрической индукции одинаков в обоих диэлектриках:

$$\mathbf{D} = \sigma \mathbf{n},$$

где σ — поверхностная плотность заряда ($\sigma > 0$) и \mathbf{n} — единичный вектор нормали к положительно заряженной обкладке, направленный в сторону диэлектрика. Поэтому напряженности поля в обоих диэлектриках составляют $E_1 = \sigma / (\epsilon_0 \epsilon_1)$, $E_2 = \sigma / (\epsilon_0 \epsilon_2)$ и разности потенциалов, приходящиеся на оба слоя, определяются формулами



2.7. Плоский слоистый конденсатор

$$\varphi_1 - \varphi_0 = \frac{\sigma}{\epsilon_0 \epsilon_1} d_1; \quad \varphi_0 - \varphi_2 = \frac{\sigma}{\epsilon_0 \epsilon_2} d_2,$$

где φ_0 — потенциал границы между диэлектриками и d_1, d_2 — толщины обоих слоев диэлектриков.

Таким образом, общая разность потенциалов на конденсаторе $\varphi_1 - \varphi_2 = \frac{\sigma}{\epsilon_0} \left(\frac{d_1}{\epsilon_1} + \frac{d_2}{\epsilon_2} \right)$, а так как заряд на обкладке $q = \sigma S$, то емкость слоистого конденсатора

$$C = \frac{q}{\varphi_1 - \varphi_2} = \frac{S \epsilon_0}{d_1/\epsilon_1 + d_2/\epsilon_2}.$$

Теперь легко вычислить энергию такого конденсатора:

$$W = \frac{q^2}{2C} = \frac{\sigma^2}{2\epsilon_0} \left(\frac{d_1}{\epsilon_1} + \frac{d_2}{\epsilon_2} \right) S.$$

Это выражение позволяет определить силы, действующие на обкладки конденсатора. В случае вакуумного конденсатора, вычисляя силу, действующую на обкладку с положительным зарядом, мы умножали этот заряд на напряженность поля, создаваемого обкладкой с отрицательным зарядом. Для слоистого конденсатора мы не знаем напряженностей полей, создаваемых каждой из обкладок в отдельности. Но для нахождения сил и не обязательно знать напряженности поля, создаваемые каждой из обкладок. Это можно сделать проще, если рассматривать электрическую энергию конденсатора как его потенциальную энергию. Такое рассмотрение законно, если конденсатор после зарядки отключен от внешних зарядов и предоставлен самому себе. Тогда работа сил, действующих на обкладки при их виртуальном смещении, равна убыли электрической энергии конденсатора. Представим себе, например, что обкладка 1 смещается так, что толщина первого слоя увеличивается на δd_1 (при неизменном d_2); если соответствующее изменение энергии W равно δW , то работа силы F_1 , действующей на обкладку 1, на перемещении δd_1 составит $F_1 \delta d_1 = -\delta W$, откуда

$$F_1 = - \frac{\partial W}{\partial d_1} = - \frac{\sigma^2}{2\epsilon_0 \epsilon_1} S$$

(при нахождении производной нужно считать толщину d_2 постоянной). Отрицательный знак в этом выражении означает, что сила F_1 направлена в сторону уменьшения d_1 , т. е. к обкладке 2.

Аналогичным образом легко показать, что на обкладку 2 действует сила $F_2 = -\partial W/\partial d_2$. Эта сила направлена в сторону уменьшения толщины d_2 , т. е. к обкладке 1. Но хотя обе эти силы и направлены противоположно друг другу, они не равны по модулю, как это имело место в случае вакуумного конденсатора. На первый взгляд кажется, что здесь нарушается закон равенства действия и противодействия. В действительности, однако, никакого нарушения этого закона нет, так как в конденсаторе имеется еще третье тело — диэлектрик. Поэтому следует считать, что сила F_1 действует на обкладку 1 не со сто-

роны обкладки 2, а со стороны примыкающего к обкладке 1 диэлектрика, а сила F_2 действует на обкладку 2 со стороны примыкающего к ней диэлектрика. В свою очередь, на диэлектрик действуют силы со стороны обкладок 1 и 2, равные соответственно $-F_1$ и $-F_2$ и направленные к обкладкам. На диэлектрик действует результирующая сила

$$F_d = -F_2 + F_1 = \frac{D^2}{2\epsilon_0} \left(\frac{1}{\epsilon_2} - \frac{1}{\epsilon_1} \right) S.$$

Она направлена в сторону диэлектрика с меньшей диэлектрической проницаемостью. Если, например, опустить одну обкладку воздушного конденсатора в сосуд с керосином так, чтобы обкладка была параллельна его поверхности, то керосин будет втягиваться в конденсатор, потому что его диэлектрическая проницаемость больше, чем у воздуха.

Таким образом, на границе двух диэлектриков действует сила. Отнесенная к площади граничного слоя, она равна разности плотностей энергии поля по обе стороны границы: $F_d = S(\omega_2 - \omega_1)$, где $\omega_1 = D^2/(2\epsilon_0\epsilon_1)$, $\omega_2 = D^2/(2\epsilon_0\epsilon_2)$.

Силы, действующие на диэлектрик, в электрическом поле, называют *пондеромоторными*. Определим объемную плотность этих сил в простейшем случае газообразного диэлектрика, когда действующее на отдельные заряды молекулы поле практически не отличается от среднего поля в среде. Считая молекулу диполем малых размеров l , можно представить силу, действующую на молекулу, в виде

$$\mathbf{f} = -q\mathbf{E}(\mathbf{r}) + q\mathbf{E}(\mathbf{r} + \mathbf{l}) = q(\nabla\mathbf{E})\mathbf{E} = (\mathbf{d}\nabla)\mathbf{E}$$

(мы пользуемся векторным дифференциальным оператором набла ∇ с компонентами $\partial/\partial x$, $\partial/\partial y$, $\partial/\partial z$). Умножая \mathbf{f} на концентрацию n молекул и вспоминая, что $n\mathbf{d}$ представляет собой поляризованность диэлектрика \mathbf{P} , получим

$$\mathbf{F} = (\mathbf{P}\nabla)\mathbf{E} = \epsilon_0(\epsilon - 1)(\mathbf{E}\nabla)\mathbf{E}.$$

Легко убедиться, что если поле \mathbf{E} направлено вдоль оси x и зависит только от x , то $F = 1/2 \epsilon_0(\epsilon - 1)dE^2/dx$. Оказывается, что и в общем случае, когда поле имеет произвольное направление и зависит от трех координат, выражение для силы \mathbf{F} может быть преобразовано к похожему виду:

$$\mathbf{F} = 1/2 \epsilon_0(\epsilon - 1) \text{grad } E^2.$$

Таким образом, сила, действующая на единичный объем диэлектрика, пропорциональна градиенту квадрата поля. Направлена эта сила в сторону увеличения модуля напряженности поля, так что диэлектрик всегда увлекается в сторону большей напряженности. Этим, в частности, объясняется притяжение заряженными проводниками кусочков бумаги или бузиновых шариков.

Заметим еще, что полученная формула для плотности пондеромоторных сил справедлива не только для газов, но и для жидкостей, если только поляризуемость пропорциональна плотности молекул.

Закон Гаусса $\int \mathbf{D} ds = \Sigma q^{(t)}, \mathbf{D} = \varepsilon_0 \varepsilon \mathbf{E}$

Граничные условия $D_{1n} = D_{2n}, E_{1t} = E_{2t}$

Связь между индукцией и поляризованностью $\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P}$

Основные законы, электростатики $\text{rot } \mathbf{E} = 0, \text{div } \mathbf{D} = \rho$



Глава 3. ПОСТОЯННЫЙ ЭЛЕКТРИЧЕСКИЙ ТОК

3.1. Плотность электрического тока

Вернемся к проводникам. Как мы уже говорили, внутренние заряды в них (в отличие от диэлектриков) могут свободно перемещаться под действием приложенного электрического поля. Движение заряженных частиц в проводниках под действием приложенного электрического поля носит общее название *электрического тока*.

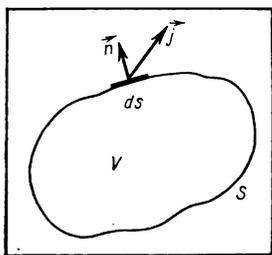
Проводниками являются металлы, полупроводники, электролиты и ионизованные газы (плазма). Подвижными заряженными частицами в металлах являются электроны, ионы же в металлах неподвижны и образуют так называемый *ионный остов*; поэтому ток в металлах представляет собой движение электронов. Аналогично обстоит дело в полупроводниках — у них также подвижны только электроны. Электропроводность электролитов обусловлена ионами. Наконец, в плазме наряду с электронами в токе участвуют и ионы.

Чтобы получить количественную характеристику тока, выделим некоторый небольшой объем проводника (для простоты будем считать его единичным) и запишем для него сумму произведений зарядов подвижных частиц на векторы их скоростей:

$$\mathbf{i} = \sum_{(V=1)} e \mathbf{v},$$

где суммирование распространяется на все заряженные частицы в рассматриваемом единичном объеме проводника. Эта величина — ее называют *плотностью тока* — служит для характеристики тока.

Ясно, что в отсутствие поля плотность тока равна нулю. Действительно, заряженные частицы проводника участвуют в тепловом движении, но в силу хаотичности этого движения среднее значение вектора скорости частицы равно нулю. Это значит, что в отсутствие поля для каждого сорта частиц в любом элементе объема количества частиц со скоростями \mathbf{v} и $-\mathbf{v}$ равны. Поэтому если $\mathbf{E} = 0$, то нулю равна сумма, определяющая плотность тока \mathbf{j} . Если же поле в проводнике отлично



3.1.
К выводу уравнения непрерывности

от нуля, то возникает направленное движение заряженных частиц проводника, при котором плотность тока отлична от нуля. Поэтому под скоростью \mathbf{v} в выражении для \mathbf{j} можно понимать среднюю скорость направленного движения частиц и значение \mathbf{v} можно вынести из-под знака суммы. Оставшаяся сумма равна qn , где n — концентрация частиц. Таким образом,

$$\mathbf{j} = qn\mathbf{v}$$

(в случае носителей тока нескольких сортов требуется еще суммирование по ним).

Так как величина \mathbf{j} определяет количество электричества, протекающего за 1 с через площадку площадью 1 м^2 , ориентированную перпендикулярно вектору \mathbf{j} , то количество электричества dI , протекающего за 1 с через произвольно ориентированную площадку ds , составит $dI = \mathbf{j} \cdot d\mathbf{s}$, где $d\mathbf{s}$ — вектор, определяемый как $d\mathbf{s} = \mathbf{n} ds$ (\mathbf{n} — единичный вектор нормали к площадке). Если задана какая-либо поверхность S , то за 1 с через нее пройдет количество электричества

$$I = \int_S \mathbf{j} \cdot d\mathbf{s}.$$

Эта величина носит название *силы тока*, проходящего через поверхность S . Таким образом, *сила тока через какую-либо поверхность представляет собой поток плотности тока через эту поверхность*.

Сила тока выражается в амперах: $1 \text{ А} = 1 \text{ Кл} \cdot \text{с}^{-1}$.

Плотность электрического тока, вообще говоря, может меняться от точки к точке и, кроме того, зависеть от времени. Можно поэтому говорить о *векторном поле плотности тока*, или сокращенно о *токовом поле*, а также о *токовых линиях*, т. е. линиях, касательная к которым в каждой их точке имеет направление существующего в этой точке вектора \mathbf{j} . Поле это удовлетворяет важному соотношению, которое вытекает из закона сохранения заряда. Чтобы получить его, рассмотрим произвольную замкнутую поверхность S (рис. 3.1) и определим изменение за 1 с заряда Q , находящегося в объеме V , ограниченном поверхностью S . Так как заряд не может ни возникнуть, ни уничтожиться, то изменение заряда Q за 1 с должно быть равно суммарной силе тока, вытекающего из объем V через поверхность S , т. е. $\dot{Q} = -I$, где I — сила тока, вытекающего из объема V . Подставляя сюда

$$Q = \int_V \rho dV, \quad I = \int_S \mathbf{j} \cdot d\mathbf{s},$$

где ρ — объемная плотность заряда ($d\mathbf{s}$ имеет направление внешней к поверхности S нормали), получим

$$\int_S \mathbf{j} \cdot d\mathbf{s} = - \int_V \frac{\partial \rho}{\partial t} dV.$$

Левая часть этого равенства представляет собой поток вектора \mathbf{i} через замкнутую поверхность. С выражением такого типа мы уже встречались при формулировке теоремы Гаусса. Используя полученные там результаты, можно преобразовать поверхностный интеграл в объемный:

$$\int_S \mathbf{j} \cdot d\mathbf{s} = \int_V \operatorname{div} \mathbf{j} dV.$$

Поэтому $\int_V \operatorname{div} \mathbf{j} dV = - \int_V \frac{\partial \rho}{\partial t} dV$, откуда ввиду произвольности объема V следует

$$\operatorname{div} \mathbf{j} = - \frac{\partial \rho}{\partial t}.$$

Это соотношение носит название *уравнения непрерывности* и выражает математически закон сохранения заряда.

3.2. Закон Ома

Вясним теперь, при каких условиях может существовать постоянный ток. Дело в том, что для протекания тока в проводнике должно присутствовать электрическое поле, причем постоянному току соответствуют, согласно уравнению непрерывности, не зависящая от времени плотность заряда и, следовательно, постоянное электрическое поле. Между тем в гл. 1 неоднократно высказывалось утверждение, что электростатическое поле в проводнике всегда равно нулю. Как же при этом может протекать постоянный ток?

И действительно, в чисто электростатическом поле существование постоянного тока невозможно. Это видно, кстати, также из того, что токовые линии постоянного тока замкнуты; поэтому заряд, движущийся вдоль токовой линии, должен непрерывно совершать работу, в окружающем же электростатическом поле при этом не будет происходить никаких изменений, что противоречит закону сохранения энергии.

Парадокс этот разрешается следующим образом. Существование постоянного тока возможно, но для этого необходимо привлечение неэлектростатических сил. Именно: необходимо, чтобы по крайней мере в некоторых местах токовых линий на подвижные заряды действовали неэлектростатические силы. Их называют *сторонними силами*. Для получения их используются так называемые *источники постоянного тока* — гальванические элементы, аккумуляторы, термоэлементы, а также динамомашины постоянного тока. Механизмы действия этих источников мы рассмотрим в следующих главах, пока же введем сторонние силы чисто формально, или, как говорят, феноменологически.

Обозначим \mathbf{E}' суммарную напряженность поля, действующую на единичный заряд в проводнике. Она включает в себя как напряженность \mathbf{E} электростатического поля, так и напряженность $\mathbf{E}^{(e)}$ поля сторонних сил, т. е. $\mathbf{E}' = \mathbf{E} + \mathbf{E}^{(e)}$. Это суммарное поле \mathbf{E}' и опре-

деляет плотность тока \mathbf{j} , а так как при $\mathbf{E}' = 0$ плотность тока, как мы уже говорили, равна нулю, то естественно предполагать, что при достаточно малых напряженностях \mathbf{E}' плотность тока должна быть пропорциональна им. Именно так и оказывается, причем для изотропных проводников существует простая пропорциональность между \mathbf{j} и \mathbf{E}' :

$$\mathbf{j} = \sigma \mathbf{E}'.$$

Для анизотропных же (т. е. монокристаллических) проводников компоненты j_i и E'_i векторов \mathbf{j} и \mathbf{E}' связаны между собой общим линейным соотношением

$$j_i = \sum_k \sigma_{ik} E'_k,$$

где индексы i и k ($i, k \equiv x, y, z$) служат для обозначения трех декартовых осей координат, при этом величины σ_{ik} удовлетворяют условию симметрии $\sigma_{ik} = \sigma_{ki}$. Эти величины, так же как и величину σ , называют *удельной электрической проводимостью* проводника. Они различны для разных проводников и зависят от ряда физических факторов, таких, как температура, наличие примесей и т. д. Удельная электрическая проводимость выражается в СИ в сименсах на метр ($\text{См} \cdot \text{м}^{-1}$, $\text{См} = \text{А} \cdot \text{В}^{-1}$).

Линейное соотношение, связывающее плотность тока и суммарную напряженность поля, называют *закон Ома* (в дифференциальной форме). Хотя этот закон и кажется очень простым, тем не менее он нуждается в некотором разъяснении. Дело в том, что если в проводнике на подвижный заряд действует постоянное поле, то заряд этот, казалось бы, должен непрерывно ускоряться, т. е. он не может обладать постоянной скоростью, а между тем согласно закону Ома мы получаем при заданном поле \mathbf{E}' конечное значение плотности тока, которому по формуле $\mathbf{j} = qn\mathbf{v}$ соответствует конечное значение скорости частицы \mathbf{v} . Каким же образом возникает равномерное, а не равномерно ускоренное движение зарядов проводников? Ответ на этот вопрос таков: *заряды в проводнике подвижны, но их нельзя считать полностью свободными.*

Разъясним это на примере электролита. Пусть в нем под действием поля \mathbf{E}' происходит движение заряженной частицы с зарядом q и массой m . Если бы частица была свободной, то движение ее определялось бы уравнением $m\dot{\mathbf{v}} = q\mathbf{E}'$. Но в действительности электролит обладает вязкостью, благодаря чему при движении частицы возникает трение. Сила трения направлена противоположно скорости и равна $-\eta\mathbf{v}$, где η — коэффициент трения. Таким образом, движение частицы определяется уравнением $m\dot{\mathbf{v}} = q\mathbf{E}' - \eta\mathbf{v}$, решение которого имеет вид

$$\mathbf{v} = \frac{q}{\eta} \mathbf{E}' + \mathbf{a} e^{-(\eta/m)t},$$

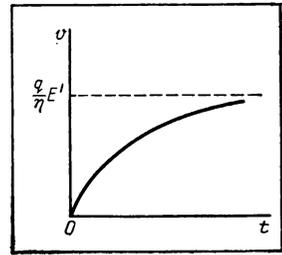
где \mathbf{a} — константа интегрирования. Ее можно найти из начальных условий. Если в начальный момент времени $t = 0$ скорость частицы была равна нулю, то, очевидно, $\mathbf{a} = -\frac{q}{\eta} \mathbf{E}'$ и, следовательно,

$$v = \frac{q}{\eta} E' (1 - e^{-(\eta/m)t}).$$

Мы видим, что с течением времени скорость увеличивается и стремится к предельному значению $(q/\eta)E'$, не зависящему от времени (рис. 3.2). По прошествии времени $\tau = m/\eta$ после включения поля (это время называют *временем релаксации*) скорость будет отличаться от предельного значения в e^{-1} раз, по прошествии же времени $2m/\eta$ различие составляет всего $e^{-2} \sim 10^{-1}$ от предельного значения. Таким образом, скорость быстро «выходит» на свое предельное значение $(q/\eta)E'$, пропорциональное напряженности поля. Подставляя это значение в общее выражение для плотности тока $\mathbf{j} = qn\mathbf{v}$, получим $\mathbf{j} = (q^2n/\eta)E'$. Сравнивая это выражение с формулой $\mathbf{j} = \sigma E'$, найдем удельную электрическую проводимость:

$$\sigma = \frac{q^2n}{\eta} = \frac{q^2n}{m} \tau.$$

Таким образом, мы показали, что учет силы трения при движении заряженных частиц в электролите действительно приводит к закону Ома. Вопрос о том, как возникает конечная проводимость в других физических системах (металлах, газовом разряде), мы рассмотрим в ч. II.



3.2.

К выводу закона Ома

3.3. Линейная электрическая цепь

Перейдем теперь к выяснению связи между силой тока в проводнике и приложенной к нему сторонней силой. Для простоты будем предполагать, что размеры сечения проводника, через которое протекает ток, малы по сравнению с длиной проводника. Такой линейный проводник можно рассматривать, очевидно, как токовую нить или токовый шнур. Из закона сохранения заряда следует, что сила тока в каждом сечении проводника одинакова. Так как проводник предполагается тонким, то плотность тока можно считать постоянной вдоль сечения. Поэтому постоянство силы тока приводит к соотношению $j = I/S$, где I — сила тока и S — площадь поперечного сечения проводника. Мы видим, что плотность тока обратно пропорциональна площади поперечного сечения проводника. Это значит, что постоянный ток ведет себя как стационарный поток несжимаемой жидкости: скорость такого потока в трубе, так же как и плотность постоянного тока, обратно пропорциональна площади сечения трубы.

Обратимся теперь к закону Ома в дифференциальной форме $\mathbf{j} = \sigma E'$, связывающему вектор плотности тока \mathbf{j} с суммарным полем E' (проводник предполагается изотропным). Так как соотношение это по самому смыслу локально, то σ может считаться функцией коор-

динат; точнее говоря, может считаться разной в разных сечениях проводника.

Записав закон Ома в виде $\mathbf{E}' = \rho \mathbf{j}$, где $\rho = 1/\sigma$ — так называемое *удельное электрическое сопротивление*, проинтегрируем это соотношение вдоль всего замкнутого проводника (проводник должен быть, очевидно, замкнутым, как и токовая линия). Элемент длины проводника, ориентированный так же, как и вектор плотности тока, обозначим через $d\mathbf{l}$, тогда $\oint \mathbf{E}' d\mathbf{l} = \oint \rho \mathbf{j} d\mathbf{l}$. Так как, по определению, $d\mathbf{l} \parallel \mathbf{j}$, то $\mathbf{j} d\mathbf{l} = j d\mathbf{l}$. Подставляя сюда $j = I/S$ и замечая, что сила тока I одинакова во всех сечениях S проводника, имеем

$$\oint \rho \mathbf{j} d\mathbf{l} = I \oint \frac{\rho d\mathbf{l}}{S},$$

где интегрирование совершается вдоль всего замкнутого проводника, причем удельное сопротивление ρ и площадь сечения S могут меняться вдоль длины проводника. Итак, $\frac{I}{S} \oint \rho d\mathbf{l} = \oint \mathbf{E}' d\mathbf{l}$, или

$$I = \mathcal{E}/R,$$

где $\mathcal{E} = \oint \mathbf{E}' d\mathbf{l}$, $R = \oint \frac{\rho d\mathbf{l}}{S}$.

Величину \mathcal{E} называют *электродвижущей силой* (э. д. с.), действующей в проводнике, а R — *электрическим сопротивлением* проводника. Таким образом, сила тока I равна отношению электродвижущей силы к сопротивлению проводника. Это соотношение известно под названием *з а к о н а О м а* для всего замкнутого проводника, или, как говорят иначе, для *электрической цепи*.

Электродвижущая сила в токовой нити представляет собой циркуляцию суммарного поля $\mathbf{E}' = \mathbf{E} + \mathbf{E}^{(e)}$ вдоль контура нити, а так как входящее сюда электростатическое поле является потенциальным, т. е. $\oint \mathbf{E} d\mathbf{l} = 0$, то $\mathcal{E} = \oint \mathbf{E}^{(e)} d\mathbf{l}$. Таким образом, электродвижущая сила представляет собой в действительности циркуляцию стороннего поля вдоль токовой нити.

Если $\mathbf{E}^{(e)} = 0$, то $\mathcal{E} = 0$. Более того, если бы поле $\mathbf{E}^{(e)}$ имело чисто электростатическое происхождение (или имело неэлектростатическое происхождение, но было бы потенциальным), то циркуляция от него была бы равна нулю и нулю была бы равна э. д. с. При этом согласно закону Ома отсутствовал бы ток. Это еще раз подтверждает, что существование постоянного тока невозможно в чисто электростатическом поле и возможно лишь при наличии непотенциальных полей.

Смысл электродвижущей силы чрезвычайно прост. Так как $e\mathbf{E}'$ представляет собой суммарную силу, действующую на заряд e , то $e\mathbf{E}' d\mathbf{l}$ представляет собой работу этой силы на пути $d\mathbf{l}$. Поэтому э. д. с. можно интерпретировать как работу сторонних сил при перемещении положительного единичного заряда вдоль всей замкнутой токовой нити. Так как в отсутствие э. д. с. постоянный ток не может протекать, то можно сказать, что источники постоянного тока представляют собой источники э. д. с.

Э. д. с. имеет, очевидно, ту же размерность, что и потенциал, и выражается в вольтах (В).

Обратимся теперь к формуле для электрического сопротивления. Если цепь состоит из отдельных проводников с различными удельными сопротивлениями ρ_1, ρ_2, \dots (рис. 3.3, где l_1 и l_2 — длины проводников, S_1 и S_2 — площади их сечений, а заштрихованный кружок обозначает источник э. д. с.), то сопротивление цепи

$$R = R^{(i)} + \frac{\rho_1 l_1}{S_1} + \frac{\rho_2 l_2}{S_2} + \dots,$$

где $R^{(i)}$ — внутреннее сопротивление источника э. д. с. Величины $R_1 = \rho_1 l_1 / S_1, R_2 = \rho_2 l_2 / S_2, \dots$ представляют собой сопротивления отдельных участков цепи, так что сопротивление всей цепи равно сумме сопротивлений отдельных ее участков, включая сопротивление источника э. д. с.:

$$R = R^{(i)} + R_1 + R_2 + \dots$$

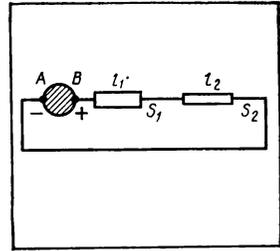
Сопротивление выражают в омах ($1 \text{ Ом} = 1 \text{ А} \cdot 1 \text{ В}^{-1}$).

Закон Ома позволяет определить силу тока и разности потенциалов на различных участках сколь угодно сложной линейной электрической сети, содержащей много соединений (*узлов*) и контуров, в которых могут действовать не одна, а целый ряд э. д. с. Вывод соответствующих правил (их называют *п р а в и л а м и К и р х г о ф а*) несложен: достаточно учесть закон сохранения заряда при каждом разветвлении тока. Сформулируем окончательные выводы:

- (1) Выберем в каждом узле произвольным образом (к узлу или от него) положительное направление тока. Тогда в каждом узле алгебраическая сумма токов равна нулю: $\Sigma I = 0$ (суммирование проводится по всем сходящимся в узле проводникам).
- (2) Выберем в каждом контуре произвольным образом положительное направление обхода. Тогда для каждого контура $\Sigma IR = \Sigma \mathcal{E}$, где суммирование распространяется по всем участкам контура и \mathcal{E} — алгебраическое значение э. д. с., действующей на этом участке.

3.4. Джоулева теплота

Постоянный ток представляет собой, как мы видели, направленное и происходящее с постоянной скоростью движение заряженных частиц в проводнике. Мы видели также, что хотя поле в случае постоянного тока и стационарно, тем не менее токовое состояние несовместимо с условиями электростатики, поэтому протекание постоянного тока возможно лишь при наличии сторонних сил, действующих на заряженные частицы и имеющих неэлектростатическое происхождение. Эти силы сами по себе непрерывно ускоряли бы заряженные частицы, т. е. если бы действовали только они, то направленное движение заряженных частиц не могло бы происходить с по-



3.3. Последовательное соединение сопротивлений

стоянной скоростью. Равномерность движения достигается благодаря эффекту торможения заряженных частиц, которое обуславливается в разных проводниках разными механизмами (см. ч. II).

Как уже указывалось, проще всего исходить из представления о существовании силы трения. При этом движение заряженной частицы описывается уравнением $m\dot{\mathbf{v}} = q\mathbf{E}' + \mathbf{f}_{\text{тр}}$, здесь \mathbf{v} — ускорение частицы, q — ее заряд, $\mathbf{E}' = \mathbf{E} + \mathbf{E}^{(e)}$ — суммарное поле, действующее на частицу, и $\mathbf{f}_{\text{тр}} = -\eta\mathbf{v}$, где $\mathbf{f}_{\text{тр}}$ — сила трения, пропорциональная при малых скоростях скорости частицы, η — коэффициент трения. При равномерном движении $\mathbf{v} = 0$ и сумма всех сил равна нулю, т. е.

$$q\mathbf{E}' = -\mathbf{f}_{\text{тр}} = \eta\mathbf{v}.$$

Умножив $q\mathbf{E}'$ на элемент пути частицы $d\mathbf{s} = \mathbf{v}dt$, мы найдем работу dA силы $q\mathbf{E}'$ над частицей за время dt :

$$dA = q\mathbf{E}'d\mathbf{s} = q\mathbf{E}'\mathbf{v}dt,$$

а так как $q\mathbf{E}' = -\mathbf{f}_{\text{тр}}$, то

$$dA = -\mathbf{f}_{\text{тр}}d\mathbf{s} = \eta v^2 dt.$$

Трение всегда сопровождается выделением теплоты, или, выражаясь более точно, работа силы трения равна количеству теплоты, получаемой окружающей средой. Мы приходим, таким образом, к заключению, что работа, совершаемая силой $q\mathbf{E}'$, полностью переходит в теплоту, выделяемую в проводнике.

Иными словами, прохождение постоянного тока сопровождается выделением в проводнике теплоты. Отнесенная к одной частице, эта теплота равна $\eta v^2 dt$. Умножив это выражение на плотность частиц n , найдем теплоту dQ_1 , выделяемую в единичном объеме проводника за время dt ; она составляет $dQ_1 = n\eta v^2 dt$. Так как плотность тока $j = qnv$, а коэффициент электропроводности σ связан с коэффициентом трения η , согласно результатам § 3.2, соотношением $\sigma = q^2 n / \eta$, то

$$dQ_1 = \frac{1}{\sigma} j^2 dt = \rho j^2 dt,$$

где $\rho = 1/\sigma$ — удельное сопротивление проводника.

Таким образом,

теплота, выделяемая в 1 м³ проводника за 1 с (джоулева теплота), равна произведению квадрата плотности тока на удельное сопротивление проводника:

$$Q_1 = \rho j^2.$$

Во всем проводнике за время t выделится теплота

$$Q = \int_V \rho j^2 t dV,$$

где интегрирование совершается по всему объему V проводника.

Рассмотрим подробнее линейный проводник. В этом случае $dV = S dl$, где dl — элемент длины проводника и S — площадь соответ-

ствующего сечения проводника. Замечая, что сила полного тока $I = jS$, перепишем выражение для Q в виде

$$Q = \oint \rho \frac{I^2}{S^2} S dl.$$

Но сила тока I постоянна вдоль всего проводника и ее можно вынести за знак интеграла, величина же $R = \oint (\rho/S) dl$ представляет собой сопротивление проводника. Итак,

$$Q = RI^2 t,$$

т. е. джоулева теплота равна произведению квадрата силы тока на сопротивление и на время t . За 1 с выделится теплота

$$W = RI^2.$$

Эта величина выражается в ваттах: $1 \text{ Вт} = 1 \text{ А}^2 \cdot \text{Ом}$.

Проходя по проводнику с сопротивлением R , ток создает на нем разность потенциалов $\varphi_1 - \varphi_2 = RI$. Поэтому джоулеву теплоту можно выразить также в виде

$$Q = I(\varphi_1 - \varphi_2)t = \frac{(\varphi_1 - \varphi_2)^2}{R} t.$$

Если в замкнутой цепи действует электродвижущая сила \mathcal{E} , то, как мы знаем, сила тока I связана с \mathcal{E} законом Ома: $I = \mathcal{E}/(R^{(e)} + R^{(i)})$, где $R^{(e)}$ — сопротивление проводника, подключенного к клеммам источника э. д. с., и $R^{(i)}$ — внутреннее сопротивление источника. Используя эту формулу, можно выразить через \mathcal{E} количества теплоты, выделяемые за 1 с во внешнем проводнике и внутри источника э. д. с.:

$$Q^{(e)} = I^2 R^{(e)} = \mathcal{E}^2 R^{(e)} / (R^{(e)} + R^{(i)})^2,$$

$$Q^{(i)} = I^2 R^{(i)} = \mathcal{E}^2 R^{(i)} / (R^{(e)} + R^{(i)})^2.$$

Суммарное количество теплоты

$$Q = Q^{(e)} + Q^{(i)} = I^2 (R^{(e)} + R^{(i)}) = \frac{\mathcal{E}^2}{R^{(e)} + R^{(i)}} = \mathcal{E}I.$$

Такую энергию должен каждую секунду доставлять источник э. д. с. Она равна произведению э. д. с. на силу тока.

Это энергетическое соотношение позволяет несколько иначе понять необходимость сторонней силы для протекания постоянного тока. Так как прохождение тока связано с непрерывным выделением теплоты, то эта теплота должна все время поставляться каким-то внешним источником. Эту функцию и выполняют источники тока. Э. д. с. представляет собой энергию, доставляемую за 1 с источником тока при силе тока 1 А.

Вектор плотности тока $\mathbf{j} = \sum e_i n_i \mathbf{v}_i$

Уравнение непрерывности $\operatorname{div} \mathbf{j} = - \frac{\partial \rho}{\partial t}$

Закон Ома $\mathbf{j} = \sigma \mathbf{E}$

Плотность джоулевой теплоты (отнесенная ко времени) $q = \rho j^2$



Глава 4. МАГНИТНОЕ ПОЛЕ В ВАКУУМЕ

4.1. Сила Лоренца

Электрическое поле и электрический заряд связаны между собой двояким образом: с одной стороны, электрическое поле порождается электрическим зарядом, а с другой — оно воздействует на электрический заряд. Но не только электрическое поле связано таким образом с электрическим зарядом. Существует еще один вид поля — *магнитное поле*, которое, так же как и электрическое поле, создается электрическим зарядом и воздействует на него.

Между электрическим и магнитным полями имеется, однако, важное различие:

если электрическое поле создается электрическим зарядом всегда, независимо от того, покоится или движется заряд, то магнитное поле создается только движущимся зарядом. Если электрическое поле воздействует на электрический заряд всегда, независимо от того, покоится или движется заряд, то магнитное поле действует только на движущийся заряд.

С другой стороны, абсолютного покоя и абсолютного движения не существует — движение всегда относительно. Поэтому если какой-либо заряд покоится в некоторой системе отсчета, то в ней он создает только электрическое поле: стоит, однако, перейти в другую систему отсчета, по отношению к которой заряд движется, и в ней этот заряд будет создавать и электрическое, и магнитное поле. Аналогичным образом, если заряд движется в магнитном поле, то на него со стороны этого поля действует некоторая сила. Стоит, однако, перейти в систему отсчета, в которой заряд покоится, как эта сила исчезнет, но заряд не станет свободным — на него начнет действовать электрическое поле, которое существует в этой системе отсчета. Это электрическое поле зависит от исходного магнитного поля и от относительной скорости обеих систем отсчета. Можно сделать вывод, что

электрическое и магнитное поля образуют единое целое — электромагнитное поле с двумя составными частями, или компонентами: полем электрическим и полем магнитным.

Оба эти поля являются векторными, и, так же как и электрическое поле, магнитное поле может меняться от точки к точке и зависеть от времени.

Магнитное поле мы должны характеризовать некоторой векторной величиной, аналогичной напряженности электрического поля \mathbf{E} . Эту величину обозначают \mathbf{B} , но называют ее по традиции не напряженностью магнитного поля, а *магнитной индукцией*. Понятие *напряженности магнитного поля* \mathbf{H} также существует, и мы далее его разъясним.

Линии, касательные к которым в каждой точке направлены вдоль \mathbf{B} , называют *магнитными силовыми линиями*.

Опыт показывает, что сила \mathbf{f} , действующая на заряд q , движущийся в магнитном поле со скоростью \mathbf{v} , пропорциональна q , направлена перпендикулярно \mathbf{v} и, кроме того, пропорциональна v . Отсюда однозначно следует, что выражение для силы должно иметь структуру векторного произведения скорости частицы \mathbf{v} и магнитной индукции \mathbf{B} . Мы положим

$$\mathbf{f} = q[\mathbf{vB}].$$

Если на заряд q действуют одновременно и магнитное \mathbf{B} и электрическое \mathbf{E} поля, то суммарная сила

$$\mathbf{f} = q(\mathbf{E} + [\mathbf{vB}]).$$

Эту силу называют *силой Лоренца*.

Из формулы для \mathbf{f} следует, что, хотя и электрическое и магнитное поля являются векторными полями, векторная природа их не одинакова. Чтобы разъяснить это важное обстоятельство, следует вспомнить определение векторного произведения, согласно которому векторное произведение ортов \mathbf{i} и \mathbf{j} (вдоль осей x и y) всегда равно орту \mathbf{k} (вдоль оси z) независимо от того, является ли система координат правовинтовой или левовинтовой. Поэтому векторные произведения двух обычных векторов в правовинтовой (R) и в левовинтовой (L) системах имеют прямо противоположные направления. Ясно, что направления таких векторов, как смещение, скорость, ускорение, сила, а также напряженность электрического поля не зависят от того, пользуемся ли мы R - или L -системой. Это значит, что если произвести пространственное отражение, при котором направления всех трех осей координат изменятся на обратные, то проекции векторов смещения, скорости, ускорения, силы, напряженности электрического поля \mathbf{E} на новые оси также изменят свой знак. Такие векторы называют *полярными*.

Но пространственное отражение представляет собой переход от R - к L -системе, при котором проекции на оси координат векторного произведения обычных, т. е. полярных, векторов не меняют, а сохраняют свой знак. Отсюда следует, что если бы магнитная индукция \mathbf{B} имела такую же векторную природу, как и напряженность электрического поля \mathbf{E} , т. е. вектор \mathbf{B} был бы полярным, то взаимная ориентация двух составляющих силы \mathbf{f} , электрической силы $q\mathbf{E}$ и магнитной силы $q[\mathbf{vB}]$, менялась бы при переходе от R - к L -системе. Такое изменение

ориентации физически абсурдно. Это значит, что вектор \mathbf{B} не может быть полярным. Иными словами, при пространственном отражении проекции вектора \mathbf{B} в отличие от вектора \mathbf{E} не должны меняться. Такие векторы называют *аксиальными*. Итак,

напряженность электрического поля — полярный, а магнитная индукция — аксиальный векторы.

Векторное произведение двух полярных или двух аксиальных векторов всегда дает аксиальный вектор; векторное же произведение полярного и аксиального векторов представляет собой полярный вектор. Если какая-либо величина представляет собой полярный вектор (например, сила \mathbf{f}), то ее нельзя представить в виде суммы полярного и аксиального векторов; все слагаемые такой величины должны быть полярными векторами.

Выясним, теперь, как происходит движение заряженной нерелятивистской частицы (т. е. частицы, скорость которой мала по сравнению со скоростью света в вакууме $c \simeq 3 \cdot 10^8 \text{ м} \cdot \text{с}^{-1}$) в заданном магнитном поле. Уравнение движения частицы имеет вид

$$m \frac{d\mathbf{v}}{dt} = q[\mathbf{v}\mathbf{B}],$$

где m — масса частицы. Из этого уравнения прежде всего следует, что *магнитное поле не может изменить кинетической энергии частицы*. Действительно, умножим скалярно обе части уравнения на вектор скорости \mathbf{v} . Справа мы получим, очевидно, нуль; поэтому $m\mathbf{v}(d\mathbf{v}/dt) = 0$, откуда следует, что кинетическая энергия частицы сохраняется: $mv^2/2 = \text{const}$.

Рассмотрим подробнее движение в постоянном и однородном магнитном поле. Очевидно, проекция скорости частицы v_{\parallel} вдоль \mathbf{B} не изменяется, а так как не изменяется квадрат ее скорости, то сохраняется и квадрат поперечной составляющей скорости частицы v_{\perp}^2 (относительно \mathbf{B}): $v_{\parallel} = \text{const}$, $v_{\perp}^2 = \text{const}$.

Проекция траектории частицы на плоскость, перпендикулярную \mathbf{B} , является окружностью. Действительно, проекция ускорения частицы на эту плоскость равна v_{\perp}^2/r , где r — радиус кривизны проекции траектории. Поэтому $m(v_{\perp}^2/r) = qv_{\perp}B$, откуда

$$r = mv_{\perp}/(qB) = \text{const}.$$

Величину r называют *ларморовым радиусом*. Вдоль поля движение происходит с постоянной скоростью v_{\parallel} ; поэтому траекторией частицы при $v_{\parallel} \neq 0$ является винтовая линия. При $v_{\parallel} = 0$ траекторией будет окружность с ларморовым радиусом.

Ларморов радиус пропорционален поперечному импульсу частицы mv_{\perp} и обратно пропорционален индукции B . Если частица движется в известном поле \mathbf{B} и мы можем измерить радиус кривизны ее траектории, то по приведенной формуле можно найти поперечный импульс частицы. Напротив, если известны mv_{\perp} и r , то можно определить значение индукции \mathbf{B} .

Разделив $2\pi r$ на v_{\perp} , получим период обращения частицы

$$T_c = \frac{2\pi r}{v_{\perp}} = 2\pi \frac{m}{qB}.$$

Эта величина не зависит ни от скорости, ни от радиуса окружности, описываемого заряженной частицей. (Этот вывод справедлив только в том случае, если скорость частицы значительно меньше c .) Легко найти также угловую скорость обращения частицы — так называемую *циклотронную частоту*:

$$\omega_c = 2\pi/T_c = qB/m.$$

Умножив T_c на v_{\parallel} , найдем шаг винтовой линии, описываемой частицей:

$$h = T_c v_{\parallel} = 2\pi \frac{mv_{\parallel}}{qB}.$$

Так как на движущийся в магнитном поле электрический заряд со стороны поля действует сила (а электрический ток — это совокупность движущихся зарядов), то на проводник с током, находящийся в магнитном поле, со стороны магнитного поля также действуют силы. Найдем силу, действующую в поле \mathbf{B} на элемент объема dV проводника. Так как на один заряд q , движущийся со скоростью \mathbf{v} в поле \mathbf{B} , действует сила $q[\mathbf{v}\mathbf{B}]$, то для 1 м^3 справедливо соотношение $\mathbf{f} = qn[\mathbf{v}\mathbf{B}]$, где n — плотность подвижных заряженных частиц. Но $qn\mathbf{v}$ представляет собой вектор плотности тока \mathbf{j} , поэтому $\mathbf{f} = [\mathbf{j}\mathbf{B}]$. Умножив \mathbf{f} на dV , найдем силу $d\mathbf{F}$, действующую на элемент объема dV проводника:

$$d\mathbf{F} = [\mathbf{j}\mathbf{B}] dV.$$

Для линейного проводника можно отсюда определить силу $d\mathbf{F}$, действующую на элемент длины dl . Записав dV в виде Sdl , где S — площадь сечения проводника, и учитывая, что $\mathbf{j}S$ представляет собой силу тока I , получим

$$d\mathbf{F} = I [d\mathbf{l}\mathbf{B}],$$

где $d\mathbf{l}$ — ориентированный вдоль направления тока элемент проводника. Мы видим, что сила $d\mathbf{F}$ направлена перпендикулярно $d\mathbf{l}$ и \mathbf{B} так, что если вращать винт с правой нарезкой от $d\mathbf{l}$ к \mathbf{B} , то поступательное движение винта будет происходить вдоль $d\mathbf{F}$. Сила $d\mathbf{F}$ пропорциональна синусу угла между $d\mathbf{l}$ и \mathbf{B} , поэтому она обращается в нуль при $d\mathbf{l} \parallel \mathbf{B}$ и максимальна при $d\mathbf{l} \perp \mathbf{B}$.

4.2. Законы Ампера и Био — Савара

Рассмотрев действие заданного магнитного поля на заряд (движущийся!), мы должны перейти к рассмотрению второй стороны связи заряд—магнитное поле: созданию магнитного поля движущимся зарядом, а следовательно, и током. Так как ток испытывает действие магнитного поля, то можно сделать вывод, что *между двумя токами должны действовать силы магнитного происхож-*

дення. Естественно взять это явление за исходный пункт изучения создания магнитного поля. Можно сказать, что здесь имеется аналогия с электростатикой, в которой введение и изучение электрического поля базировалось на изучении взаимодействия между двумя зарядами. Но там имелся закон Кулона, и поэтому нам нужен аналогичный закон для взаимодействия токов. Такой закон был открыт Ампером, установившим, что

между двумя очень длинными параллельными проводами, по которым протекают токи I и I' , возникает сила взаимодействия, пропорциональная произведению сил токов I и I' и обратно пропорциональная расстоянию между проводами. Параллельные токи притягиваются, а антипараллельные — отталкиваются. Сила взаимодействия пропорциональна также длине проводов.

Таким образом, силу, отнесенную к длине проводников, можно записать в виде

$$\frac{dF_2}{dl} = 2k_2 \frac{II'}{d},$$

где k_2 — некоторая константа, аналогичная константе k_1 в законе Кулона, зависящая от выбора системы единиц (множитель 2 введен для удобства).

Из сравнения размерностей заряда и силы тока вытекает, что отношение k_1/k_2 имеет размерность квадрата скорости. Это обстоятельство имеет глубокий физический смысл — в теории электромагнетизма должна содержаться величина с размерностью скорости. Такой величиной является скорость света в вакууме $c = 2,9979 \cdot 10^8$ м·с⁻¹. И действительно, сравнение механических сил F_1 и F_2 при известных зарядах и токах показало, что

$$k_1/k_2 = c^2.$$

Это соотношение показывает, что при выборе системы единиц величины k_1 и k_2 не могут выбираться произвольно, а всегда должна учитываться связь между ними. Если выбрать $k = 1$, то для k_2 мы получим значение $k_2 = c^{-2}$. Такой выбор соответствует гауссовой системе единиц. В ней, следовательно,

$$\frac{dF_2}{dl} = \frac{2II'}{c^2 d}.$$

Выбрав $k_1 = (4\pi\epsilon_0)^{-1}$, получим

$$k_2 = \frac{k_1}{c^2} = (4\pi\epsilon_0 c^2)^{-1} \equiv \frac{\mu_0}{4\pi},$$

где μ_0 — новая константа, аналогичная ϵ_0 и связанная с ней соотношением $\epsilon_0\mu_0 = c^{-2}$. Такой выбор k_2 делается в СИ; в ней

$$\frac{dF_2}{dl} = \frac{\mu_0}{2\pi} \frac{II'}{d}$$

и $\mu_0 = 4\pi \cdot 10^{-7}$ Гн·м⁻¹ (Гн — единица индуктивности генри).

Теперь мы можем перейти к введению магнитной индукции. Напряженность электрического поля мы вводили на основании закона

Кулона как силу, отнесенную к единичному заряду. Магнитную индукцию можно определить как величину, пропорциональную dF_2/dl . Иными словами, индукцию, порождаемую током I на расстоянии d , можно определить как

$$B = 2k_2\alpha I/d,$$

где α — новая константа. Полагая $\alpha = 1$, получим определение B в СИ:

$$B = \frac{\mu_0}{2\pi} \frac{I}{d}.$$

В гауссовой системе полагают $\alpha = c$ и получают следующее определение B :

$$B = \frac{2}{c} \frac{I}{d}.$$

В этой системе размерности B и E совпадают.

Что касается направления индукции, то вектор \mathbf{B} лежит в плоскости P , перпендикулярной проводнику с током, и силовые линии имеют вид концентрических окружностей, центры которых расположены на проводнике (рис. 4.1). Так как $B \sim d^{-1}$, то произведение длины силовой линии $2\pi d$ на значение магнитной индукции B есть константа, определяющаяся только силой тока I , т. е. $2\pi dB = \mu_0 I$. Направление магнитной индукции связано с направлением тока так же, как связаны между собой направления вращения и поступательного движения винта с правой нарезкой.

Написанному соотношению можно придать другую форму:

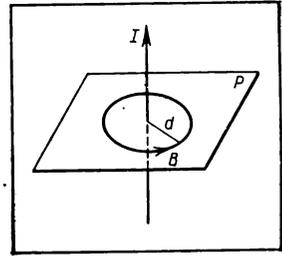
$$\oint_L \mathbf{B} d\mathbf{l} = \mu_0 I,$$

где интегрирование совершается вдоль контура, совпадающего с силовой линией. Оказывается, что контур не обязательно должен совпадать с силовой линией: он может быть выбран произвольно, лишь бы он охватывал ток I . Более того, контур L могут пронизывать много проводников, по которым текут токи I_1, I_2, \dots , и тогда $\oint_L \mathbf{B} d\mathbf{l} = \mu_0 \sum_n (\pm I_n)$, где суммирование распространяется на все токи, причем токи, направления которых связаны с направлением обхода контура правилом правого винта, берутся со знаком плюс, а токи, текущие в обратном направлении, — со знаком минус. Мы можем, таким образом, сказать, что

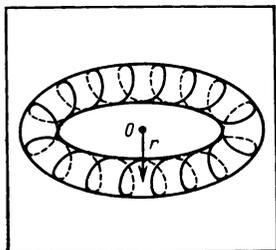
циркуляция магнитной индукции вдоль произвольного контура равна умноженной на μ_0 алгебраической сумме сил токов, пронизывающих этот контур.

Это соотношение носит название **закона полного тока**.

Проиллюстрируем закон полного тока на примере *соленоида*, т. е. согнутого в виде спирали проводника, по которому протекает



4.1. Магнитное поле прямого постоянного тока



4.2.
Соленоид в форме тора

электрический ток. На рис. 4.2 изображен круговой соленоид, в котором витки спирали равномерно намотаны на тор. При тесной намотке поле соленоида сосредоточено в основном внутри соленоида. Силовыми линиями являются, очевидно, окружности с центром в центре тора O . Циркуляция вдоль окружности радиуса r равна $2\pi rB$. Согласно закону полного тока, ее следует приравнять μ_0IN , где N — полное число витков соленоида, откуда $B = \mu_0IN/(2\pi r)$. Таким образом, магнитная индукция в соленоиде обратно пропорциональна радиусу кривизны силовой

линии. В случае достаточно тонкого соленоида значения r меняют в небольших пределах и индукция B практически однородна. Ее можно, очевидно, представить в виде

$$B = \mu_0In,$$

где n — число витков, приходящееся на единицу длины соленоида. В таком виде формула пригодна и для длинного прямого соленоида.

Итак, мы нашли магнитное поле в двух простейших случаях — прямого провода и соленоида. Непосредственное применение при этом закона полного тока было возможно благодаря тому, что нам была известна симметрия магнитных силовых линий. Однако в общем случае произвольных токов непосредственное применение закона полного тока становится невозможным. Поэтому необходимо придать этому закону такую форму, которая не была бы связана с конкретным видом контура. Для этого следует воспользоваться ф о р м у л о й С т о к с а

$$\oint_L \mathbf{B} d\mathbf{l} = \int_S \text{rot } \mathbf{B} d\mathbf{s},$$

причем S — поверхность, опирающаяся на контур L .

Обратимся теперь к закону полного тока и применим его к бесконечно малому контуру, ограничивающему бесконечно малую площадь Δs в сечении проводника. Ток, проходящий через эту площадь, можно записать в виде $\Delta s \mathbf{j}_t$, где \mathbf{j}_t — вектор плотности тока. Так как магнитное поле создается всеми движущимися зарядами, то под \mathbf{j}_t мы должны понимать общую плотность потока движущихся зарядов, складывающуюся из плотности тока, обусловленного свободными зарядами, т. е. плотности тока проводимости \mathbf{j} , и плотности тока, обусловленного связанными зарядами \mathbf{j}_b (ее называют плотностью молекулярного тока): $\mathbf{j}_t = \mathbf{j} + \mathbf{j}_b$. Плотность тока проводимости определяется известным уже нам законом Ома, определение же плотности тока связанных зарядов более сложно, и мы вернемся к нему в гл. 6 и 12. Там мы, в частности, убедимся, что для постоянных полей полный ток связанных зарядов через произвольное сечение тела равен нулю.

Итак, вводя пока формально \mathbf{j}_b , мы можем утверждать, что согласно закону полного тока $\text{rot } \mathbf{B} \Delta s = \mu_0 \mathbf{j}_t \Delta s$, а так как вектор Δs произволен, то

$$\text{rot } \mathbf{B} = \mu_0 \mathbf{j}_t.$$

Это уравнение, связывающее ротор магнитной индукции с полной плотностью тока \mathbf{j}_t , представляет собой *дифференциальную форму закона полного тока в случае постоянных полей*.

Однако закона полного тока даже в дифференциальной форме все еще недостаточно для однозначного определения магнитного поля. Дело в том, что для определения любого векторного поля нужно знать две его характеристики — дивергенцию и ротор. Поэтому нам нужно еще знать дивергенцию магнитной индукции, или, что то же самое, поток магнитной индукции через произвольную замкнутую поверхность.

Поток электрического поля через произвольную замкнутую поверхность определяется находящимся внутри нее суммарным электрическим зарядом. Поэтому естественно, что поток магнитной индукции через замкнутую поверхность должен определяться содержащимся внутри нее суммарным «магнитным зарядом». Но магнитных зарядов в природе не существует и, следовательно, поток магнитной индукции \mathbf{B} через произвольную замкнутую поверхность равен нулю:

$$\int_S \mathbf{B} ds = 0,$$

а это означает, что дивергенция магнитной индукции равна нулю: $\text{div } \mathbf{B} = 0$. Итак, мы имеем

два уравнения:

$$\text{rot } \mathbf{B} = \mu_0 \mathbf{j}_t, \quad \text{div } \mathbf{B} = 0;$$

они полностью определяют магнитное поле, порождаемое любыми постоянными токами.

Подобно тому, как электростатическое поле $\mathbf{E} = -\text{grad } \varphi$ можно представить в виде градиента некоторой скалярной функции — электростатического потенциала φ , так и магнитную индукцию \mathbf{B} можно представить в виде ротора некоторой векторной функции \mathbf{A} , которую называют *векторным потенциалом*:

$$\mathbf{B} = \text{rot } \mathbf{A}.$$

Однако при заданном \mathbf{B} функция \mathbf{A} определяется этим соотношением неоднозначно. Действительно, имеет место тождество $\text{rot grad } \chi = 0$, где χ — произвольная скалярная функция координат. Поэтому ясно, что векторные потенциалы \mathbf{A} и $\mathbf{A}' = \mathbf{A} + \text{grad } \chi$ будут приводить к одной и той же магнитной индукции. Этим обстоятельством можно воспользоваться для того, чтобы наложить на функцию \mathbf{A} какое-либо ограничение. Удобно наложить на \mathbf{A} условие $\text{div } \mathbf{A} = 0$, тогда, поскольку $\text{rot } \mathbf{B} = \mu_0 \mathbf{j}_t$ и $\text{rot rot } \mathbf{A} = \text{grad div } \mathbf{A} - \Delta \mathbf{A} = \text{rot } \mathbf{B}$,

$$\Delta \mathbf{A} = \mu_0 \mathbf{j}_t.$$

Это уравнение для векторного потенциала \mathbf{A} аналогично уравнению Пуассона $\Delta\varphi = \rho/\epsilon_0$ для электростатического потенциала φ .

Как мы знаем, заряд $dq = \rho dV$ порождает потенциал $d\varphi = dq/(4\pi\epsilon_0 r)$, где r — расстояние между элементом объема dV и точкой, в которой ищется потенциал. Поэтому потенциал φ при заданном распределении плотности заряда ρ определяется формулой

$$\varphi = \frac{1}{4\pi\epsilon_0} \int \frac{\rho dV}{r}.$$

Так как уравнение для \mathbf{A} аналогично уравнению для φ с той лишь разницей, что вместо ρ/ϵ_0 в уравнение для \mathbf{A} входит $\mu_0 \mathbf{j}_t$, то можно сразу написать решение уравнения для векторного потенциала:

$$\mathbf{A} = \frac{\mu_0}{4\pi} \int \frac{\mathbf{j}_t dV}{r},$$

где r — расстояние между элементом объема dV (с плотностью тока \mathbf{j}_t) и точкой, в которой ищется векторный потенциал \mathbf{A} .

Воспользуемся этой формулой для определения магнитной индукции, создаваемой постоянным током I , протекающим по некоторому замкнутому линейному (т. е. очень тонкому) проводнику. Интегрирование при этом должно производиться по объему, занимаемому проводником, и поэтому элемент объема в формуле для \mathbf{A} можно представить в виде $dV = ds dl$, где ds — элемент площади сечения и dl — элемент длины проводника. Выполним сперва интегрирование по площади сечения проводника S . Так как расстояние r от точки наблюдения до всех точек сечения S практически одинаково, то мы придем к выражению $r^{-1} \int \mathbf{j}_t ds$. Но, как уже указывалось, $\int \mathbf{j}_t ds = 0$, поэтому мы получим $r^{-1} I \boldsymbol{\tau}$, где $\boldsymbol{\tau}$ — единичный вектор вдоль нормали к S , т. е. вдоль dl . Остается выполнить интегрирование по dl :

$$\mathbf{A} = \frac{\mu_0 I}{4\pi} \oint \frac{dl}{r};$$

оно проводится здесь вдоль токового контура в направлении протекающего тока (r — расстояние от dl до точки, в которой ищется \mathbf{A}). Тогда

$$\mathbf{B} = \text{rot } \mathbf{A} = \frac{\mu_0 I}{4\pi} \text{rot} \oint \frac{dl}{r}.$$

Операции дифференцирования, содержащиеся в rot , должны производиться по координатам точки, в которой ищется \mathbf{B} . Поэтому rot можно внести под знак интеграла и считать dl постоянным вектором. Но в случае постоянного вектора \mathbf{a} имеет место формула

$$\text{rot} \frac{\mathbf{a}}{r} = \frac{[\mathbf{a}\mathbf{r}]}{r^3}.$$

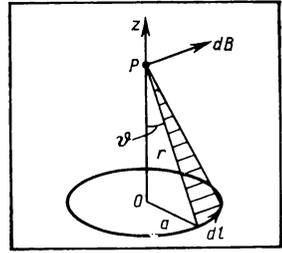
Поэтому окончательно

$$\mathbf{B} = \frac{\mu_0 I}{4\pi} \oint \frac{[dl\mathbf{r}]}{r^3},$$

где интегрирование совершается вдоль токового контура. Эта формула выражает закон Био — Савара.

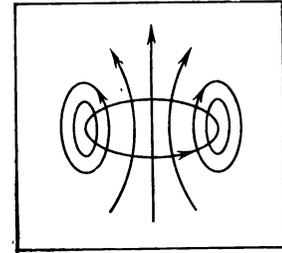
4.3. Магнитный момент тока

Применим закон Био — Савара для определения магнитного поля, создаваемого круговым током (рис. 4.3). Из соображений симметрии ясно, что во всех точках, лежащих на линии Oz , перпендикулярной плоскости кругового тока и проходящей через центр круга O , магнитное поле направлено вдоль этой линии (причем вверх, если ток течет против часовой стрелки). Это значит, что эта линия является силовой линией. Она начинается и кончается на бесконечности. Другие силовые линии изображены на рис. 4.4. Все они являются замкнутыми линиями, охватывающими токовое кольцо.



4.3. К определению магнитного поля кругового тока

Определить магнитное поле кругового тока в произвольной точке довольно сложно, и мы ограничимся здесь вычислением \mathbf{B} только на линии Oz (см. рис. 4.3). Согласно закону Био—Савара, магнитная индукция $d\mathbf{B}$, создаваемая элементом проводника $d\mathbf{l}$ в точке P , перпендикулярна плоскости ($d\mathbf{l}; \mathbf{r}$) (на рис. 4.3 она заштрихована), а так как $\mathbf{r} \perp d\mathbf{l}$, то



4.4. Магнитные силовые линии кругового тока

$$d\mathbf{B} = \frac{\mu_0 I}{4\pi} \frac{d\mathbf{l}}{r^2}.$$

Результирующее поле в точке P направлено вдоль OP ; поэтому нужно найти проекцию $d\mathbf{B}$ на ось z и потом сложить все проекции, соответствующие различным элементам $d\mathbf{l}$ токового контура. Проекция dB_z равна, очевидно,

$$dB \sin v = \frac{\mu_0 I}{4\pi} \frac{dl}{r^2} \sin v,$$

а так как угол v одинаков для всех элементов контура $d\mathbf{l}$, то результирующая индукция

$$B = \frac{\mu_0 I}{2} \frac{a}{r^2} \sin v,$$

где a — радиус токового контура. Замечая, что $\sin v = a/r$, получим

$$B = \frac{\mu_0 I}{2} \frac{a^2}{r^3}.$$

В частности, в точке O имеем $B_0 = \mu_0 I / (2a)$.

Формулу для \mathbf{B} можно переписать в виде

$$\mathbf{B} = [\mu_0 / (2\pi r^3)] \mathbf{m},$$

где $\mathbf{m} = I\mathbf{S}$ и \mathbf{S} — вектор площади, обтекаемой током, т. е. вектор, равный по модулю $S = \pi a^2$ и направленный по нормали к площадке (т. е. вдоль оси z). Вектор \mathbf{m} называют *магнитным моментом тока**. Мы видим, что на оси z поле кругового тока определяется его магнитным моментом. Вне оси такое утверждение в общем случае не справедливо. Но, как можно показать, вдали от кругового тока магнитная индукция, хотя и отличается от индукции на оси, может быть тем не менее выражена через \mathbf{m} . Более того, вдали от произвольного токового контура магнитная индукция всегда определяется магнитным моментом тока \mathbf{m} , а именно

$$\mathbf{B} = \mu_0 \frac{3\mathbf{r}(\mathbf{m}\mathbf{r}) - m r^2}{4\pi r^5},$$

где \mathbf{r} — радиус-вектор, проведенный из места нахождения токового контура в точку, где рассматривается поле (предполагается, что $r \gg a$, где a — размеры контура). По модулю

$$B = \frac{\mu_0 m}{4\pi r^3} \sqrt{1 + 3 \cos^2 \Theta},$$

где Θ — угол между \mathbf{m} и \mathbf{r} . Обратим внимание на аналогию между этими формулами и формулами для электрического дипольного момента (см. § 1.3).

Сравним в заключение этого параграфа электростатическое поле с магнитным полем постоянного тока. Электростатическое поле порождается электрическими зарядами — источниками поля, и дивергенция его напряженности определяется плотностью электрических зарядов: $\operatorname{div} \mathbf{E} = \rho/\epsilon_0$. Поле это безвихревое, т. е. ротор его напряженности равен нулю: $\operatorname{rot} \mathbf{E} = 0$. Магнитное поле не имеет источников, и дивергенция магнитной индукции равна нулю: $\operatorname{div} \mathbf{B} = 0$. Это поле является вихревым, и ротор магнитной индукции определяется плотностью электрических токов: $\operatorname{rot} \mathbf{B} = \mu_0 \mathbf{j}$.

Там, где находятся электрические заряды, начинаются или кончаются электрические силовые линии. Но магнитных зарядов не существует. Поэтому магнитные силовые линии не могут иметь ни начала, ни конца. Отсюда, однако, не следует, что магнитные силовые линии должны быть обязательно замкнутыми или начинаться и кончаться на бесконечности. Оба эти случая, конечно, возможны, но существует еще и третья возможность: силовая линия, не имея ни начала, ни конца, плотно заполняет целую поверхность, называемую *магнитной поверхностью*.

4.4. Взаимодействие токов

Так как ток создает вокруг себя магнитное поле, то, как уже отмечалось, два тока электродинамически взаимодействуют друг с другом. Именно: на элемент проводника $d\mathbf{l}_1$ с током I_1 со стороны тока I_2 действует сила

* Эту величину обозначают также \mathbf{p}_m .

$$d\mathbf{F}_1 = I_1 [d\mathbf{l}_1 \mathbf{B}_2],$$

где \mathbf{B}_2 — магнитная индукция, создаваемая током I_2 в том месте, где находится элемент $d\mathbf{l}_1$ тока I_1 . Согласно закону Био—Савара,

$$\mathbf{B}_2 = \frac{\mu_0 I_2}{4\pi} \oint_{L_2} \frac{[d\mathbf{l}_2 \mathbf{r}_{21}]}{r_{21}^3},$$

где \mathbf{r}_{21} — радиус-вектор, проведенный из элемента $d\mathbf{l}_2$ контура с током I_2 в элемент $d\mathbf{l}_1$ контура с током I_1 . Таким образом,

$$d\mathbf{F}_1 = \frac{\mu_0}{4\pi} I_1 I_2 \left[d\mathbf{l}_1, \oint_{L_2} \frac{[d\mathbf{l}_2 \mathbf{r}_{21}]}{r_{21}^3} \right].$$

Эта формула выражает общий закон Ампера для взаимодействия токов. Мы видим, что сила взаимодействия пропорциональна произведению сил токов.

Рассмотрим подробнее тот случай, когда ток I_2 течет по очень длинному прямолинейному проводнику. Тогда

$$\mathbf{B}_2 = \frac{\mu_0}{2\pi} \frac{I_2}{r^2} [\mathbf{n}_2 \mathbf{r}],$$

где \mathbf{n}_2 — единичный вектор, направленный вдоль проводника с током I_2 (в направлении тока), r — расстояние между $d\mathbf{l}_1$ и вторым проводником, причем вектор \mathbf{r} направлен от второго проводника к $d\mathbf{l}_1$. Поэтому

$$d\mathbf{F}_1 = \frac{\mu_0 I_1 I_2}{2\pi r^2} [d\mathbf{l}_1 [\mathbf{n}_2 \mathbf{r}]].$$

Замечая, что $[d\mathbf{l}_1 [\mathbf{n}_2 \mathbf{r}]] = \mathbf{n}_2 (d\mathbf{l}_1 \mathbf{r}) - \mathbf{r} (d\mathbf{l}_1 \mathbf{n}_2)$, и предполагая, что $d\mathbf{l}_1 \perp \mathbf{r}$, получим окончательно

$$d\mathbf{F}_1 = - \frac{\mu_0 I_1 I_2}{2\pi r^2} d\mathbf{l}_1 (\mathbf{n}_1 \mathbf{n}_2) \mathbf{r},$$

где \mathbf{n}_1 — единичный вектор, направленный вдоль $d\mathbf{l}_1$.

Мы видим, что сила, действующая на элемент $d\mathbf{l}_1$ первого проводника со стороны второго длинного линейного тока, пропорциональна произведению сил токов, текущих в проводниках, и обратно пропорциональна расстоянию между проводниками. Сила пропорциональна косинусу угла между проводниками и направлена либо против \mathbf{r} (при $\mathbf{n}_1 \mathbf{n}_2 > 0$), либо по \mathbf{r} (при $\mathbf{n}_1 \mathbf{n}_2 < 0$). В первом случае проводники притягиваются, а во втором — отталкиваются. Притягиваются параллельные токи, а отталкиваются — антипараллельные.

Возвратимся к общей формуле для силы, действующей на элемент линейного проводника в магнитном поле. Если мы хотим определить общее электродинамическое воздействие, испытываемое контуром с током, находящимся в магнитном поле, то следует разбить контур с током, находящимся в магнитном поле, на отдельные элементы, найти силы, действующие на каждый элемент, а затем исследовать воздействие этих сил на контур. Но такая процедура, как правило, очень

сложна и неудобна; проще поступить иначе, а именно определить сперва потенциальную функцию тока в магнитном поле. Представим себе с этой целью, что токовый контур, находящийся в магнитном поле, несколько сместился под действием электромагнитных сил (и, вообще говоря, деформировался). Если элемент контура $d\mathbf{l}$ сместился на отрезок \mathbf{a} , то работа силы $d\mathbf{F} = I[d\mathbf{l}\mathbf{B}]$, действующей на этот элемент, составит

$$dA = d\mathbf{F} \cdot \mathbf{a} = I\mathbf{a} [d\mathbf{l}\mathbf{B}] = I\mathbf{B} [d\mathbf{l}\mathbf{a}].$$

Но $[d\mathbf{l}\mathbf{a}]$ представляет собой вектор площади ds , описанной элементом $d\mathbf{l}$ при его смещении на \mathbf{a} . Поэтому $dA = I\mathbf{B}ds$ и суммарная работа сил, действующих на отдельные элементы контура,

$$A = I \int_{S_b} \mathbf{B}ds,$$

где интегрирование совершается по поверхности «пояска» S_b , ограниченного начальным и конечным положениями контура.

Работа сил, действующих на контур с током, может быть выражена через магнитный поток $\Phi = \int_S \mathbf{B}ds$, пронизывающий контур (S — какая-либо поверхность, охватываемая контуром). Заметим сперва, что

для магнитного потока, как и для электрического потока, справедлива теорема Гаусса, но магнитных зарядов, в отличие от электрических, не существует, в силу чего магнитный поток через произвольную замкнутую поверхность должен обращаться в нуль.

Поэтому, если взять какой-либо контур и натянуть на него две произвольные поверхности S и S' , магнитные потоки через них будут одинаковы. (При этом следует лишь одинаково ориентировать элементы поверхностей S и S' , а именно мы будем всегда предполагать, что вектор площади и направление обхода контура образуют правовинтовую систему.) Таким образом, можно говорить о магнитном потоке, пронизывающем данный контур.

Обратимся теперь к формуле для работы A . Натянем на кривые 1 и 2, соответствующие двум положениям токового контура, произвольные поверхности S_1 и S_2 , и рассмотрим замкнутую поверхность $S_1 + S_2 + S_b$, образованную поверхностями S_1 и S_2 и поверхностью «пояска» S_b . Магнитный поток через эту поверхность равен нулю:

$$\int_{S_1+S_2+S_b} \mathbf{B}ds = 0$$

(ds — вектор площади, направленный всюду по внешней нормали к поверхности). Обозначим далее Φ_1 и Φ_2 магнитные потоки, пронизывающие токовый контур в двух его положениях: $\Phi_1 = \int_{S_1} \mathbf{B}ds_1$, $\Phi_2 = \int_{S_2} \mathbf{B}ds_2$ (направление обхода токового контура считается совпадающим с направлением тока). Векторы ds_b и ds_1 направлены по

внешней нормали к замкнутой поверхности, вектор же ds_2 направлен по внутренней нормали к ней; поэтому $\Phi_1 - \Phi_2 + \Phi_b = 0$, где $\Phi_b = \int S_b \mathbf{B} ds_b$. Таким образом, $\Phi_b = \Phi_2 - \Phi_1$ и, следовательно,

$$A = I (\Phi_2 - \Phi_1).$$

Итак, работа сил магнитного поля при перемещении контура с током равна произведению силы тока на приращение магнитного потока, пронизывающего токовый контур. Так как силы направлены всегда так, чтобы работа была положительной, то $\Phi_2 > \Phi_1$, т. е. контур перемещается и деформируется таким образом, чтобы магнитный поток через него был максимальным. В частности,

контур с током перемещается всегда в направлении более сильного поля. Если проводник, составляющий контур, может деформироваться (но не растягиваться), то действующие на него силы стремятся придать контуру форму круга, так как при заданном периметре круг имеет наибольшую площадь, а магнитный поток при прочих равных условиях пропорционален площади, охватываемой токовым контуром. Наконец, контур с током стремится ориентироваться относительно поля таким образом, чтобы направление поля образовывало с направлением тока праввинтовую систему.

Работа при движении в потенциальном поле равна убыли потенциальной энергии. Хотя магнитное поле и не является потенциальным, мы можем тем не менее ввести *потенциальную функцию тока в магнитном поле*, уменьшение которой определит совершенную силами поля работу. Эта функция играет роль потенциальной энергии тока во внешнем магнитном поле. Из формулы для работы видно, что потенциальную функцию тока U следует определить как

$$U = -I\Phi,$$

где Φ — магнитный поток, пронизывающий контур; при этом работа связана с потенциальной функцией соотношением

$$A = U_1 - U_2.$$

Рассмотрим в качестве примера рамку с током I , находящуюся в однородном поле с индукцией B и могущую вращаться около оси. Если \mathbf{S} — вектор, направленный по нормали к рамке и равный по модулю охватываемой ею площади, то рамку пронизывает магнитный поток $\Phi = BS = BS \cos \vartheta$, где ϑ — угол между нормалью к плоскости рамки и направлением поля. Умножив Φ на $-I$, найдем потенциальную функцию тока:

$$U = -IBS = -mB = -mB \cos \vartheta,$$

где $\mathbf{m} = IS$ — магнитный момент тока. Мы видим, что функция U равна в случае рамки взятому с обратным знаком скалярному произведению магнитного момента тока на вектор магнитной индукции. Функция U зависит от угла между нормалью к плоскости рамки и магнитным полем. Взяв производную от U по углу ϑ , найдем момент силы K , вращающий рамку: $K = -\partial U / \partial \vartheta = -mB \sin \vartheta$, или в векторной форме

$$K = [mB].$$

Силы стремятся повернуть рамку так, чтобы угол ϑ стал равным нулю, т. е. чтобы потенциальная функция стала минимальной.

4.5. Индуктивность

Определим теперь потенциальную функцию двух токов I_1 и I_2 . Функция U_{12} тока I_1 в магнитном поле $B^{(2)}$, порождаемом током I_2 , есть $U_{12} = -I_1\Phi_{21}$, где Φ_{21} — магнитный поток, создаваемый током I_2 и пронизывающий контур первого тока. Так как индукция $B^{(2)}$ пропорциональна I_2 , то и поток Φ_{21} пропорционален I_2 , т. е. $\Phi_{21} = L_{21}I_2$, где L_{21} — некоторый коэффициент, зависящий только от геометрии контуров, т. е. от их формы, размеров и взаимного расположения. Его называют *взаимной индуктивностью* контуров. Подставляя это выражение в формулу для U_{12} , получим

$$U_{12} = -L_{21}I_1I_2.$$

Аналогично можно определить функцию U_{21} тока I_2 в поле $B^{(1)}$, порождаемом током I_1 :

$$U_{21} = -L_{12}I_1I_2.$$

Но функции U_{12} и U_{21} должны быть равны между собой (иначе не будет выполняться закон равенства действия и противодействия): $U_{12} = U_{21}$; следовательно, $L_{12} = L_{21}$.

Легко найти явный вид L_{12} в случае достаточно тонких проводников. Запишем для этого Φ_{21} в виде $\Phi_{21} = \int_{S_1} B^{(2)} ds_1$, где ds_1 — ориентированный элемент поверхности, опирающийся на контур L_1 , обтекаемый током I_1 . Полагая $B^{(2)} = \text{rot } A^{(2)}$, где $A^{(2)}$ — векторный потенциал, создаваемый током I_2 , получим согласно формуле Стокса

$$\Phi_{21} = \int_{S_1} \text{rot } A^{(2)} ds_1 = \int_{L_1} A^{(2)} dl_1$$

(dl_1 — ориентированный элемент контура L_1). В случае линейных проводников

$$A^{(2)} = \frac{\mu_0 I_2}{4\pi} \oint_{L_2} \frac{dl_2}{r_{21}},$$

где r_{21} — расстояние между элементами dl_2 и dl_1 . Поэтому

$$\Phi_{21} = \frac{\mu_0 I_2}{4\pi} \oint_{L_1} \oint_{L_2} \frac{dl_1 dl_2}{r_{21}},$$

откуда

$$L_{12} = \frac{\mu_0}{4\pi} \oint_{L_1} \oint_{L_2} \frac{dl_1 dl_2}{r_{21}}.$$

Мы видим, что величина L_{12} может быть как положительной, так и отрицательной или равной нулю.

Наряду с взаимной потенциальной функцией U_{12} мы должны ввести также собственные потенциальные функции U_{11} и U_{22} токов. Действительно, ток I_1 создает собственную магнитную индукцию $\mathbf{B}^{(1)}$, в которой находится сам. Поэтому он будет обладать некоторой собственной потенциальной функцией, связанной с электродинамическими силами взаимодействия между отдельными элементами контура. Чтобы определить эту функцию, казалось бы, нужно найти собственный магнитный поток и умножить его на $-I_1$. В действительности, однако, нужно взять половину этой величины, т. е.

$$U_{11} = -\frac{1}{2} I_1 \Phi_{11}.$$

Чтобы понять смысл введения множителя $\frac{1}{2}$, рассмотрим детальнее происхождение собственной потенциальной функции. Разобьем для этого ток I_1 на тонкие токовые нити ΔI_1 . Тогда потенциальная функция ΔU_{11} какой-либо нити ΔI_1 в поле остального тока $I_1 - \Delta I_1$ составит $\Delta U_{11} = -\Delta I_1 \Phi'_{11}$, где Φ'_{11} — магнитный поток, пронизывающий нить и создаваемый остальным током $I_1 - \Delta I_1$. Но при бесконечно малом ΔI_1 поток Φ'_{11} бесконечно мало отличается от потока Φ_{11} . Последний же, как и индукция $\mathbf{B}^{(1)}$, пропорционален вызвавшему его току I_1 , т. е.

$$\Phi_{11} = L_1 I_1,$$

где L_1 — некоторый коэффициент, зависящий от размеров и формы контура. Его называют *индуктивностью контура*. Заметим, что $L_1 > 0$. Итак, при бесконечно малом токе нити dI_1 имеем $dU_{11} = -L_1 I_1 dI_1$. Просуммировав это выражение по всем нитям, найдем всю собственную потенциальную функцию первого контура: $U_{11} = -L_1 I_1^2 / 2$. Заменяя здесь $L_1 I_1$ на Φ_{11} , получим приведенную выше формулу для U_{11} . Аналогичным образом вводится собственная потенциальная функция второго тока

$$U_{22} = -\frac{1}{2} L_2 I_2^2 = -\frac{1}{2} I_2 \Phi_{22},$$

где L_2 — индуктивность второго контура, связанная с собственным потоком Φ_{22} второго контура соотношением $\Phi_{22} = L_2 I_2$.

Если сложить собственные потенциальные функции U_{11} и U_{22} и взаимную потенциальную функцию U_{12} , то мы найдем суммарную потенциальную функцию обоих токов: $U = U_{11} + U_{12} + U_{22}$. Ее можно представить в двух видах:

$$U = -\frac{1}{2} (I_1 \Phi_{11} + I_1 \Phi_{21} + I_2 \Phi_{22} + I_2 \Phi_{12});$$

$$U = -\frac{1}{2} (L_1 I_1^2 + 2L_{12} I_1 I_2 + L_2 I_2^2),$$

где $L_{12} = L_{21}$ — *взаимная индуктивность*.

Мы видим, что U представляет собой квадратичную форму токов, причем эта форма, как мы сейчас разъясним, является отрицательной. Отсюда легко заключить, что всегда должно выполняться условие

$$|L_{12}| \leq \sqrt{L_1 L_2}.$$

Определим индуктивность соленоида. Соленоид может быть либо прямым, либо круговым; мы будем лишь считать его достаточно длинным и тонким, чтобы можно было предполагать, что магнитное поле внутри соленоида однородно и, кроме того, равно нулю вне соленоида: $B = \mu_0 In$. Поэтому магнитный поток через один виток $BS = \mu_0 nSI$, где S — площадь поперечного сечения соленоида. Умножив эту величину на общее число витков N , получим собственный магнитный поток $\Phi = \mu_0 nNSI$. Наконец, приравняв эту величину LI , найдем индуктивность соленоида: $L = \mu_0 nNS$. Так как $n = N/l$ (l — длина соленоида), то

$$L = \mu_0 N^2 S / l.$$

Таким образом, индуктивность соленоида пропорциональна квадрату числа его витков.

Найдем теперь собственную потенциальную функцию соленоида. Подставив в формулу $U = -1/2 LI^2$ найденное выражение для L , получим $U = -\mu_0 SN^2 I^2 / (2l)$. Замечая, что $I = B / (\mu_0 n)$, представим U в виде

$$U = -B^2 V / (2\mu_0),$$

где $V = Sl$ — объем соленоида.

Рассмотрим далее соленоид с двумя обмотками, содержащими соответственно N_1 и N_2 витков, и пусть по ним протекают разные токи I_1 и I_2 . Токи могут течь либо в одном и том же, либо в разных направлениях. Спрашивается: какова взаимная индуктивность L_{12} обмоток? Чтобы найти ее, заметим, что ток I_2 создает индукцию $B^{(2)} = \mu_0 n_2 I_2$ (n_2 — число витков второй обмотки, приходящееся на единичную длину соленоида). Поэтому ток I_2 , протекающий через один виток первой обмотки, создает магнитный поток $SB^{(2)} = \mu_0 n_2 SI_2$. Умножив эту величину на общее число витков N_1 первой обмотки, найдем поток Φ_{21} , создаваемый током I_2 и пронизывающий первую обмотку: $\Phi_{21} = \mu_0 n_2 SN_1 I_2$. Сравнивая это выражение с формулой $\Phi_{21} = L_{12} I_2$, имеем $L_{12} = \mu_0 n_2 N_1 S$. Эта величина должна браться со знаком плюс, если токи текут в одном направлении, и со знаком минус, если токи текут в разных направлениях (это правило следует из того, что в общем выражении для L_{21} входит под знаком интеграла скалярное произведение элементов проводников $d\mathbf{l}_1 d\mathbf{l}_2$. Подставляя в формулу для коэффициента взаимной индукции $n_2 = N_2/l$ (N_2 — число витков второй обмотки), получим окончательно

$$L_{12} = \pm \mu_0 N_1 N_2 S / l = \pm \mu_0 n_1 n_2 V.$$

Мы видим, что взаимная индуктивность пропорциональна произведению чисел витков обеих обмоток. Сравнение формулы для L_{12} с формулами для индуктивностей L_1 , L_2 обмоток показывает, что

$$|L_{12}| = \sqrt{L_1 L_2}.$$

Найдем, наконец, потенциальную функцию обоих токов:

$$U = -1/2 (L_1 I_1^2 + 2L_{12} I_1 I_2 + L_2 I_2^2).$$

Подставляя сюда формулы для L_{12} и $L_1 L_2$ и учитывая, что $\mu_0 n_1 I_1 = B^{(1)}$; $\mu_0 n_2 I_2 = B^{(2)}$, получим

$$U = -\frac{1}{2\mu_0} (B^{(1)} \pm B^{(2)})^2 V.$$

А так как суммарная магнитная индукция $B = B^{(1)} \pm B^{(2)}$ (знак плюс соответствует одинаковым, а знак минус — противоположным направлениям токов), то окончательно имеем

$$U = -\frac{1}{2\mu_0} B^2 V.$$

Это та же формула, что и формулы для собственной потенциальной функции. Более того, можно показать, что формулой такого же типа определяется силовая функция в любом случае, для любых токов, а именно U всегда выражается в виде интеграла от квадрата магнитной индукции:

$$U = -\frac{1}{2\mu_0} \int B^2 dV,$$

где интегрирование совершается по всему объему (вектор магнитной индукции может меняться от точки к точке). Мы видим, что величина U может быть отрицательной или равной нулю (но не положительной) в соответствии со сказанным выше.

Как мы увидим далее, эта формула имеет глубокий физический смысл: величина $B^2/(2\mu_0)$ может интерпретироваться как плотность энергии магнитного поля в вакууме.

Приведем в заключение единицы СИ для введенных в этой главе величин. Магнитный поток Φ выражается в веберах (Вб): $1 \text{ Вб} = 1 \text{ м}^2 \cdot \text{кг} \cdot \text{с}^{-2} \text{ А}^2 = \text{В} \cdot \text{с}$. Единицу индукции называют теслой (Тл): $1 \text{ Тл} = 1 \text{ Вб/м}^2$. Единицу индуктивности называют генри (Гн): $\text{Гн} = \text{Вб/А}^{-1}$.

ОСНОВНЫЕ ФОРМУЛЫ

Сила Лоренца $\mathbf{f} = e(\mathbf{E} + [\mathbf{v}\mathbf{B}])$

Циклотронная частота $\omega_e = qB/m$

Сила, действующая на элемент тока в магнитном поле $d\mathbf{F} = I [d\mathbf{l}\mathbf{B}]$

Закон Био—Савара $\mathbf{B} = \frac{\mu_0 I}{4\pi} \oint \frac{[d\mathbf{l}\mathbf{r}]}{r^3}$

Сила взаимодействия между токами $d\mathbf{F}_i = \frac{\mu_0}{4\pi} I_1 I_2 \left[d\mathbf{l}_1 \oint \frac{[d\mathbf{l}_2 \mathbf{r}_{21}]}{r_{21}^3} \right]$

Коэффициент взаимной индукции между контурами $L_{12} = \frac{\mu_0}{4\pi} \oint \oint \frac{d\mathbf{l}_1 d\mathbf{l}_2}{r_{21}}$

Глава 5. ЭЛЕКТРОМАГНИТНАЯ ИНДУКЦИЯ И ПЕРЕМЕННЫЕ ТОКИ

5.1. Закон Фарадея

До сих пор при изучении электрических и магнитных явлений мы молчаливо предполагали, что электрическое и магнитное поля никак не связаны между собой. Между тем такое предположение справедливо только в том случае, когда поля не меняются с течением времени. Если же поля не постоянны, то они уже не могут быть независимыми. Дело в том, что

переменное магнитное поле всегда создает переменное же электрическое поле, а переменное электрическое поле создает переменное магнитное поле.

Первая сторона этой взаимосвязи проявляется в замечательном явлении *электромагнитной индукции*. Явление это, открытое Фарадеем в 1831 г., состоит в том, что если возле проводника двигать магнит либо проводник, по которому течет постоянный ток, то в первом проводнике возникает электрический ток без всяких посторонних источников. Ток возникает и тогда, когда проводник движется, а магнит неподвижен. Наконец, ток возникает в неподвижном проводнике и в том случае, если возле него находится другой неподвижный проводник, по которому течет не постоянный, а переменный ток.

Общим между магнитом и проводником с постоянным током является то, что магнит и проводник создают вокруг себя магнитное поле. Поэтому можно утверждать, что возникновение индукционного тока в опытах с магнитом и постоянным током связано с наличием магнитного поля и относительностью движения проводника и поля.

В опыте с проводником, по которому течет переменный ток, имеется только магнитное поле тока, но поле это является переменным. Вместе с током возникает и электрическое поле. Поэтому опыт с переменным током явно показывает, что переменное магнитное поле окружено электрическим полем. Как мы убедимся далее, явление электромагнитной индукции при относительном движении проводника и магнита означает, по существу, то же самое, т. е. создание переменным магнитным полем поля электрического.

Переходя к изучению электромагнитной индукции, мы начнем с того случая, когда магнит (или проводник с постоянным током) покоится и в его поле движется замкнутый проводник. Причину появления тока в движущемся проводнике легко понять. Действительно, если, например, проводник ab (рис. 5.1) перемещается перпендикулярно магнитному полю с индукцией \mathbf{B} со скоростью \mathbf{v} , то заряд e , находящийся в проводнике, под действием силы Лоренца

$$\mathbf{f} = e [\mathbf{vB}]$$

движется вдоль проводника. Если теперь разрешить проводнику ab свободно скользить вдоль металлической рамки $ABCD$ (рис. 5.2), то в замкнутой цепи $abCD$ потечет электрический ток. Чтобы вычис-

лить силу этого тока, найдем сперва электродвижущую силу (э. д. с.) индукции \mathcal{E} . Э. д. с. всегда определяется как интеграл по проводящему контуру от силы f/e , действующей на единичный заряд, т. е.

$$\mathcal{E} = \frac{1}{e} \oint_C f dl,$$

где интегрирование проводится вдоль контура C в направлении, совпадающем с направлением тока. Поэтому в рассматриваемом случае

$$\mathcal{E} = \oint_{C_1} [v\mathbf{B}] dl,$$

а так как движется только один участок цепи ab и $\mathbf{v} \perp \mathbf{B}$, а $dl \perp \mathbf{B}$, то, очевидно,

$$\mathcal{E} = vBl,$$

где l — длина подвижного участка цепи ab . Разделив \mathcal{E} на полное сопротивление R всей замкнутой цепи, найдем в соответствии с законом Ома силу тока в цепи:

$$I = \mathcal{E}/R = vBl/R.$$

Полученное выражение для э. д. с. можно преобразовать, если заметить, что $v = dx/dt$, где x — путь, пройденный проводником ab . Обозначим $S = xl$ площадь рамки, тогда

$$\mathcal{E} = \frac{d}{dt} BS.$$

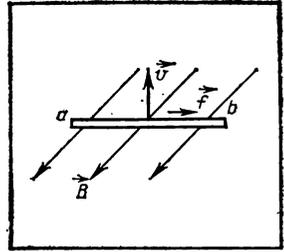
Но BS представляет собой магнитный поток через проводящую рамку. Поэтому можно сказать, что

э. д. с. индукции пропорциональна производной по времени от магнитного потока, пронизывающего проводящий контур.

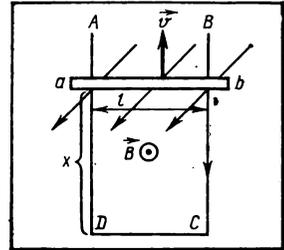
Уточним этот вывод, вспомнив определение магнитного потока. Мы условились понимать под магнитным потоком, пронизывающим контур C , поверхностный интеграл

$$\Phi = \int_{\Sigma} \mathbf{B} ds,$$

где интегрирование совершается по произвольной поверхности Σ , опирающейся на контур C , причем элемент поверхности ds считается ориентированным таким образом, чтобы его направление и направление обхода контура C образовали правовинтовую систему. Поэтому



5.1. На заряд в проводнике ab действует сила \mathbf{f} Лоренца



5.2. В замкнутой цепи $ABCD$ индуцируется ток

при определении магнитного потока через рамку с направлением обхода, совпадающим с направлением индуцированного тока, мы должны покрыть ее поверхность, скажем, плоскостью и считать нормаль к ней \mathbf{n} направленной от нас, за плоскость чертежа. А так как магнитная индукция \mathbf{B} направлена перпендикулярно плоскости чертежа на нас, то для потока мы получим выражение $\Phi = \mathbf{BS} = -BS$. Поэтому э. д. с. индукции можно представить в виде

$$\mathcal{E} = -d\Phi/dt.$$

Таким образом,

э. д. с. индукции равна взятой с обратным знаком производной по времени от магнитного потока, пронизывающего проводящий контур.

Знак минус в этой формуле означает, что индуцируется ток такого направления, при котором

магнитное поле, создаваемое током, стремится ослабить изменение магнитного потока, вызвавшего индукционный ток (правило Ленца).

Мы определили э. д. с. индукции для прямоугольной рамки, но точно такой же формулой определяется э. д. с. и в общем случае для произвольного проводящего контура, как угодно двигающегося в постоянном магнитном поле, причем контур может и деформироваться. Э. д. с. индукции всегда равна взятой с обратным знаком производной по времени от магнитного потока, пронизывающего контур, и индуцируемый ток своим собственным магнитным полем всегда стремится ослабить изменение внешнего магнитного потока, вызывающего ток.

До сих пор мы считали, что проводящий контур движется, а магнит (или проводник с постоянным током) неподвижен. Но проводящий контур можно не двигать, а двигать только магнит (или другой проводник, по которому течет постоянный ток); в этом случае в контуре также будет индуцироваться ток. Более того, э. д. с. индукции по-прежнему определится формулой $\mathcal{E} = -d\Phi/dt$, где Φ — магнитный поток, пронизывающий контур, в данном случае неподвижный. Итак,

для возникновения э. д. с. индукции необходимо лишь, чтобы изменялся во времени магнитный поток, пронизывающий проводящий контур, но движется или покоится контур, безразлично. Э. д. с. индукции всегда пропорциональна производной по времени от этого потока.

В этом заключается общий закон электромагнитной индукции, открытый Фарадеем.

5.2. Относительность электрического и магнитного полей

Как же объяснить явление электромагнитной индукции в случае неподвижного проводника? Если проводник движется, а магнит неподвижен, то возникновение тока в проводнике

мы связывали с силой Лоренца $e[\mathbf{vB}]$. Но если проводник неподвижен, то этой силы нет, почему же тогда возникает ток? Если проводник неподвижен, то на его заряды может действовать только электрическое поле, и мы приходим к выводу, что в этом случае в проводнике должно индуцироваться электрическое поле. Электрическое поле должно создаваться при движении магнита, т. е. должно создаваться переменным магнитным полем.

Э. д. с., действующая в контуре, всегда выражается в виде интеграла по контуру от силы, действующей на единичный заряд. Поэтому э. д. с. индукции \mathcal{E} в случае неподвижного контура C должна иметь вид

$$\mathcal{E} = \oint_C \mathbf{E} d\mathbf{l},$$

где \mathbf{E} — напряженность электрического поля, порождаемого в неподвижном контуре переменным магнитным полем движущегося магнита, и интегрирование совершается вдоль проводящего контура в направлении, совпадающем с направлением индуцированного тока.

Итак, мы приходим к следующей формулировке закона электромагнитной индукции в случае неподвижного проводника:

циркуляция электрического поля, индуцируемого в проводящем контуре переменным магнитным полем, совпадает с взятой с обратным знаком производной по времени от магнитного потока, пронизывающего контур:

$$\oint_C \mathbf{E} d\mathbf{l} = -d\Phi/dt,$$

векторы $d\mathbf{l}$ и $d\mathbf{s}$ образуют вместе правовинтовую систему.

Обратим внимание на следующее обстоятельство. Циркуляция электрического поля, возникающего при электромагнитной индукции, может быть отличной от нуля (если только отлично от нуля изменение магнитного потока через контур), в то время как циркуляция электростатического поля всегда равна нулю. Иными словами,

при электромагнитной индукции возникает не потенциальное, а вихревое электрическое поле.

Вернемся теперь к выражению для э. д. с. индукции. В случае движущегося проводника она определяется формулой

$$\mathcal{E} = \int_C [\mathbf{vB}] d\mathbf{l},$$

а в случае покоящегося проводника — формулой

$$\mathcal{E} = \oint_C \mathbf{E} d\mathbf{l}.$$

По внешнему виду это две совершенно различные и никак не связанные между собой формулы. Поэтому создается впечатление, что при разъяснении явления электромагнитной индукции в случае движущегося

гося проводника и покоящегося магнита и в случае покоящегося проводника и движущегося магнита нужно пользоваться различными подходами и различными формулами. Между тем явление в обоих случаях одно и то же, ибо движение по сути своей всегда относительно, и совершенно безразлично, что двигать — проводник или магнит.

Однако впечатление это не соответствует действительности, и обе формулы для э. д. с., различные по виду, по сути дела тождественны. Дело в том, что электрическое и магнитное поля образуют единое целое — *электромагнитное поле*, разделение же этого единого поля на поле электрическое и поле магнитное имеет относительный характер; иными словами, разделение это происходит по-разному в разных системах отсчета. Чтобы убедиться в этом, обратимся к уравнению движения заряженной частицы в электрическом и магнитном полях:

$$m \frac{d\mathbf{v}}{dt} = e(\mathbf{E} + [\mathbf{v}\mathbf{B}])$$

Это уравнение, так же как и уравнение движения Ньютона, должно быть справедливо в любой инерциальной системе отсчета. Однако при переходе от одной инерциальной системы отсчета к другой скорость частицы изменяется, а ускорение остается неизменным. Поэтому левая часть уравнения останется неизменной, правая же изменится, если только соответствующим образом не изменятся поля.

Итак, если написанное уравнение движения относится к системе K , то в системе K' уравнение движения имеет вид

$$m \frac{d\mathbf{v}'}{dt'} = e(\mathbf{E}' + [\mathbf{v}'\mathbf{B}']),$$

где штрих служит для обозначения величин в системе K' . Значение \mathbf{v}' связано с \mathbf{v} законом сложения скоростей: $\mathbf{v}' = \mathbf{v} - \mathbf{V}$ (\mathbf{V} — скорость системы K' относительно K). Так как ускорение в обеих системах одинаково, то мы приходим к соотношению

$$\mathbf{E} + [\mathbf{v}\mathbf{B}] = \mathbf{E}' + [\mathbf{v}'\mathbf{B}'],$$

которое должно выполняться для любых \mathbf{v} и \mathbf{V} . Отсюда может быть найдена связь между штрихованными и нештрихованными полями.

Нужно, однако, иметь в виду, что в действительности скорости \mathbf{v} и \mathbf{V} не могут быть любыми, так как в исходное уравнение движения, и закон сложения скоростей справедливы только при малых по сравнению с c скоростях. Поэтому написанное соотношение позволяет найти связь между полями в системах K и K' только в случае малых V , т. е. при $V \ll c$. Связь эта такова:

$$\mathbf{E}' = \mathbf{E} + [\mathbf{V}\mathbf{B}]; \quad \mathbf{B}' = \mathbf{B} - [\mathbf{V}\mathbf{E}]$$

(В гл. 7 будет дано обобщение этих формул на любые скорости.)

Из этих формул следует, что если в системе K электрическое поле отсутствует, но отлично от нуля магнитное поле, то в системе K' отлично от нуля электрическое поле (если только $[\mathbf{V}\mathbf{B}] \neq 0$).

Поэтому если проводник движется относительно магнита, то в системе отсчета, связанной с проводником, существует электрическое поле $\mathbf{E}' = [\mathbf{v}\mathbf{B}]$, а так как в этом случае э. д. с. индукции

$$\mathcal{E} = \oint_C [\mathbf{v}\mathbf{B}] \, d\mathbf{l},$$

то ее можно представить в виде

$$\mathcal{E} = \oint_C \mathbf{E}' \, d\mathbf{l}.$$

Но это та же формула, что и формула для э. д. с. индукции в случае неподвижного проводника. Таким образом, э. д. с. индукции всегда равна циркуляции электрического поля в системе отсчета, связанной с проводником, т. е. в системе, где проводник неподвижен. Закон же электромагнитной индукции утверждает, что эта циркуляция всегда определяется полным изменением магнитного потока, пронизывающего контур.

5.3. Индукционные генераторы тока

Явление электромагнитной индукции может быть использовано для построения *генераторов тока*. Простейшим является *генератор переменного тока*, работа которого основана на том, что при равномерном вращении проводящей рамки в постоянном магнитном поле в ней индуцируется переменная э. д. с., синусоидально изменяющаяся со временем. В самом деле, если магнитное поле однородно, то рамку пронизывает магнитный поток Φ :

$$\Phi = \mathbf{B}\mathbf{S} = BS \cos \vartheta,$$

где \mathbf{B} — магнитная индукция, S — площадь рамки и ϑ — угол между \mathbf{B} и нормалью к рамке. При равномерном вращении рамки с угловой частотой ω имеем $\vartheta = \omega t + \vartheta_0$ (ϑ_0 — начальное значение ϑ) и

$$\Phi = BS \cos (\omega t + \vartheta_0).$$

Э. д. с. индукции, которую мы обозначим здесь $e(t)$, связана с Φ законом Фарадея:

$$e(t) = -d\Phi/dt.$$

Подставляя сюда выражение для Φ , получим

$$e(t) = BS \omega \sin (\omega t + \vartheta_0).$$

Таким образом, при равномерном вращении рамки (витка) в постоянном магнитном поле в рамке возникает э. д. с., периодически изменяющаяся во времени по закону синуса. Период изменения э. д. с. равен, очевидно, $T = 2\pi/\omega$, а ее амплитуда \mathcal{E}_m определяется при однородном поле произведением индукции, частоты и площади рамки: $\mathcal{E}_m = BS\omega$. Для того чтобы наводимую в витке э. д. с. передать во внеш-

нюю цепь, достаточно концы витка присоединить к двум кольцам, вращающимся вместе с витком, внешнюю же цепь связать с кольцами с помощью двух щеточных контактов.

Как было уже разъяснено, для получения э. д. с. индукции не обязательно двигать проводник, с равным успехом можно получить э. д. с. индукции и в случае неподвижного проводника, если двигать источник магнитного поля. Именно этот принцип и используется во всех современных мощных генераторах переменного тока (*синхронных генераторах*). Строятся такие генераторы главным образом для получения так называемого *трехфазного тока*, использование которого дает ряд преимуществ.

Трехфазным током называют совокупность трех синусоидальных токов, сдвинутых по фазе на угол $2\pi/3$ друг относительно друга.

При одинаковой амплитуде этих токов их мгновенные значения составляют

$$i_1 = I \sin \omega t, \quad i_2 = I \sin (\omega t - 2/3 \pi), \quad i_3 = I \sin (\omega t + 2/3 \pi),$$

где I — амплитуда тока, ω — его частота. Легко проверить непосредственным расчетом, что сумма этих токов равна нулю: $i_1 + i_2 + i_3 = 0$.

Основное преимущество трехфазного тока и связано с этим равенством. Дело в том, что для использования трех постоянных токов требуется шесть проводов — по паре проводов на каждый ток. Если же сумма токов равна нулю, то можно использовать только три провода.

В обычных генераторах электрического тока э. д. с. индукции наводится в твердых металлических проводниках, которые движутся в магнитном поле, для чего используются механические двигатели. Но можно построить генератор и иначе, применяя в качестве рабочего тела не твердый металлический проводник, а ионизованный газ. Такой генератор носит название *магнитогидродинамического* или сокращенно *МГД-генератора*.

Рассмотренные генераторы переменного и постоянного токов обладают важным свойством обратности: если в генератор подать извне соответствующее напряжение, то подвижная часть генератора начнет вращаться. Таким образом, генератор может, в принципе, работать как электрический двигатель. Мы не будем останавливаться на этом вопросе подробно.

5.4. Э.д.с. взаимной индукции и самоиндукции

Рассмотрим два проводящих контура, и пусть в одном из них, скажем в первом, действует некоторая сторонняя э. д. с. $e_1^{(e)}(t)$, зависящая от времени. Тогда ток в первом контуре i_1 также зависит от времени. От времени зависит и магнитное поле $\mathbf{B}^{(1)}$, создаваемое током i_1 , а следовательно, и магнитный поток Φ_{21} , пронизывающий второй контур и создаваемый током в первом контуре. Это значит, что во втором контуре возникает э. д. с. индукции

$$e_{21}(t) = -d\Phi_{21}/dt.$$

Эта э. д. с. носит название *э. д. с. взаимной индукции*.

В гл. 4 мы показали, что $\Phi_{21} = L_{21}i_1$, где L_{21} — взаимная индуктивность контуров. Строго говоря, эта формула справедлива для постоянного тока, но ею можно пользоваться также и в случае переменного тока, если только он изменяется достаточно медленно (в этом случае говорят о *квазистационарном токе*); необходимо, чтобы время, в течение которого ток претерпевает существенное изменение, было велико по сравнению с временем, которое требуется свету, чтобы пройти от одного контура к другому, а также пройти расстояние порядка размеров контуров. Мы будем предполагать это условие выполненным. (Заметим, что если контуры могут двигаться и изменять свою форму и размеры, то следует дифференцировать все произведение $L_{21}i_1$.)

Э. д. с. взаимной индукции вызовет во втором контуре ток i_2 , который, так же как и ток i_1 , будет переменным. Поэтому он, в свою очередь, вызовет в первом контуре э. д. с. взаимной индукции

$$e_{12}(t) = -\frac{d}{dt} \Phi_{12} = -\frac{d}{dt} (L_{12}i_2).$$

Кроме e_{21} и e_{12} в контурах возникнут еще две э. д. с. индукции, связанные с изменением собственных магнитных потоков. Действительно, собственные магнитные потоки Φ_{11} и Φ_{22} , вызываемые токами i_1 и i_2 в своих контурах, пропорциональны токам:

$$\Phi_{11} = L_1 i_1, \quad \Phi_{22} = L_2 i_2.$$

(L_1 и L_2 — индуктивности контуров). Поэтому они меняются со временем и вызывают в контурах э. д. с. индукции

$$e_{11} = -\frac{d\Phi_{11}}{dt} = -\frac{d}{dt} (L_1 i_1),$$

$$e_{22} = -\frac{d\Phi_{22}}{dt} = -\frac{d}{dt} (L_2 i_2).$$

Эти э. д. с. называют *э. д. с. самоиндукции*. Если контуры не изменяют своих размеров и формы, то значения L постоянны, в противном случае следует учитывать изменение L .

Итак, в первом контуре действует суммарная э. д. с.

$$e_1 = e_1^{(e)} + e_{11} + e_{12} = e_1^{(e)} - \frac{d}{dt} (L_1 i_1) - \frac{d}{dt} (L_{12} i_2),$$

а во втором — суммарная э. д. с.

$$e_2 = -d(L_2 i_2)/dt - d(L_{21})/dt.$$

5.5. Бетатрон

Как уже говорилось в предыдущих параграфах, возникновение индукционного тока есть, по сути дела, вторичный эффект,

первичным же является создание переменным магнитным полем вихревого электрического поля. Чтобы убедиться в этом непосредственно, следует видоизменить опыт, в котором наблюдается электромагнитная индукция, а именно нужно сделать так, чтобы в нем не было проводника, а остались только свободные заряды. Тогда если переменное магнитное поле действительно создает вокруг себя вихревое электрическое поле, то свободные заряды должны ускоряться этим полем. Такой процесс и происходит в бетатроне, в котором свободные электроны вращаются по круговой орбите, пронизываемой переменным магнитным потоком.

Покажем, как определить ускоряющее электрическое поле в бетатроне. Для этого следует воспользоваться законом электромагнитной индукции

$$\oint_C \mathbf{i} \mathbf{E} d\mathbf{l} = - \frac{d\Phi}{dt}$$

Но прежде всего возникает вопрос, что понимать в бетатроне под контуром C : здесь нет обычного проводящего контура, с которым мы имели дело в индукционных опытах. Под C следует всегда понимать *материальный контур*, т. е. контур, образуемый определенными частицами, или элементами среды. В случае бетатрона мы возьмем в качестве частиц ускоряемые электроны, и тогда контуром C будет их круговая траектория.

В бетатроне вектор магнитной индукции \mathbf{B} перпендикулярен плоскости траектории и является функцией как времени t , так и расстояния r до центра траектории: $\mathbf{B} = \mathbf{B}(r, t)$. Поэтому магнитный поток, пронизывающий траекторию, определяется формулой

$$\Phi = \int \mathbf{B} ds = \int_0^R B(r, t) 2\pi r dr,$$

где R — радиус круговой орбиты электрона.

Силовые линии индуцируемого электрического поля \mathbf{E} в бетатроне, является, очевидно, окружностями, concentрическими с круговой орбитой электрона, причем напряженность поля зависит как от r , так и от t . Поэтому циркуляция электрического поля вдоль этой орбиты равна $2\pi ER$. Приравняв циркуляцию $d\Phi/dt$, мы найдем ускоряющее электрическое поле на орбите электрона:

$$E = \frac{1}{2\pi R} \frac{d}{dt} \int_0^R B(r, t) 2\pi r dr = \frac{R}{2} \langle \dot{B}(t) \rangle,$$

где

$$\langle B(t) \rangle = \frac{\Phi}{\pi R^2} = \frac{2}{R^2} \int_0^R B(r, t) r dr.$$

Эта величина имеет, очевидно, смысл средней индукции по площади, ограниченной орбитой. Производная по времени от средней индукции и определяет ускоряющее поле.

Под действием электрического поля E электрон ускоряется и импульс его p (направленный вдоль траектории) изменяется по закону

$$\frac{dp}{dt} = eE = \frac{eR}{2} \langle \dot{B}(t) \rangle.$$

Отсюда следует, что

$$p \equiv p(t) = \frac{eR}{2} (\langle B(t) \rangle - \langle B(0) \rangle) + p_0,$$

где p_0 — начальное значение импульса. Таким образом, импульс электрона растет пропорционально средней магнитной индукции. Поэтому если за период

обращения электрона средняя магнитная индукция возрастет на $\Delta\langle B \rangle$, то за каждый период обращения электрона импульс электрона увеличивается на $\Delta p = (eR/2)\Delta\langle B \rangle$.

Теперь мы должны выяснить, как сделать движение электрона круговым. Точнее говоря, можно ли сделать движение электрона круговым, используя то же магнитное поле $\mathbf{B}(r, t)$, которое служит для ускорения электрона?

Под действием поля $\mathbf{B}(r, t)$ электрон испытывает действие силы Лоренца, равной $e\mathbf{v}B(r, t)$ и направленной к центру орбиты. Поэтому если движение электрона происходит по окружности радиуса R , то мы должны приравнять эту силу произведению массы электрона m на центростремительное ускорение v^2/R , т. е. $mv^2/R = evB(R, t)$, откуда

$$mv = p(t) = eRB(R, t),$$

где $B(R, t)$ — мгновенное значение индукции на орбите.

Таким образом, для того чтобы движение электрона в переменном магнитном поле происходило по окружности, необходимо, чтобы импульс электрона изменялся пропорционально индукции магнитного поля на орбите. С другой стороны, мы уже нашли, как при ускорении меняется импульс электрона. Поэтому мы можем сравнить обе формулы, которые должны давать одинаковый результат, т. е.

$$\frac{eR}{2} (\langle B(t) \rangle - \langle B(0) \rangle) + p_0 = eRB(R, t).$$

Считая, что $p_0 = 1/2 eR\langle B(0) \rangle$, мы приходим к соотношению

$$\langle B(t) \rangle = 2B(R, t),$$

которое и представляет собой *условие существования круговой орбиты электронов в бетатроне*. Мы видим, что для этого необходимо, чтобы средняя магнитная индукция была в два раза больше значения магнитной индукции на орбите.

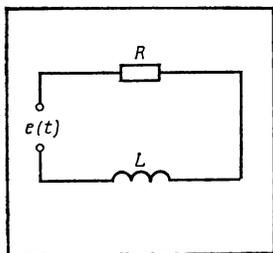
Бетатрон позволяет ускорять электроны до очень больших энергий — порядка сотен миллионов электрон-вольт, но для получения еще больших энергий он уже непригоден, главным образом из-за эффекта излучения.

5.6. Индуктивность и емкость в цепи переменного тока

Простота генерирования переменного тока, легкость его трансформации и возможность превращения его энергии в механическую работу привели к широчайшему использованию переменных токов на практике. Переходя к изучению закономерностей, которым подчиняются такие токи, начнем с выяснения роли индуктивности. Рассмотрим проводящий контур, в котором действует переменная э. д. с. или приложено переменное напряжение. В случае постоянной э. д. с. сила тока была бы равна \mathcal{E}/R , где R — сопротивление контура. Но при переменной э. д. с. $e(t)$ или переменном напряжении $u(t)$ переменным будет и ток, а в таком случае в контуре возникнет добавочная э. д. с. — э. д. с. самоиндукции $-Ldi/dt$, так что в контуре действует суммарная э. д. с. $e(t) - Ldi/dt$. Поэтому сила тока i определится из уравнения $i = (e - Ldi/dt)/R$, или

$$L \frac{di}{dt} + Ri = e(t).$$

Этому уравнению соответствует схема цепи с рассредоточенными индуктивностью L и сопротивлением R , изображенная на рис. 5.3.



5.3.

Последовательно соединенные сопротивление и индуктивность

Можно сказать, что напряжение $u \equiv e(t)$ складывается на два слагаемых $u_R = Ri$ и $u_L = L(di/dt)$, представляющих собой напряжения на сопротивлении и индуктивности: $u = u_R + u_L$. В действительности, конечно, не всегда можно в цепи реально разделить u на u_R и u_L . Но если цепь состоит из реостата с малой индуктивностью и катушки с большой индуктивностью, то разности потенциалов на реостате и катушке равны u_R и u_L . Для того чтобы реостат имел малую индуктивность, его следует намотать согнутой вдвое проволокой (*бифилярная намотка*). В этом случае противоположно направленные токи всюду оказываются рядом, благодаря чему

их магнитные поля почти компенсируются.

Уравнение, которое мы получили для определения силы тока, представляет собой неоднородное дифференциальное уравнение первого порядка с известной правой частью. Решение уравнений такого типа складывается из частного решения неоднородного уравнения и общего решения однородного уравнения (без правой части), причем сумма подбирается так, чтобы удовлетворить начальным условиям. Например, в случае постоянной э. д. с. $e(t) = \mathcal{E}$ естественно выбрать в качестве частного решение $i_1 = \mathcal{E}/R$ — значение тока, даваемое законом Ома; общее же решение однородного уравнения имеет вид $i = \text{const} e^{-t/\tau}$, где $\tau = L/R$ — так называемая *постоянная контура*. Выбирая в качестве $t = 0$ начальный момент времени, когда $i = 0$, получим $\text{const} = -\mathcal{E}/R$ и, следовательно,

$$i = \frac{\mathcal{E}}{R} (1 - e^{-t/\tau}).$$

Второе слагаемое в скобках становится малым через время порядка нескольких τ ; мы можем сказать, что «память» о начальном значении тока стирается за время, характеризующееся постоянной контура.

Рассмотрим теперь случай, когда в цепи действует синусоидальное напряжение $u(t) = U \cos \omega t$ (U — амплитуда, ω — частота). Чтобы найти частное решение в этом случае, удобно воспользоваться так называемым *комплексным методом*, в котором ток ищется в виде $i(t) = \text{Re} I e^{i\omega t}$, где I — некоторая комплексная константа, называемая *комплексной амплитудой тока*. Полагая $u(t) = \text{Re} U e^{i\omega t}$ и учитывая линейность дифференциального уравнения для тока, можно, очевидно, опустить в этом уравнении знак Re :

$$L \frac{d}{dt} (I e^{i\omega t}) + R I e^{i\omega t} = U e^{i\omega t}.$$

Отсюда находим

$$I = U/Z; \quad Z = R + i\omega L \quad (i \equiv \sqrt{-1}).$$

Величину Z называют *комплексным сопротивлением* (или *импедансом*)

цепи. (Что касается решения однородного уравнения, то мы не будем его выписывать, так как за время порядка нескольких τ оно обращается в нуль.)

Остановимся на энергетических соотношениях в рассматриваемой цепи. Очевидно, внешний источник э. д. с. подает за время dt в цепь энергию, равную $uidt$. Эта величина может быть на основании дифференциального уравнения для тока представлена в виде

$$uidt = \left(L \frac{di}{dt} + Ri \right) idt = Ri^2 dt + d \left(\frac{1}{2} Li^2 \right).$$

Первое слагаемое здесь определяет джоулеву теплоту, выделяющуюся за время dt , а второе слагаемое — изменение за dt магнитной энергии цепи. Произведение Ri^2 всегда положительно; величина же $d(\frac{1}{2}Li^2)/dt$ может иметь оба знака. Если $i = I \cos(\omega t - \varphi)$, то

$$\frac{d}{dt} \left(\frac{1}{2} Li^2 \right) = - \frac{1}{2} \omega LI^2 \sin(2\omega t - 2\varphi).$$

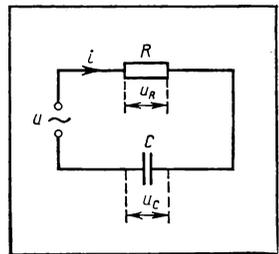
Это выражение отрицательно в первой и третьей четвертях периода изменения тока и положительно во второй и четвертой четвертях. Это значит, что во вторую четверть периода тока в цепи накапливается магнитная энергия, однако в течение третьей четверти периода она возвращается внешнему источнику э. д. с. Она вновь накапливается в цепи в четвертую четверть и, наконец, снова отдается источнику э. д. с. в первую четверть.

Таким образом, магнитная энергия непрерывно переходит из источника э. д. с. в цепь и обратно из цепи в источник э. д. с., среднее же значение ее за период тока (и даже за половину этого периода) равно нулю.

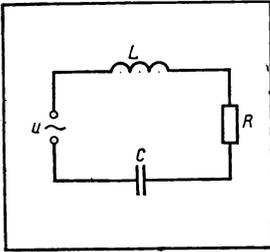
Переменный ток в отличие от постоянного может протекать по цепи, в которую последовательно включен конденсатор. Это связано с возникновением переменного заряда на обкладках конденсатора. Рассмотрим сперва простейший случай, когда индуктивность цепи ничтожно мала, так что схема цепи имеет вид, изображенный на рис. 5.4. Здесь R — сопротивление цепи, C — емкость конденсатора. Приложенное внешнее напряжение u равно, очевидно, сумме напряжений u_R и u_C на сопротивлении и на конденсаторе: $u = u_R + u_C$. Ясно, что $u_R = iR$, $u_C = q/C$, где i — сила тока и q — заряд на обкладке конденсатора. Поэтому $u = Ri + q/C$, а так как i и q связаны очевидным соотношением $i = dq/dt$, то

$$Rdq/dt + q/C = u.$$

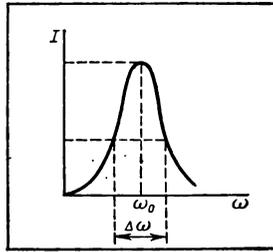
Это дифференциальное уравнение позволяет найти заряд q , а следовательно, и ток i . Пусть в цепи действует синусоидальное напряжение $u = \operatorname{Re} U e^{i\omega t}$. Тогда, используя комплексный метод, легко убедиться, что частным решением уравнения будет $q =$



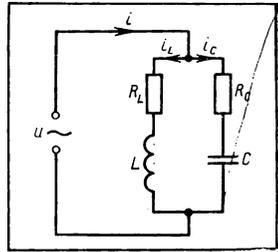
5.4. Последовательно соединенные сопротивление и емкость



5.5. Последовательно соединенные индуктивность, сопротивление и емкость



5.6. Частотная зависимость амплитуды колебаний в контуре



5.7. Простейший колебательный контур

$= \operatorname{Re} Q e^{i\omega t}$, где $Q = U(i\omega R + C)^{-1}$. Соответственно ток определится формулой $i = \operatorname{Re} I e^{i\omega t}$, где

$$I = U/Z, \quad Z = R + 1/(i\omega C).$$

Такой ток (и заряд) установится в цепи за время порядка нескольких постоянных времени $\tau = RC$. Величина Z представляет собой, очевидно, импеданс рассматриваемой цепи.

Выяснив, как ведут себя порознь индуктивность и емкость при прохождении переменного тока, можно рассчитать любую цепь, в которой действует синусоидальное напряжение. При этом следует лишь иметь в виду, что основными элементами любой цепи являются активное сопротивление R , индуктивность L и емкость C и что им соответствуют импедансы R , $i\omega L$, $1/(i\omega C)$. Каждая цепь строится из этих элементов, которые могут соединяться последовательно или параллельно. Импедансы последовательно соединенных участков складываются ($Z = Z_1 + Z_2 + \dots$), так как через них проходит общий ток. В случае параллельно соединенных импедансов общим для них всех является напряжение, поэтому складываются обратные импедансы ($1/Z = 1/Z_1 + 1/Z_2 + \dots$). В простейшем случае контура, изображенного на рис. 5.5, $Z = R + i\omega L + 1/(i\omega C)$.

Амплитуда тока I как функция частоты представлена на рис. 5.6. Она обращается в нуль при $\omega = 0$ и $\omega \rightarrow \infty$ и достигает максимума при резонансной частоте $\omega_0 = 1/\sqrt{LC}$. Ширина этого максимума $\Delta\omega \sim R/L$, максимальное же значение тока $I_{\max} = U/R$; поэтому если активное сопротивление R очень мало, то максимум будет очень резким. О резонансе в такой цепи говорят как о *резонансе напряжений*.

Контур, изображенный на рис. 5.7, характеризуется импедансом

$$Z = \frac{(R_L + i\omega L)(R_C + 1/(i\omega C))}{R_L + R_C + i\omega L + 1/(i\omega C)}.$$

Если оба активных сопротивления малы ($R_{L,C} \ll \sqrt{L/C}$), то ими можно пренебречь в числителе (но не в знаменателе!) этого выражения; вводя $R = R_L + R_C$, получим

$$Z = \frac{L/C}{R + i\omega L + 1/(i\omega C)}.$$

Амплитуда напряжения $|U|$ как функция частоты (при заданном I) имеет в этом случае тот же вид, что амплитуда тока $|I|$ (при заданном U) для контура, изображенного на рис. 5.5 (см. рис. 5.6). Напряжение достигает максимума при той же резонансной частоте $\omega_0 = 1/\sqrt{LC}$; в этом случае говорят о резонансе токов. Легко убедиться, что если $R \rightarrow 0$, то при резонансе токов можно положить $I = 0$, $U = 0$; при этом токи в обеих ветвях цепи I_L , I_C останутся конечными. Отсюда следует, что при $R \rightarrow 0$ можно отключить источник внешнего напряжения, в контуре же, состоящем из емкости и индуктивности, будет циркулировать переменный ток частоты ω_0 . Такой контур называют *колебательным*, и мы перейдем теперь к его изучению.

5.7. Колебательный контур

Выясним прежде всего, как происходит разряд конденсатора не просто на сопротивление, а на катушку, обладающую индуктивностью. Иными словами, мы рассматриваем цепь, изображенную на рис. 5.5 или 5.7, в которой нет внешнего источника переменной э. д. с., но конденсатор обладает в начальный момент некоторым зарядом q_0 . Спрашивается, каков заряд конденсатора $q(t)$ в последующие моменты времени и каков ток $i(t)$ в этом контуре? В контуре действует э. д. с. самоиндукции $-Ldi/dt$, поэтому $-Ldi/dt = Ri + u_C$, где $u_C = q/C$ — напряжение на конденсаторе. Учитывая, что $\dot{i} = dq/dt$, получим следующее дифференциальное уравнение для определения заряда $q(t)$:

$$L \frac{d^2q}{dt^2} + R \frac{dq}{dt} + \frac{1}{C} q = 0.$$

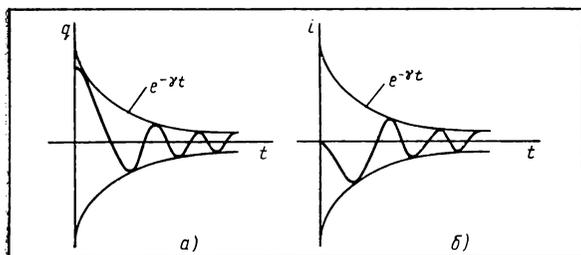
Это уравнение в отличие от неоднородных уравнений, с которыми мы имели дело выше, является однородным уравнением второго порядка. Решение уравнения следует искать в виде $q(t) = Ae^{st}$, где s и A — некоторые константы. Подстановка этого выражения в уравнение для q дает $Ls^2 + Rs + 1/C = 0$, откуда мы получаем два значения s :

$$s_{1,2} = -\frac{R}{2L} \pm \sqrt{\left(\frac{R}{2L}\right)^2 - \frac{1}{LC}}.$$

Таким образом, общее решение уравнения имеет вид

$$q(t) = A_1 e^{s_1 t} + A_2 e^{s_2 t},$$

где A_1 и A_2 — константы, определяемые начальными условиями. В начальный момент $t = 0$ заряд $q(t)$ имеет заданное значение q_0 , начальный же ток следует считать равным нулю, так как иначе в момент включения конденсатора в цепи возникла бы бесконечная э. д. с. самоиндукции. Итак,



5.8. Зависимости заряда (а) и тока (б) от времени в периодическом режиме

$$q(0) = q_0, \quad i(0) = \left(\frac{dq}{dt}\right)_{t=0} = 0.$$

Отсюда можно найти A_1 и A_2 .

При $R < 2\sqrt{LC}$ величины s_1 и s_2 будут комплексно-сопряженными. Записав их в виде

$$s_{1,2} = -\gamma \pm i\omega'_0, \quad \gamma = R/(2L), \quad \omega'_0 = \sqrt{\omega_0^2 - \gamma^2}, \quad \omega_0 = 1/\sqrt{LC},$$

получим следующие выражения для силы тока и заряда:

$$i(t) = -\left(1 + \frac{\gamma^2}{\omega_0^2}\right) \omega'_0 q_0 e^{-\gamma t} \sin \omega'_0 t;$$

$$q(t) = \sqrt{1 + \frac{\gamma^2}{\omega_0^2}} q_0 e^{-\gamma t} \cos(\omega'_0 t + \varphi),$$

причем $\operatorname{tg} \varphi = \gamma/\omega'_0$. Мы видим, что заряд и ток уменьшаются с течением времени — затухают, но затухание происходит не монотонно. Это затухание определяется экспонентой $\exp(-\gamma t)$, потому величину γ называют *коэффициентом затухания*. Она пропорциональна активному сопротивлению контура. Величину $\delta = 2\pi\gamma/\omega_0$ называют *логарифмическим декрементом колебаний*. Синусоидальные функции, входящие в выражения для $i(t)$ и $q(t)$, имеют период

$$T' = \frac{2\pi}{\omega_0} = 2\pi \sqrt{\frac{LC}{1 - R^2C/L}}.$$

Зависимость заряда и тока от времени графически изображена на рис. 5.8, а, б.

Если активное сопротивление R мало, то мал и логарифмический декремент и мы получим почти незатухающие колебания с постоянными амплитудами. В предельном случае $R \rightarrow 0$ частота колебаний совпадает с $\omega_0 = 1/\sqrt{LC}$ и $q(t) = q_0 \cos \omega_0 t$, $i(t) = -\omega_0 q_0 \sin \omega_0 t$. Таким образом, при $R \rightarrow 0$ в контуре без внешнего источника переменной э. д. с. возникают *собственные колебания*, частота которых совпадает с *резонансной частотой* ω_0 контура.

Найдем электрическую энергию, W_e , запасаемую в конденсаторе, и магнитную энергию W_m , запасаемую в индуктивности:

$$W_e = \frac{1}{2C} q^2, \quad W_m = \frac{1}{2} Li^2.$$

Полная электромагнитная энергия, находящаяся в конденсаторе и катушке, равна $W_e + W_m = q_0^2/(2C)$. Эта величина не меняется с течением времени, причем чем больше W_m , тем меньше W_e , и наоборот, так что колебания сопровождаются перекачкой энергии из конденсатора, т. е. электрического резервуара, в индуктивность, т. е. магнитный резервуар, и обратно из магнитного в электрический резервуар. В некотором смысле эта перекачка энергии аналогична переходу кинетической энергии в потенциальную энергию и обратно при обычных колебаниях маятника.

Наш результат о неизменности полной энергии $W = W_e + W_m$ связан с предположением, что $R \rightarrow 0$. При конечном активном сопротивлении колебания становятся затухающими и сумма электрической и магнитной энергий уменьшается со временем по экспоненциальному закону $e^{-2\gamma t}$.

Уменьшение энергии, т. е. диссипация, происходит из-за непрерывного выделения джоулевой теплоты в проводах катушки.

Ясно, что если бы можно было непрерывно компенсировать эту диссипацию энергии в контуре, то в нем происходили бы незатухающие колебания, несмотря на наличие активного сопротивления. Такая компенсация потерь энергии могла бы в принципе производиться с помощью обычного источника постоянной э. д. с., если бы он в нужные моменты времени, в такт с уменьшением амплитуды колебаний тока в контуре, подавал энергию в контур. Но частота колебаний в контуре очень велика. Поэтому должна быть очень большой частота подачи энергии от источника в контур, а в таком случае желательно, чтобы это происходило автоматически, причем не механическим, а электрическим путем.

Таким электрическим вентилем, автоматически открывающим и закрывающим источник энергии постоянного тока, соединенный с колебательным контуром, может служить специальная электронная лампа — так называемый *триод*. Простейшая электронная лампа — *диод* — содержит два электрода: раскаленный катод, испускающий электроны, и анод с положительным потенциалом относительно катода, воспринимающий электроны. Сила тока, текущего между анодом и катодом (его называют *анодным током*), зависит от разности потенциалов между электродами, причем зависимость эта нелинейна в отличие от обычных проводников, подчиняющихся закону Ома. В триоде между катодом и анодом помещают еще третий электрод — *управляющую сетку*, почти полностью пропускающую ток электронов от катода к аноду. Смысл работы сетки заключается в том, что на нее автоматически поступает информация о том, что делается в контуре, и сетка, реагируя на полученную информацию, в нужные моменты времени увеличивает анодный ток, а тем самым и ток в контуре. Для этого сетка связана с контуром, и наличие этой так называемой *обратной связи* позволяет непрерывно, автоматически и безынерционно подавать

в нужных количествах энергию от источника постоянной э. д. с. в колебательный контур, где она преобразуется в колебательную энергию переменного тока.

5.8. Электромагнитные флуктуации

Определяя токи в различных цепях, мы считали, что в них действуют вполне определенные э. д. с. — постоянные или переменные, но всегда являющиеся заданными функциями времени. Между тем в цепях неизбежно действуют и случайные э. д. с., точный вид которых принципиально не может быть установлен. Достаточно сказать, что посторонние токи наводят в каждой цепи э. д. с. индукции, которые могут обладать произвольными частотами, фазами и амплитудами. Точнее знать их мы не можем, и описывать их можно только статистически, используя понятие вероятности. Но если даже представить себе некоторую цепь электрически изолированной от окружающего мира, то и в этом случае в цепи будут течь случайные токи. Дело в том, что в системе зарядов токи отсутствуют только в том случае, если система находится в состоянии статистического, или теплового, равновесия. В действительности же в любой системе обязательно имеют место отклонения от статистического равновесия, а при таких отклонениях (их называют *тепловыми флуктуациями*) всегда возникают так называемые *флуктуационные электрические токи*. Поэтому

ток в цепи, в которой действует вполне определенная заданная э. д. с., отличается от значения, определяемого этой э. д. с., согласно обобщенному закону Ома: на это значение тока накладывается флуктуационный ток, возникающий из-за тепловых флуктуаций.

Покажем, как находятся флуктуационные токи. Начнем с цепи, состоящей из конденсаторов с емкостью C и резистора с сопротивлением R . В отсутствие внешней э. д. с. напряжение на конденсаторе u удовлетворяет уравнению

$$RC \frac{du}{dt} + u = 0,$$

решение которого имеет вид

$$u(t) = u(0) e^{-t/(RC)} \quad (t \geq 0),$$

где $u(0)$ — начальное значение u . Будем считать в этом решении $u(0)$ случайной величиной. Случайность процесса означает, что среднее значение $u(0)$, которое берется по всем возможным значениям $u(0)$ (или, как говорят, по ансамблю значений $u(0)$, не по времени!), равно нулю: $\langle u(0) \rangle = 0$. Поэтому и $\langle u(t) \rangle = 0$ (усреднение не затрагивает фиксированного множителя $\exp[-t/(RC)]$). Но среднее значение квадратичной величины $u(0)^2$ не равно нулю. Действительно, конденсатор мы можем рассматривать как гигантскую броуновскую частицу, находящуюся во внешней среде с температурой T . Поэтому можно считать, что в состоянии статистического равновесия энергия конденсатора (тепловая) должна быть равна $1/2 kT$, где k — постоянная Больцмана (на конденсаторе нет внешнего заряда!). Но энергия конденсатора равна половине произведения его емкости на квадрат напряжения, поэтому

$$1/2 C \langle u^2(0) \rangle = 1/2 kT,$$

откуда $\langle u^2(0) \rangle = kT/C$. Учитывая это соотношение, умножим $u(t)$ на $u(0)$ и усредним произведение по ансамблю $u(0)$:

$$\langle u(0) u(t) \rangle = \frac{kT}{C} e^{-t/(RC)} \quad (t \geq 0).$$

Начальный момент времени можно, очевидно, выбирать произвольно. Поэтому эту формулу перепишем в более общем виде:

$$\langle u(t_1) u(t_2) \rangle = \frac{kT}{C} e^{-|t_1 - t_2|/(RC)},$$

где t_1 и t_2 — два произвольных момента времени.

Величину $\langle u(t_1)u(t_2) \rangle$ называют *корреляционной функцией напряжения*. Она зависит от разности моментов времени и определяется в случае RC -контура емкостью и сопротивлением.

Выясним теперь, каков частотный состав, или, как говорят, *частотный спектр*, флуктуаций напряжения. Для этого нужно флуктуации разложить в интеграл Фурье. Но если раскладывать просто

$$u(t) = \int_{-\infty}^{\infty} u_{\omega} e^{i\omega t} d\omega,$$

то среднее значение каждой компоненты Фурье u_{ω} равно нулю: $\langle u_{\omega} \rangle = 0$. Поэтому в интеграл Фурье нужно раскладывать не линейную, а квадратичную функцию напряжения — корреляционную функцию $\langle u(0)u(t) \rangle$:

$$\langle u(0) u(t) \rangle = \int_{-\infty}^{\infty} (u^2)_{\omega} e^{i\omega t} d\omega.$$

Здесь символ $(u^2)_{\omega}$ служит для обозначения компоненты Фурье корреляционной функции напряжения. Используя известную формулу обращения

$$(u^2)_{\omega} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \langle u(0) u(t) \rangle e^{-i\omega t} dt$$

и подставляя сюда выражение для корреляционной функции, получим

$$(u^2)_{\omega} = \frac{kT}{2\pi C} \int_{-\infty}^{\infty} e^{-\frac{|t|}{RC} - i\omega t} dt.$$

Вычислив интеграл, найдем

$$(u^2)_{\omega} = \frac{kTR}{\pi(1 + \omega^2 R^2 C^2)}.$$

Это выражение и определяет частотную характеристику флуктуаций напряжения на конденсаторе.

Если умножить это выражение на $\omega^2 C^2$, то мы найдем компоненту Фурье $(i^2)_{\omega}$ корреляционной функции тока $\langle i(t_1)i(t_2) \rangle$:

$$(i^2)_{\omega} = \frac{kT}{\pi} \frac{R}{R^2 + 1/(\omega^2 C^2)}.$$

Вспомнив, что комплексное сопротивление цепи, состоящей из емкости и сопротивления, есть $Z = R + 1/(i\omega C)$, перепишем $(i^2)_{\omega}$ в виде

$$(i^2)_{\omega} = \frac{kT}{\pi} \frac{R}{|Z|^2} = \frac{kT}{\pi} \operatorname{Re} \frac{1}{Z}.$$

Найдем теперь корреляционную функцию флуктуационного тока в цепи, состоящей из индуктивности и сопротивления. В этом случае в отсутствие внешней э. д. с. сила тока $i(t)$ удовлетворяет уравнению $L(di/dt) + Ri = 0$, откуда

$$i(t) = i(0) e^{-(R/L)t}$$

В этом выражении мы будем рассматривать $i(0)$ как случайную величину, среднее от которой равно нулю: $\langle i(0) \rangle = 0$. Величина же $\langle i^2(0) \rangle$ в случае тепловых флуктуаций не равна нулю, а может быть просто связана с температурой T . Действительно, катушку самоиндукции можно рассматривать как единое целое, подобное частице, совершающей броуновское движение в среде с температурой T . Поэтому мы должны считать среднюю тепловую энергию катушки равной $\frac{1}{2} kT$. Но эту энергию можно отождествить со средней магнитной энергией $\frac{1}{2} L \langle i^2(0) \rangle$, имеющей флуктуационное происхождение, т. е. $\frac{1}{2} L \langle i^2(0) \rangle = kT/2$, откуда $\langle i^2(0) \rangle = kT/L$.

Умножим $i(t)$ на $i(0)$ и усредним произведение по возможным значениям $i(0)$. Получим

$$\langle i(0) i(t) \rangle = \frac{kT}{L} e^{-(R/L)t} \quad (t \geq 0).$$

Эта величина и представляет собой корреляционную функцию флуктуационных токов в цепи, состоящей из катушки и сопротивления. Поскольку начальный момент времени можно выбирать произвольно, корреляционную функцию можно записать в общем виде:

$$\langle i(t_1) i(t_2) \rangle = \frac{kT}{L} e^{-\frac{R}{L} |t_1 - t_2|}$$

Найдем теперь частотный спектр флуктуаций тока. Для этого нужно разложить корреляционную функцию $\langle i(0) i(t) \rangle$ в интеграл Фурье:

$$\langle i(0) i(t) \rangle = \int_{-\infty}^{\infty} (i^2)_{\omega} e^{i\omega t} d\omega.$$

Компонента Фурье $(i^2)_{\omega}$ корреляционной функции равна

$$(i^2)_{\omega} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \langle i(0) i(t) \rangle e^{-i\omega t} dt.$$

Подставляя сюда выражение для $\langle i(0) i(t) \rangle$, получим

$$(i^2)_{\omega} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{kT}{L} e^{-R|t|/L} e^{-i\omega t} dt,$$

откуда

$$(i^2)_{\omega} = \frac{kT}{\pi} \frac{R}{R^2 + \omega^2 L^2}.$$

Вспомним теперь, что импеданс рассматриваемой цепи определяется формулой $Z = R + i\omega L$. Поэтому $(i^2)_{\omega}$ можно представить в виде

$$(i^2)_{\omega} = \frac{kT}{\pi} \frac{R}{|Z|^2} = \frac{kT}{\pi} \operatorname{Re} \frac{1}{Z}.$$

Это выражение имеет ту же структуру, что и выражение для спектрального распределения флуктуаций тока в цепи, состоящей из C, R . Такой же формулой

определяется частотный спектр флуктуаций тока в самом общем случае для любой цепи:

$$(i^2)_\omega = \frac{kT}{\pi} \frac{R}{|Z|^2} = \frac{kT}{\pi} \frac{R}{R^2 + X^2},$$

где $Z = R + iX$ — импеданс цепи, R и $X \equiv X(\omega)$ — активное и реактивное сопротивление цепи. Эта формула носит название формулы Найквиста.

Если активное сопротивление очень мало, то и компонента Фурье $(i^2)_\omega$ мала почти для всех частот. Исключение составляют только те частоты, для которых реактивное сопротивление обращается в нуль: $X(\omega) = 0$. Это уравнение определяет резонансные частоты цепи. Например, в случае контура, состоящего из катушки и конденсатора, $X(\omega) = \omega L - 1/(\omega C)$ и X обращается в нуль при $\omega = (LC)^{-1/2}$. Для более сложных цепей резонансных частот может быть несколько и все они являются корнями уравнения $X(\omega) = 0$. Компонента Фурье корреляционной функции флуктуационных токов в случае малого сопротивления цепи R имеет резкие максимумы при резонансных частотах.

ОСНОВНЫЕ ФОРМУЛЫ

Закон электромагнитной индукции

$$\mathcal{E} = -\frac{d\Phi}{dt}, \quad \Phi = \int \mathbf{B} ds$$

Импедансы цепей (L, R) и (C, R)

$$Z = R + i\omega L, \quad Z = R + \frac{1}{i\omega C}$$

Сила тока в цепи

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

Частота колебательного контура

$$\omega_0 = 1/\sqrt{LC}$$

Коэффициент затухания

$$\gamma = R/(2L)$$

Корреляционная функция флуктуаций тока в цепи (L, R)

$$\langle i(t_1) i(t_2) \rangle = \frac{kT}{L} e^{-\frac{R}{L} |t_1 - t_2|}$$

Формула Найквиста

$$(i^2)_\omega = \frac{kT}{\pi} \frac{R}{|Z|^2}$$



Глава 6. ЗАКОНЫ ЭЛЕКТРОМАГНЕТИЗМА

6.1. Уравнения Максвелла

В предыдущей главе мы нашли ускоряющее вихревое электрическое поле в бетатроне. Но можно установить в самом общем виде связь между вихревым электрическим полем и вызы-

вающим его переменным магнитным полем. Обратимся для этого к закону электромагнитной индукции, согласно которому

$$\oint_L \mathbf{E} d\mathbf{l} = - \frac{d}{dt} \int_{\Sigma} \mathbf{B} ds.$$

Считая контур L неподвижным, перепишем это уравнение в виде

$$\oint_L \mathbf{E} d\mathbf{l} = - \int_{\Sigma} \frac{\partial \mathbf{B}}{\partial t} ds$$

и преобразуем, согласно формуле Стокса, линейный интеграл в поверхностный:

$$\oint_L \mathbf{E} d\mathbf{l} = \int_{\Sigma} \text{rot } \mathbf{E} ds$$

(напомним, что элемент контура $d\mathbf{l}$ и элемент поверхности ds образуют правовинтовую систему). Получим

$$\int_{\Sigma} \left(\text{rot } \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} \right) ds = 0,$$

а так как контур L и опирающаяся на него поверхность Σ могут быть произвольными, то отсюда следует, что

$$\text{rot } \mathbf{E} = - \frac{\partial \mathbf{B}}{\partial t}.$$

Это фундаментальное соотношение носит название **первого уравнения Максвелла**.

Подчеркнем еще раз, что электрическое поле, порождаемое переменным магнитным полем, является вихревым, а не потенциальным, поэтому его силовые линии замкнуты в отличие от силовых линий электростатического поля, которые всегда разомкнуты.

Из первого уравнения Максвелла легко вывести закон электромагнитной индукции — для этого стоит лишь проинтегрировать обе части этого уравнения по какой-либо поверхности Σ , опирающейся на некоторый проводящий контур L . Воспользовавшись далее формулой Стокса, можно преобразовать поверхностный интеграл от ротора электрического поля в циркуляцию этого поля вдоль контура L . Таким образом, мы получим э. д. с., действующую в контуре, которая выразится через временную производную магнитного потока, пронизывающего контур. На этом основании можно сказать, что первое уравнение Максвелла представляет собой закон электромагнитной индукции в дифференциальной форме. В действительности, однако, в уравнении Максвелла содержания больше, чем в законе электромагнитной индукции. Дело в том, что закон электромагнитной индукции неотъемлем от проводящего или (в более общей формулировке) материального контура. Уравнение же Максвелла никак не связано с материальным контуром — в него входят в чистом виде сами поля; иными словами это

полевое уравнение, связывающее поля в данном месте пространства и в данный момент времени.

Взяв дивергенцию от обеих частей уравнения Максвелла и замечая, что $\operatorname{div} \operatorname{rot} \mathbf{E} \equiv 0$, получим $(\partial/\partial t)\operatorname{div} \mathbf{B} = 0$, откуда следует, что дивергенция магнитной индукции не зависит от времени. Такой вывод вытекает из первого уравнения Максвелла, но этот вывод не означает, что

$$\operatorname{div} \mathbf{B} = 0.$$

Последнее соотношение выполняется, но

оно не вытекает из первого уравнения Максвелла, а выражает отдельный закон природы — не закон электромагнитной индукции, а закон отсутствия в природе магнитных зарядов.

Напомним в этой связи, что магнитные силовые линии не имеют ни начала, ни конца.

Переменное магнитное поле создает вокруг себя вихревое электрическое поле. Но электрическое и магнитное поля во многом аналогичны. Поэтому естественен вопрос: не создает ли и переменное электрическое поле поля магнитного? То, что такая связь между переменными полями должна существовать, показывает элементарный анализ процесса разряда конденсатора. Действительно, соединим обкладки заряженного конденсатора проводом. Тогда по проводу пойдет разрядный ток, который будет создавать магнитное поле. Силовые линии поля охватывают проводник. Если выбрать какой-либо контур L , охватывающий провод, то циркуляция магнитной индукции вдоль L должна определяться, согласно закону полного тока, силой тока, проходящего через произвольную поверхность, опирающуюся на L . Но если взять две поверхности — поверхность Σ_1 , пересекающую провод, и поверхность Σ_2 , проходящую в пространстве между обкладками конденсатора, — то через поверхность Σ_1 протекает разрядный ток I , через поверхность же Σ_2 никакого тока не течет. Налицо, таким образом, нарушение закона полного тока. Дело здесь, очевидно, в том, что конденсатор нарушает замкнутость тока проводимости. Но магнитное поле незамкнутого разрядного тока тем не менее несомненно существует и, более того, циркуляция индукции \mathbf{B} вдоль контура L равна $\mu_0 I$. Ток I есть интеграл по поверхности Σ_1 от плотности тока \mathbf{j} , т. е.

$$\oint_L \mathbf{B} d\mathbf{l} = \mu_0 \int_{\Sigma_1} \mathbf{j} ds.$$

Но циркуляция не может зависеть от вида поверхности, важно лишь, чтобы последняя опиралась на L . Поэтому должно существовать соотношение

$$\oint_L \mathbf{B} d\mathbf{l} = \mu_0 \int_{\Sigma_2} \mathbf{X} ds,$$

где Σ_2 — произвольная поверхность, проходящая между обкладками

конденсатора, и \mathbf{X} — некоторая неизвестная векторная величина. Для ее определения мы имеем уравнение

$$\int_{\Sigma_1} \mathbf{j} ds = \int_{\Sigma_2} \mathbf{X} ds,$$

которое годится для всех поверхностей Σ , пересекающих провод, включая поверхность обкладки конденсатора; поверхность же Σ_2 , лежащая внутри конденсатора, может сколь угодно близко подходить к обкладке конденсатора.

Возьмем теперь в качестве Σ_1 поверхность обкладки конденсатора. На ней плотность тока j , очевидно, совпадает с $d\sigma/dt$, где σ — поверхностная плотность заряда, или в векторной форме $\mathbf{j} = \mathbf{n}(d\sigma/dt)$, где \mathbf{n} — единичный вектор нормали к обкладке. Плотность заряда σ можно выразить через напряженность поля \mathbf{E} вблизи обкладки: $\sigma = \epsilon_0 \mathbf{E} \mathbf{n}$, поэтому на поверхности обкладки

$$\mathbf{j} = \epsilon_0 \dot{\mathbf{E}}.$$

Учтем теперь, что поверхность Σ_2 может изнутри подходить сколь угодно близко к поверхности обкладки и что поле внутри плоского конденсатора всюду одинаково. Поэтому, по крайней мере в случае плоского конденсатора, мы можем считать, что

$$\mathbf{X} = \epsilon_0 \dot{\mathbf{E}}$$

и, следовательно,

$$\oint_L \mathbf{B} dl = \mu_0 \epsilon_0 \int_{\Sigma_2} \dot{\mathbf{E}} ds.$$

Преобразуем теперь, согласно формуле Стокса, циркуляцию магнитной индукции в интеграл по поверхности Σ_2 . Мы получим тогда соотношение

$$\int_{\Sigma_2} \text{rot } \mathbf{B} ds = \mu_0 \epsilon_0 \int_{\Sigma_2} \frac{\partial \mathbf{E}}{\partial t} ds,$$

а так как поверхность Σ_2 внутри конденсатора произвольна, то должно выполняться соотношение

$$\text{rot } \mathbf{B} = \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}.$$

В это соотношение не входят ни емкость конденсатора, ни сопротивление провода, соединяющего его обкладки ни разрядный ток, а входят лишь локально магнитная индукция и электрическое поле. Поэтому возникает вопрос: не является ли это соотношение универсальным и справедливым во всех тех случаях, когда имеется переменное электрическое поле и нет движущихся электрических зарядов? Максвелл впервые предположил, что дело обстоит именно таким образом, т. е. что переменное электрическое поле всегда создает вокруг

себя магнитное поле в соответствии с написанным соотношением. Иными словами, переменное электрическое поле с точки зрения создания магнитного поля эквивалентно некоторому току, плотность которого

$$\mathbf{j}_c = \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}.$$

Этот ток носит название *тока смещения*, а \mathbf{j}_c называют *плотностью тока смещения*.

Если имеются также и движущиеся заряды, которым соответствует суммарная плотность тока \mathbf{j}_t (она включает в себя как плотность тока проводимости, так и плотность молекулярных токов), то должен быть справедлив закон полного тока

$$\text{rot } \mathbf{B} = \mu_0 (\mathbf{j}_t + \mathbf{j}_c),$$

в который входит сумма плотности потока заряда движущихся частиц \mathbf{j}_t и плотности тока смещения \mathbf{j}_c . Смысл этого соотношения, которое называют вторым уравнением Максвелла, заключается в том, что *магнитное поле порождается как движущимися зарядами, так и переменным электрическим полем*.

На основании своих уравнений (сформулированных им в 1860—1865 гг.) Максвелл в 1865 г. предсказал существование электромагнитных волн, распространяющихся со скоростью света, и предположил, что сам видимый свет — это электромагнитная волна определенной длины.

6.2. Свободное электромагнитное поле в вакууме

Итак, переменное электрическое поле создает вокруг себя (подобно переменному току) переменное магнитное поле. В свою очередь, переменное магнитное поле создает вокруг себя электрическое поле, которое также переменное. Поэтому возникает вопрос: не могут ли переменные электрическое и магнитное поля, поддерживая друг друга, существовать в вакууме без зарядов и токов? Математически вопрос сводится к тому, имеют ли уравнения Максвелла для полей в вакууме

$$\begin{aligned} \text{rot } \mathbf{E} &= -\partial \mathbf{B} / \partial t; & \text{div } \mathbf{B} &= 0; \\ \text{rot } \mathbf{B} &= \mu_0 \epsilon_0 \partial \mathbf{E} / \partial t; & \text{div } \mathbf{E} &= 0 \end{aligned}$$

решения, отличающиеся от тривиального $\mathbf{E} = \mathbf{B} = 0$? Ответ гласит, что такие решения существуют и их бесконечно много.

Чтобы убедиться в этом, исключим из уравнений одно из полей, например индукцию \mathbf{B} . Применим для этого операцию rot к обеим частям первого уравнения:

$$\text{rot rot } \mathbf{E} = -\frac{\partial}{\partial t} \text{rot } \mathbf{B}.$$

Используя далее второе уравнение, получим $\text{rot rot } \mathbf{E} = -\mu_0 \epsilon_0 \partial^2 \mathbf{E} / \partial t^2$.

Учтем, наконец, что $\text{rot rot } \mathbf{E} = \text{grad div } \mathbf{E} - \Delta \mathbf{E}$. Так как в вакууме $\text{div } \mathbf{E} = 0$, то мы получим окончательно

$$\Delta \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0,$$

где $c^2 = (\epsilon_0 \mu_0)^{-1}$. Легко проверить, что такому же уравнению — его называют в о л н о в ы м — удовлетворяет в вакууме и индукция \mathbf{B} .

Покажем теперь, что волновое уравнение имеет бесчисленное множество нетривиальных решений. Рассмотрим для простоты случай, когда поля зависят только от одной декартовой координаты, скажем x . Тогда все отличные от нуля компоненты полей \mathbf{E} , \mathbf{B} (обозначим их общей буквой f) удовлетворяют одному и тому же одномерному волновому уравнению

$$\frac{\partial^2 f}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 f}{\partial t^2} = 0.$$

Решение этого уравнения имеет вид

$$f(x, t) = F_1(x - ct) + F_2(x + ct),$$

где F_1 и F_2 — произвольные функции. При этом функция $F_1(x - ct)$ описывает волну, распространяющуюся со скоростью c вдоль положительной оси x , а функция $F_2(x + ct)$ — волну, распространяющуюся с той же скоростью в противоположном направлении. Эти волны называют *электромагнитными волнами*. Величина c равна скорости света в вакууме. Поэтому мы приходим к заключению, что скорость электромагнитных волн совпадает со скоростью света в вакууме. Более того,

сам свет представляет собой электромагнитную волну, и как световая волна, так и любая электромагнитная волна в вакууме обладает свойством поперечности.

Это немедленно вытекает из уравнений $\text{div } \mathbf{E} = 0$, $\text{div } \mathbf{B} = 0$. Действительно, подставляя в них $\mathbf{E} = \mathbf{E}(x \mp ct)$, $\mathbf{B} = \mathbf{B}(x \mp ct)$, мы получим $\partial E_x / \partial x = \partial B_x / \partial x = 0$, откуда $E_x = \text{const}$, $B_x = \text{const}$. Но мы рассматриваем здесь переменные, а не статистические поля. Поэтому следует считать $E_x = B_x = 0$, т. е. переменные поля не имеют составляющих вдоль направления распространения волны и являются, следовательно, поперечными. Отличными от нуля в них могут быть только y - и z -составляющие.

Мы показали, что одномерное волновое уравнение имеет бесчисленное множество нетривиальных решений. Этим же свойством обладает и трехмерное волновое уравнение, и все эти решения имеют вид волн, распространяющихся в разных направлениях со скоростью света c .

Существование нетривиальных решений уравнений Максвелла в отсутствие токов и зарядов имеет фундаментальное значение. Оно означает, что

переменные поля могут существовать в вакууме без зарядов и токов. Это значит, что электромагнитное поле следует рассматривать как физическую реальность, а не как атрибут зарядов.

Для того чтобы окончательно убедиться в этом, нужно еще доказать, что электромагнитное поле, как и вещество, обладает определенной энергией, импульсом и моментом. Сейчас мы убедимся в этом, после чего можно будет с полным основанием утверждать, что электромагнитное поле действительно материально.

С энергией электростатического поля мы уже встречались: рассматривая процесс зарядки конденсатора, мы показали, что затрачиваемая при этом работа против сил электростатического расталкивания зарядов не исчезает и что можно считать, что она запасается в виде энергии электростатического поля W_e , плотность которой равна $\epsilon_0 E^2/2$. Однако с равным правом мы могли считать, что величина W_e представляет собой энергию зарядов, а не энергию поля. Поэтому мы рассмотрим теперь вопрос об энергии поля — электрического и магнитного — с более общей точки зрения, исходя из баланса энергии в системе, состоящей из поля и заряженных частиц.

Если в электромагнитном поле \mathbf{E} , \mathbf{B} движутся частицы, то их движение определяется уравнением

$$\frac{d}{dt} \mathbf{p} = q (\mathbf{E} + [\mathbf{v}\mathbf{B}]),$$

где q — заряд частицы, \mathbf{v} и \mathbf{p} — ее скорость и импульс. Само поле мы будем теперь считать создаваемым зарядами. Это значит, что в уравнениях Максвелла

$$\text{rot } \mathbf{E} = - \frac{\partial \mathbf{B}}{\partial t}; \quad \text{div } \mathbf{B} = 0;$$

$$\text{rot } \mathbf{B} = \mu_0 \left(\mathbf{j} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right); \quad \epsilon_0 \text{div } \mathbf{E} = \rho$$

плотности тока и заряда определяются рассматриваемыми зарядами, т. е.

$$\rho = \sum q_a n_a, \quad \mathbf{j} = \sum q_a n_a \mathbf{v}_a,$$

где n_a — плотность частиц с зарядом q_a и скоростью \mathbf{v}_a , и суммирование распространяется на различные сорта a частиц, находящихся в единичном объеме. Вопрос теперь состоит в следующем: как изменится кинетическая энергия частиц

$$K = \sum_{(V)} mv^2/2,$$

находящихся в некотором объеме V , если учесть воздействие поля на частицы?

Изменение кинетической энергии одной частицы за 1 с составляет

$$\frac{d}{dt} \frac{1}{2} mv^2 = \mathbf{v} \frac{d}{dt} \mathbf{p},$$

а так как сила, действующая на частицу со стороны магнитного поля, направлена перпендикулярно скорости частицы, то согласно уравнению движения частицы

$$\frac{d}{dt} \frac{1}{2} m v^2 = q E v.$$

Поэтому

$$\frac{d}{dt} K = \sum_{(V)} q E v.$$

Можно перейти в этом выражении от суммирования по всем частицам к интегрированию по объему V , занимаемому частицами, с последующим суммированием только по различным сортам частиц:

$$\frac{d}{dt} K = \sum_a \int_V q_a n_a v_a E dV.$$

Вспоминая определение плотности тока \mathbf{j} , можно переписать это выражение в виде

$$\frac{d}{dt} K = \int_V \mathbf{j} E dV.$$

Но согласно второму уравнению Максвелла

$$\mathbf{j} E = \left(\frac{1}{\mu_0} \operatorname{rot} \mathbf{B} - \epsilon_0 \dot{\mathbf{E}} \right) E.$$

Замечая, что согласно первому уравнению Максвелла $\dot{\mathbf{B}} = -\mathbf{B} \operatorname{rot} \mathbf{E}$, получим

$$\mathbf{j} E = -\epsilon_0 \mathbf{E} \frac{\partial E}{\partial t} - \frac{1}{\mu_0} \mathbf{B} \frac{\partial B}{\partial t} - \frac{1}{\mu_0} (\mathbf{B} \operatorname{rot} \mathbf{E} - \mathbf{E} \operatorname{rot} \mathbf{B}).$$

Это выражение нужно теперь проинтегрировать по объему V . В силу теоремы Гаусса

$$\int_V \operatorname{div} [\mathbf{E} \mathbf{B}] dV = \int_{\Sigma} [\mathbf{E} \mathbf{B}] ds,$$

где Σ — поверхность, ограничивающая объем V , и ds — элемент поверхности Σ , направленный вдоль внешней нормали к ней. В результате получим

$$\frac{d}{dt} K \equiv \int_V \mathbf{j} E dV = -\frac{d}{dt} \int_V \frac{1}{2} (\mathbf{E} \mathbf{D} + \mathbf{B} \mathbf{H}) dV - \int_{\Sigma} [\mathbf{E} \mathbf{H}] ds,$$

где $\mathbf{D} = \epsilon_0 \mathbf{E}$, $\mathbf{H} = \mathbf{B} / \mu_0$.

Проанализируем теперь это выражение. Предположим сперва, что на границе объема V поля обращаются в нуль. Тогда

$$\frac{d}{dt} K = -\frac{d}{dt} W, \quad W \equiv \frac{1}{2} \int (\mathbf{E} \mathbf{D} + \mathbf{B} \mathbf{H}) dV.$$

Мы видим, что в этом случае $dK = -dW$, т. е. приращение K равно убыли W , так что сумма K и W не изменяется с течением времени: $K + W = \text{const}$.

Так как K представляет собой кинетическую энергию частиц, то полученное соотношение, определяющее баланс энергии, не что иное, как закон сохранения энергии. Но мы рассматриваем только частицы и электромагнитное поле, которые взаимодействуют между собой и образуют замкнутую систему. Поэтому $K + W$ есть полная энергия нашей системы, а величина W , зависящая только от \mathbf{E} , \mathbf{B} , \mathbf{D} , \mathbf{H} , есть не что иное, как энергия электромагнитного поля. Она представляет собой интеграл от плотности электромагнитной энергии w по объему, занимаемому полем:

$$W = \int_V w dV, \quad w = \frac{1}{2} (\mathbf{E}\mathbf{D} + \mathbf{B}\mathbf{H}),$$

Плотность энергии равна сумме плотностей электрической $w_e = 1/2 \epsilon_0 E^2$ и магнитной $w_m = (2\mu_0)^{-1} B^2$ энергий. Эти величины пропорциональны квадратам полей. Напомним, что с плотностями электрической и магнитной энергии мы уже сталкивались в § 1.7, 4.5.

Обратимся теперь снова к общему выражению для $(d/dt)K$, содержащему поверхностный интеграл:

$$J = \int_{\Sigma} \mathbf{S} ds, \quad \mathbf{S} = [\mathbf{E}\mathbf{H}].$$

С учетом этого интеграла баланс энергии рассматриваемой нами динамической системы — заряженные частицы плюс электромагнитное поле — выглядит следующим образом:

$$\frac{d}{dt} \left(K + \int_V w dV \right) = - \int_{\Sigma} \mathbf{S} ds.$$

Слева здесь стоит производная по времени от полной энергии системы, т. е. суммы кинетической энергии заряженных частиц и энергии электромагнитного поля. Эта производная не равна нулю, т. е. полная энергия системы не сохраняется. Ясно, что такая ситуация может возникнуть только потому, что энергия покидает или, наоборот, входит в объем V . Движение энергии происходит через поверхность Σ , ограничивающую объем V , и поэтому в баланс энергии входит именно поверхностный интеграл J .

Таким образом, мы должны интерпретировать величину J как поток энергии, уходящей из объема V , т. е. количество электромагнитной энергии, покидающей объем V за 1 с. Соответственно величину \mathbf{S} нужно интерпретировать как *плотность потока электромагнитной энергии*, т. е. как энергию, проходящую за 1 с через единичную площадку, ориентированную перпендикулярно вектору \mathbf{S} . Этот вектор носит название *вектора Пойнтинга*.

Кроме энергии электромагнитное поле обладает еще *импульсом* (количеством движения). Плотность импульса поля $\mathbf{g} = \mathbf{S}/c^2$ отличается

от вектора Пойнтинга только множителем $1/c^2$, так что импульс электромагнитного поля, находящегося в объеме V , есть

$$\mathbf{G} = \int_V \mathbf{g} dV = c^{-2} \int_V [\mathbf{E}\mathbf{H}] dV.$$

Поэтому полный импульс \mathbf{P} системы, состоящей из частиц и электромагнитного поля, определяется формулой

$$\mathbf{P} = \sum_{(V)} \mathbf{p} + \frac{1}{c^2} \int_V [\mathbf{E}\mathbf{H}] dV,$$

где первое слагаемое представляет собой импульс частиц в объеме V . Полный импульс в конечном объеме V не сохраняется, так как электромагнитное поле проходит через границу объема и переносит с собой часть импульса.

Наконец, поле обладает еще и определенным *моментом импульса* (моментом количества движения), плотность которого

$$[\mathbf{r}\mathbf{g}] = c^{-2} [\mathbf{r} [\mathbf{E}\mathbf{H}]],$$

так что момент \mathbf{L} поля в объеме V составляет

$$\mathbf{L} = c^{-2} \int [\mathbf{r} [\mathbf{E}\mathbf{H}]] dV.$$

6.3. Усредненные микроскопические поля

Уравнения Максвелла, сформулированные выше, связывают электрическое и магнитное поля в вакууме между собой и с заданными токами и зарядами. Их можно рассматривать как законы движения электромагнитного поля в вакууме. Теперь наша задача будет состоять в том, чтобы установить аналогичные законы для электромагнитного поля в веществе — поля в общем случае переменного и неоднородного. При этом следует учитывать, что вещество построено из атомов, которые, в свою очередь, состоят из заряженных частиц — электронов и ядер. Поэтому мы должны исходить из представления об электромагнитном поле в вакууме, в котором находятся, пребывая в постоянном движении, частицы вещества, а так как они заряжены, то, с одной стороны, они испытывают влияние поля в вакууме, а с другой — модифицируют это поле.

Таким образом, возникает задача о самосогласованном взаимодействии электромагнитного поля в вакууме и электронов и ядер вещества. Задача эта была впервые сформулирована Лоренцем в конце прошлого столетия, и он же указал путь ее принципиального решения.

Сложность задачи связана как с колоссальным числом электронов и ядер в любом макроскопическом участке вещества, так и с невероятно запутанным характером их движения. Но именно эта сложность в конечном счете и упрощает задачу, ибо она позволяет использовать для решения задачи статистический метод. Дело в том, что истин-

ные поля в вакууме, модифицированные заряженными частицами вещества, — будем называть эти поля *микроскопическими* — характеризуются чрезвычайной нерегулярностью. Нерегулярность эта — и в пространстве, и во времени — обусловлена атомной структурой вещества. Если взять, например, отдельный ион, то электрическое поле, порождаемое им, меняет свое направление на расстоянии порядка размеров иона, т. е. 10^{-10} м. Ясно, далее, что и электрическое и магнитное поля, создаваемые атомными электронами, претерпевают существенные изменения в течение характерного атомного времени, составляющего примерно 10^{-16} с (такое время соответствует прохождению атомного расстояния $\sim 10^{-10}$ м с характерной для электронов атома скоростью порядка 10^6 м/с). Но все эти нерегулярности полей для многих задач несущественны. Поэтому нужно их исключить из рассмотрения, а это значит, что интересны не сами микроскопические поля, а усредненные их значения, не содержащие уже мелких нерегулярностей, или, образно выражаясь, «шероховатостей» полей, обусловленных атомной структурой вещества.

Итак, мы должны ввести усредненные, или, как их называют, *макроскопические, поля*. Усреднение должно производиться как по пространственным координатам, так и по времени. Именно: если $f(\mathbf{r}, t)$ — некоторая микроскопическая величина, содержащая атомные нерегулярности, то соответствующая усредненная макроскопическая величина определяется как

$$\langle f(\mathbf{r}, t) \rangle = \frac{1}{\Delta V \Delta t} \int f(\mathbf{r} + \delta \mathbf{r}, t + \delta t) d(\delta \mathbf{r}) d(\delta t),$$

где интегрирование производится по некоторому пространственному объему ΔV возле точки \mathbf{r} и некоторому временному интервалу Δt возле момента времени t . Объем усреднения ΔV должен быть большим по сравнению с объемом атома a^3 , но малым по сравнению с той областью пространства L^3 , где уже существенно меняется сама усредненная величина (она ведь остается функцией координат и времени!). Аналогичным образом время усреднения Δt должно быть большим по сравнению с характерным атомным временем $\tau_a = a/v_a$, но малым по сравнению с временем T существенного временного изменения самой усредненной величины (v_a — средняя скорость электронов в атоме). Если соблюдены условия

$$a \ll (\Delta V)^{1/3} \ll L, \quad \tau_a \ll \Delta t \ll T,$$

то усредненная величина $\langle f(\mathbf{r}, t) \rangle$ не содержит атомных «шероховатостей», а является плавной функцией координат и времени. Именно такие усредненные, или макроскопические, величины мы и будем рассматривать.

Введем теперь усредненные поля — электрическое $\mathbf{E}(\mathbf{r}, t)$ и магнитное $\mathbf{B}(\mathbf{r}, t)$ (вектор \mathbf{B} , как уже указывалось, называется, по исторической традиции, магнитной индукцией). С микроскопическими полями, которые мы обозначим $\mathbf{e}(\mathbf{r}, t)$ и $\mathbf{b}(\mathbf{r}, t)$, они связаны соотношениями

$$\mathbf{E}(\mathbf{r}, t) = \langle \mathbf{e}(\mathbf{r}, t) \rangle; \quad \mathbf{B}(\mathbf{r}, t) = \langle \mathbf{b}(\mathbf{r}, t) \rangle.$$

Задача заключается в том, чтобы выяснить, каким закономерностям подчиняются макроскопические поля в среде $\mathbf{E}(\mathbf{r}, t)$ и $\mathbf{B}(\mathbf{r}, t)$. При этом исходным пунктом должны быть закономерности для микроскопических полей $\mathbf{e}(\mathbf{r}, t)$ и $\mathbf{b}(\mathbf{r}, t)$ в вакууме. Эти закономерности выражаются уравнениями электромагнитного поля в вакууме. Первая группа этих уравнений записывается просто:

$$\text{rot } \mathbf{e}(\mathbf{r}, t) = -\frac{\partial}{\partial t} \mathbf{b}(\mathbf{r}, t), \quad \text{div } \mathbf{b}(\mathbf{r}, t) = 0.$$

Чтобы написать вторую пару уравнений поля, нужно иметь в виду, что вещество может содержать заряженные частицы различных сортов, и если плотность частиц, обладающих зарядом q_a , равна n_a , то микроскопическая плотность электрического заряда равна $\sum q_a n_a$, где суммирование распространяется на все сорта a заряженных частиц. Если далее \mathbf{v}_a — скорость частиц сорта a , то микроскопическая плотность тока $\sum q_a n_a \mathbf{v}_a$. Поэтому вторая пара уравнений поля может быть записана в виде

$$\begin{aligned} \text{rot } \mathbf{b}(\mathbf{r}, t) &= \mu_0 \left(\sum q_a n_a \mathbf{v}_a + \epsilon_0 \frac{\partial}{\partial t} \mathbf{e}(\mathbf{r}, t) \right); \\ \epsilon_0 \text{div } \mathbf{e}(\mathbf{r}, t) &= \sum q_a n_a. \end{aligned}$$

Обе эти пары уравнений представляют собой, по существу, уравнения Максвелла для полей в вакууме, в котором «плавают» заряженные частицы вещества. Но часто их называют уравнениями Максвелла — Лоренца, чтобы подчеркнуть тот подход, который был указан Лоренцом для вывода уравнений поля в средах на основе электронной картины материи.

Мы должны теперь усреднить уравнения Максвелла—Лоренца. При этом, очевидно, усреднение никак не связано ни с операциями rot и div , т. е. дифференцированием по координатам, ни с дифференцированием по времени. Поэтому усреднение дает

$$\begin{aligned} \text{rot } \langle \mathbf{e}(\mathbf{r}, t) \rangle &= -\frac{\partial}{\partial t} \langle \mathbf{b}(\mathbf{r}, t) \rangle, \quad \text{div } \langle \mathbf{b}(\mathbf{r}, t) \rangle = 0; \\ \mu_0^{-1} \text{rot } \langle \mathbf{b}(\mathbf{r}, t) \rangle &= \sum \langle q_a n_a \mathbf{v}_a \rangle + \epsilon_0 \frac{\partial}{\partial t} \langle \mathbf{e}(\mathbf{r}, t) \rangle; \\ \epsilon_0 \text{div } \langle \mathbf{e}(\mathbf{r}, t) \rangle &= \sum \langle q_a n_a \rangle, \end{aligned}$$

или, так как $\langle \mathbf{e} \rangle = \mathbf{E}$, $\langle \mathbf{b} \rangle = \mathbf{B}$, усредненная первая пара уравнений поля имеет вид

$$\text{rot } \mathbf{E}(\mathbf{r}, t) = -\frac{\partial}{\partial t} \mathbf{B}(\mathbf{r}, t); \quad \text{div } \mathbf{B}(\mathbf{r}, t) = 0,$$

а усредненная вторая пара уравнений — вид

$$\mu_0^{-1} \text{rot } \mathbf{B}(\mathbf{r}, t) = \sum \langle q_a n_a \mathbf{v}_a \rangle + \epsilon_0 \frac{\partial}{\partial t} \mathbf{E}(\mathbf{r}, t);$$

$$\epsilon_0 \operatorname{div} \mathbf{B}(\mathbf{r}, t) = \sum \langle q_a n_a \rangle.$$

Первая пара имеет простой физический смысл: первое из них выражает для макроскопических полей закон электромагнитной индукции в дифференциальной форме, а второе — отсутствие в природе магнитных зарядов.

Рассмотрим теперь вторую пару усредненных уравнений и выясним прежде всего смысл величины $\sum \langle q_a n_a \mathbf{v} \rangle$, представляющей собой усредненную плотность микроскопического тока. Чтобы определить ее, нужно знать структуру рассматриваемого вещества. Но при общем феноменологическом описании мы должны различать плотность тока, обусловленного свободными зарядами, и плотность тока, обусловленного связанными зарядами. Обозначив обе эти величины соответственно \mathbf{j} и \mathbf{j}_b и предполагая для общности, что оба типа зарядов имеются в теле, мы можем написать

$$\sum \langle q_a n_a \mathbf{v}_a \rangle = \mathbf{j} + \mathbf{j}_b.$$

В действительности, конечно, первое слагаемое присутствует только в проводниках, и в этом случае для достаточно медленных полей и неподвижных проводников плотность тока \mathbf{j} (ее называют *плотностью тока проводимости*) пропорциональна напряженности макроскопического электрического поля \mathbf{E} , т. е. $\mathbf{j} = \sigma \mathbf{E}$, где σ — удельная электрическая проводимость вещества. Но, несмотря на это, мы включили слагаемое \mathbf{j} в сумму $\sum \langle q_a n_a \mathbf{v}_a \rangle$, чтобы получить в дальнейшем общие соотношения, пригодные как для проводников, так и непроводников.

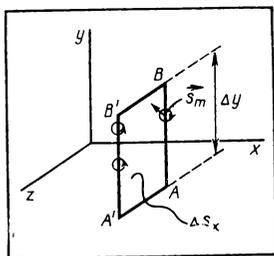
Рассмотрим далее плотность тока, обусловленного связанными зарядами. Этот ток можно разделить на две части — ток, обусловленный смещением связанных зарядов, или, иначе, их поляризацией, и ток, обусловленный вращением заряженных частиц, в первую очередь вращением электронов в атомах. Соответствующие плотности токов мы обозначим \mathbf{j}_p и \mathbf{j}_J , так что $\mathbf{j}_b = \mathbf{j}_p + \mathbf{j}_J$. Первая из этих величин — плотность поляризационного тока \mathbf{j}_p — может быть выражена через вектор поляризации \mathbf{P} , который определяется как усредненная сумма дипольных моментов $\mathbf{d}_a = q_a \mathbf{r}_a$ связанных частиц, содержащихся в единичном объеме (\mathbf{r}_a — радиус-вектор, определяющий положение связанного заряда a), т. е.

$$\mathbf{P} = \sum_{(V=1)} \langle \mathbf{d}_a \rangle.$$

Взяв производную от \mathbf{P} по времени, мы и получим плотность тока:

$$\mathbf{j}_p = \frac{\partial}{\partial t} \mathbf{P}.$$

Обратимся к плотности тока \mathbf{j}_J , обусловленного связанными вращающимися зарядами. Каждый такой заряд эквивалентен, очевидно, некоторому замкнутому молекулярному току, и вопрос, сводится к тому, как связать величину \mathbf{j}_J с этими молекулярными токами. Определим с этой целью суммарный ток $\mathbf{j}_J \Delta s$, проходящий через некоторую площадку Δs , например через площадку Δs_x , ориен-



6.1.

Плотность тока, обусловленного вращающимися связанными зарядами

тированную перпендикулярно оси x (рис. 6.1). Так как молекулярные токи замкнуты, то, очевидно, суммарный ток через площадку Δs_x определяется только теми индивидуальными молекулярными токами, которые пересекают площадку Δs_x не два, а только один раз; иными словами, существенны только те молекулярные токи, которые расположены на границе площадки Δs_x . Если взять границу AB , то число таких токов равно $\Delta y s_y n$, где n — число молекулярных токов (т. е. атомов) в единичном объеме, s_y — проекция площади s_m , обтекаемой молекулярным током, на плоскость xz , и Δy — длина AB . Умножив эту величину на силу отдельного молекулярного тока i_m , найдем общий молекулярный ток возле границы AB площадки Δs_x . Замечая, что величина $\mathbf{m} = i_m s_m$ представляет собой магнитный момент отдельного молекулярного тока, мы можем записать общий молекулярный ток вблизи границы AB в виде $\Delta y (n\mathbf{m})_y$. Аналогичного типа выражения справедливы для общих молекулярных токов вблизи границ $A'A$, $B'B$, $A'B'$. Все эти выражения нужно сложить, в результате чего получим циркуляцию вектора $n\mathbf{m}$ вдоль контура $ABA'B'$, ограничивающего площадку Δs . Наконец, так как \mathbf{j}_J по самому смыслу — макроскопическая величина, то величину $n\mathbf{m}$ нужно еще усреднить, чтобы в нее не входили атомные нерегулярности. Замечая, что общий молекулярный ток через площадку Δs равен $\mathbf{j}_J \Delta s$, получим соотношение

$$\mathbf{j}_J \Delta s = \oint_{ABA'B'} \langle n\mathbf{m} \rangle d\mathbf{l}.$$

где $d\mathbf{l}$ — элемент контура $ABA'B'$, ограничивающего площадку. Циркуляцию можно выразить согласно формуле Стокса через ротор вектора $n\mathbf{m}$, а так как площадка макроскопически мала, то

$$\mathbf{j}_J \Delta s = \text{rot} \langle n\mathbf{m} \rangle \Delta s,$$

откуда

$$\mathbf{j}_J = \text{rot} \mathbf{J}, \quad \mathbf{J} = \langle n\mathbf{m} \rangle.$$

Величину \mathbf{J} называют намагниченностью (вектором намагниченя). Ее можно записать также в виде

$$\mathbf{J} = \sum_{(V=1)} \langle \mathbf{m}_a \rangle,$$

где \mathbf{m}_a — магнитный момент a -го вращающегося заряда, а суммирование производится по всем связанным зарядам, находящимся в единичном объеме.

Итак, плотность макроскопического тока, связанного с вращающимися связанными зарядами, равна ротору намагниченности \mathbf{J} .

Отсюда легко заключить, что полный ток $\int_{\Sigma} \mathbf{j}_J ds$, обусловленный вращающимися зарядами, через произвольное сечение Σ тела равен нулю. Действительно, так как ток \mathbf{j}_J отличен от нуля только там, где $\mathbf{J} \neq 0$, а вне тела $\mathbf{J} = 0$, то

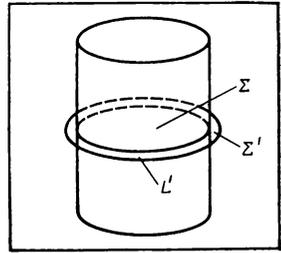
$$\int_{\Sigma} \mathbf{j}_J ds = \int_{\Sigma'} \mathbf{j}_J ds,$$

где Σ' — поверхность, включающая в себя поверхность Σ и лежащая частично вне тела (рис. 6.2). Но по формуле Стокса,

$$\int_{\Sigma'} \mathbf{j}_J ds = \int_{\Sigma'} \text{rot } \mathbf{J} ds = \oint_{L'} \mathbf{J} dl,$$

где L' — контур, ограничивающий Σ' , а так как на нем $\mathbf{J} = 0$, то мы и приходим к равенству

$$\int_{\Sigma} \mathbf{j}_J ds = 0.$$



6.2.
К определению полного тока, обусловленного вращающимися зарядами

6.4. Уравнения Максвелла для макроскопических полей в веществе

Теперь мы можем вернуться к второй группе усредненных уравнений. Первое из них, представляющее собой дифференциальную форму закона полного тока в среде, можно записать в виде

$$\mu_0^{-1} \text{rot } \mathbf{B} = \mathbf{j} + \mathbf{j}_P + \mathbf{j}_J + \epsilon_0 \frac{\partial}{\partial t} \mathbf{E}$$

или, учитывая, что $\mathbf{j}_P = (\partial/\partial t)\mathbf{P}$, $\mathbf{j}_J = \text{rot } \mathbf{J}$, в виде

$$\text{rot } \mathbf{H} = \mathbf{j} + \frac{\partial}{\partial t} \mathbf{D},$$

где

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}, \quad \mathbf{H} = \mu_0^{-1} \mathbf{B} - \mathbf{J}.$$

Вектор \mathbf{D} называют *электрической индукцией*, а вектор \mathbf{H} — *напряженностью магнитного поля* (напомним, что \mathbf{B} — магнитная индукция).

Итак, ротор напряженности магнитного поля в среде определяется плотностью тока проводимости \mathbf{j} и производной по времени от электрической индукции \mathbf{D} . Если ввести плотность \mathbf{j}_c тока смещения в среде $\mathbf{j}_c = (\partial/\partial t)\mathbf{D}$, то полученному соотношению можно придать вид

$$\text{rot } \mathbf{H} = \mathbf{j} + \mathbf{j}_c.$$

Это соотношение представляет собой окончательную формулировку закона полного тока для материальных сред в дифференциальной форме.

Обратимся, наконец, к последнему уравнению второй группы усредненных уравнений, связывающему дивергенцию усредненного электрического поля \mathbf{E} с усредненной микроскопической плотностью заряда $\Sigma \langle q_a n_a \rangle$. Эту величину можно представить в виде суммы плотности ρ свободных зарядов и плотности ρ_b связанных зарядов: $\Sigma \langle q_a n_a \rangle = \rho + \rho_b$. Производные по времени от плотностей зарядов ρ и ρ_b можно просто выразить через плотности токов \mathbf{j} и \mathbf{j}_b . Для этого следует воспользоваться законом сохранения заряда, согласно которому отношение к времени изменения общего заряда Q внутри некоторого объема V равно суммарной силе тока I , втекающего в этот объем через окружающую его поверхность Σ . Так как

$$Q = \int_V \rho dV, \quad I = - \int_{\Sigma} \mathbf{j} ds,$$

где ρ — объемная плотность заряда и \mathbf{j} — вектор плотности тока (ds имеет направление внешней к Σ нормали), то

$$\frac{d}{dt} \int_V \rho dV = - \int_{\Sigma} \mathbf{j} ds.$$

Учитывая, что $\int_{\Sigma} \mathbf{j} ds = \int_V \operatorname{div} \mathbf{j} dV$, получим

$$\frac{\partial}{\partial t} \rho = - \operatorname{div} \mathbf{j}.$$

Это соотношение (его называют уравнением непрерывности), представляющее собой закон сохранения заряда в дифференциальной форме, написано для свободных зарядов. Но такое же соотношение справедливо и для связанных зарядов, т. е. $\frac{\partial}{\partial t} \rho_b = - \operatorname{div} \mathbf{j}_b$. Подставляя сюда $\mathbf{j}_b = \mathbf{j}_P + \mathbf{j}_J$, получим

$$\frac{\partial}{\partial t} \rho_b = - \operatorname{div} \mathbf{j}_P - \operatorname{div} \mathbf{j}_J,$$

а так как $\mathbf{j}_P = \frac{\partial}{\partial t} \mathbf{P}$, $\mathbf{j}_J = \operatorname{rot} \mathbf{J}$ и $\operatorname{div} \operatorname{rot} \mathbf{J} = 0$, то

$$\frac{\partial}{\partial t} \rho_b = - \frac{\partial}{\partial t} \operatorname{div} \mathbf{P}.$$

Отсюда следует, что усредненная плотность связанных зарядов

$$\rho_b = - \operatorname{div} \mathbf{P}$$

(поскольку плотность ρ_b равна, очевидно, нулю при $\mathbf{P} = 0$).

Проинтегрировав ρ_b по объему V , найдем суммарный усредненный связанный заряд Q_b в этом объеме:

$$Q_b = \int_V \rho_b dV = - \int_V \operatorname{div} \mathbf{P} dV.$$

Согласно теореме Гаусса, этот интеграл равен

$$Q_b = - \int_{\Sigma} \mathbf{P} ds,$$

т. е. связанный заряд определяется потоком вектора поляризации через поверхность Σ , окружающую объем, и не зависит от намагниченности. Итак,

$$\Sigma \langle q_a n_a \rangle = \rho - \operatorname{div} \mathbf{P}.$$

Поэтому последнее уравнение второй группы усредненных уравнений может быть записано в виде

$$\epsilon_0 \operatorname{div} \mathbf{E} = \rho - \operatorname{div} \mathbf{P},$$

или

$$\operatorname{div} \mathbf{D} = \rho,$$

где \mathbf{D} , как и ранее, вектор электрической индукции. Таким образом, мы пришли к следующим четырем уравнениям для усредненных полей:

$$\operatorname{rot} \mathbf{E} = - \frac{\partial}{\partial t} \mathbf{B}, \quad \operatorname{div} \mathbf{B} = 0,$$

$$\operatorname{rot} \mathbf{H} = \mathbf{j} + \frac{\partial}{\partial t} \mathbf{D}, \quad \operatorname{div} \mathbf{D} = \rho.$$

Эти уравнения называют уравнениями Максвелла для материальных сред.

Мы видим, что в них входит не два, как в случае вакуума, а четыре поля: \mathbf{E} , \mathbf{B} и \mathbf{D} , \mathbf{H} . Из них непосредственный физический смысл имеют поля \mathbf{E} и \mathbf{B} . Они представляют собой усредненные микроскопические электрическое и магнитное поля. Поля \mathbf{D} и \mathbf{H} являются вспомогательными, и, чтобы найти их, нужно кроме \mathbf{E} и \mathbf{B} знать еще поляризованность \mathbf{P} и намагниченность \mathbf{J} :

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}, \quad \mathbf{H} = \mu_0^{-1} \mathbf{B} - \mathbf{J}.$$

Величины \mathbf{P} и \mathbf{J} зависят от \mathbf{E} и \mathbf{B} , но нахождение этих зависимостей представляет собой, вообще говоря, сложную задачу, которая может быть решена только на основе атомной динамики, т. е. квантовой механики.

Так как зависимость величин \mathbf{P} и \mathbf{J} от \mathbf{E} и \mathbf{B} не одинакова для различных веществ, то на границе раздела двух разных тел поля \mathbf{E} , \mathbf{D} , \mathbf{B} , \mathbf{H} , вообще говоря, не непрерывны, а испытывают скачки при переходе от одного тела к другому. При этом выполняются вполне определенные граничные условия. Именно:

всегда непрерывны тангенциальные составляющие напряженностей \mathbf{E} , \mathbf{H} и нормальные составляющие индукций \mathbf{D} , \mathbf{B} :

$$\mathbf{E}_{1t} = \mathbf{E}_{2t}, \quad \mathbf{H}_{1t} = \mathbf{H}_{2t}, \quad D_{1n} = D_{2n}, \quad B_{1n} = B_{2n},$$

где индексы 1, 2 служат для обозначения двух соприкасающихся тел, а индексы t, n обозначают тангенциальную и нормальную (к границе раздела) составляющие векторов.

Соотношения эти вытекают из уравнений Максвелла, если записать их в интегральной форме. Обратимся сперва к одному из уравнений, содержащих дивергенцию индукции, например $\operatorname{div} \mathbf{D} = \rho$, и проинтегрируем его по объему малого плоского цилиндра, одно из оснований которого находится в среде 1, а второе — в среде 2. Преобразуя согласно теореме Гаусса объемный интеграл в поверхностный, имеем $\int_{\Sigma} \mathbf{D} ds = Q$, где Q — сторонний заряд внутри цилиндра, Σ — поверхность цилиндра и ds — ориентированный элемент поверхности. Устремим теперь высоту цилиндра к нулю. При этом $Q \rightarrow 0$ и на одном из оснований цилиндра $ds \rightarrow nds$, а на другом $ds \rightarrow -nds$, где \mathbf{n} — вектор нормали к границе и ds — элемент площади. Поэтому $\int_{\Delta s} D_{1n} ds = \int_{\Delta s} D_{2n} ds$ (Δs — основание цилиндра), или в силу произвольности в выборе цилиндра $D_{1n} = D_{2n}$. Аналогично получается граничное условие $B_{1n} = B_{2n}$.

Обратимся теперь к одному из уравнений, содержащих ротор поля, например

$$\operatorname{rot} \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{j},$$

и проинтегрируем его по поверхности Δs , ограниченной малым прямоугольником L , две стороны которого параллельны границе и лежат по обе ее стороны. Преобразуя, согласно теореме Стокса, поверхностный интеграл к интегралу по контуру, имеем

$$\oint_L \mathbf{H} dr = \int_{\Delta s} \left(\frac{\partial \mathbf{D}}{\partial t} + \mathbf{j} \right) ds.$$

Устремим теперь нормальные к поверхности стороны прямоугольника к нулю. При этом поверхностный интеграл обратится в нуль, а циркуляция поля в силу произвольности выбора прямоугольника L даст $\mathbf{H}_{1t} = \mathbf{H}_{2t}$. Аналогично получается граничное условие $\mathbf{E}_{1t} = \mathbf{E}_{2t}$.

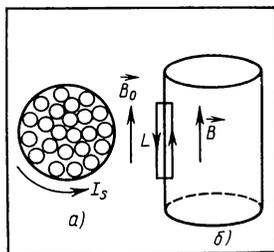
Нормальные составляющие полей и тангенциальные составляющие индукций на границе двух сред испытывают, вообще говоря, скачок. В гл. 3 мы уже сталкивались со скачком E_n на границе диэлектрика с металлом и видели, что

скачок связан с возникновением поверхностного заряда.

Этот результат сохраняет силу для любой границы раздела. В самом деле, согласно определению макроскопического поля, $\epsilon_0 \operatorname{div} \mathbf{E} = \Sigma \langle q_a n_a \rangle$ (сумма берется по зарядам среды). Применив к этому уравнению процедуру, использованную выше при выводе граничного условия для D_n , получим $\epsilon_0 (E_{1n} - E_{2n}) = \rho_s$, где ρ_s — поверхностная плотность заряда. Покажем, что

скачок тангенциальной составляющей магнитной индукции выражается через плотность поверхностных токов на границе раздела сред.

Рассмотрим для этого тело в форме цилиндра, помещенное в магнитное поле \mathbf{B}_0 , параллельное оси цилиндра. Картина молекулярных токов в сечении цилиндра схематически изображена на рис. 6.3, а. Мы видим, что молекулярные токи приводят к существованию поверхностного тока I_s на границе цилиндра. Ясно, что ток этот пропорционален длине l цилиндра. Рассмотрим теперь замкнутый контур L в виде прямоугольника, две длинные стороны которого, параллельные образующей цилиндра, находятся в непосредственной близости к границе цилиндра, причем одна сторона лежит внутри цилиндра, а другая — вне его (рис. 6.3, б). Циркуляция магнитной индукции вдоль контура L равна $(B - B_0)l$, где B — индукция магнитного поля внутри цилиндра. С другой стороны, согласно закону полного тока, ее следует приравнять $\mu_0 I$, где I — ток, пронизывающий контур L . Но этот ток представляет собой поверхностный ток I_s . Поэтому,



6.3. К выводу соотношения между скачком \mathbf{B}_t и плотностью поверхностных токов на границе двух сред

$$\mu_0^{-1} (B - B_0) l = I_s,$$

а так как поверхностный ток I_s пропорционален длине цилиндра l , то можно ввести плотность поверхностного тока \mathbf{j}_s , т. е. поверхностный ток, приходящийся на единичную длину цилиндра. В результате получим соотношение

$$\mu_0^{-1} (\mathbf{B}_t - \mathbf{B}_0 t) = \mathbf{j}_s,$$

где мы снабдили величины \mathbf{B} и \mathbf{B}_0 индексом t , чтобы подчеркнуть, что речь идет о тангенциальных составляющих вектора магнитной индукции. В общем случае как угодно граничащих друг с другом двух сред это соотношение имеет вид

$$\mu_0^{-1} [\mathbf{n}_{12} (\mathbf{B}_2 - \mathbf{B}_1)] = \mathbf{j}_s,$$

где \mathbf{n}_{12} — единичный вектор нормали, направленный из среды 1 в среду 2.

Заметим в заключение этого параграфа, что все приведенные граничные условия, так же как и сами уравнения Максвелла для материальных сред, справедливы в том случае, если пространственный и временной интервалы существенного изменения макроскопических полей велики по сравнению с соответствующими атомными величинами.

6.5. Баланс энергии в среде при наличии электромагнитного поля

В § 6.2 мы дали определение энергии электромагнитного поля в вакууме, основываясь на рассмотрении энергетического баланса в динамической системе, состоящей из поля в вакууме и заряженных частиц. Представим себе теперь, что вакуум заменяется материальной средой, в которой существует некоторое электромагнитное поле, и поставим вопрос о балансе энергии в этом случае.

Будем для определенности предполагать, что имеется источник сторонней э. д. с. мощностью

$$P^{(e)} = \int_V \mathbf{j} \mathbf{E}^{(e)} dV,$$

где $\mathbf{j} = \sigma(\mathbf{E} + \mathbf{E}^{(e)})$ — плотность тока проводимости, $\mathbf{E}^{(e)}$ — напряженность электрического поля сторонних источников и σ — проводимость среды. Вопрос тогда сводится к тому, как расходуется эта мощность. Чтобы получить ответ на этот вопрос, положим в формуле для $P^{(e)}$

$$\mathbf{j} = \text{rot } \mathbf{H} - \dot{\mathbf{D}}, \quad \mathbf{E}^{(e)} = \mathbf{j}/\sigma - \mathbf{E}$$

и учтем, что $\text{rot } \mathbf{E} = -\dot{\mathbf{B}}$. Замечая еще, что

$$\mathbf{H} \text{ rot } \mathbf{E} - \mathbf{E} \text{ rot } \mathbf{H} = \text{div } [\mathbf{E}\mathbf{H}],$$

получим

$$P^{(e)} = \int_V \frac{j^2}{\sigma} dV + \int_V \left(\mathbf{E} \frac{\partial}{\partial t} \mathbf{D} + \mathbf{H} \frac{\partial}{\partial t} \mathbf{B} \right) dV + \int_V \text{div } \mathbf{S} dV,$$

где $\mathbf{S} = [\mathbf{E}\mathbf{H}]$. Последний интеграл можно согласно теореме Гаусса преобразовать в поверхностный интеграл от \mathbf{S} , так что

$$P^{(e)} = \int_V \frac{1}{\sigma} j^2 dV + \int_V \left(\mathbf{E} \frac{\partial}{\partial t} \mathbf{D} + \mathbf{H} \frac{\partial}{\partial t} \mathbf{B} \right) dV + \int_{\Sigma} \mathbf{S} ds,$$

где ds — ориентированный вдоль внешней нормали элемент поверхности Σ , ограничивающей объем V .

Это соотношение и определяет баланс энергии в среде. Первое слагаемое справа представляет собой джоулеву теплоту, выделяемую в среде и обусловленную существованием проводимости, а третье слагаемое — плотность потока электромагнитной энергии, выходящего из объема V . В § 6.2 мы назвали вектор \mathbf{S} , определяющий плотность потока, вектором Пойнтинга.

Наконец, второе слагаемое справа представляет собой мощность, которую нужно затратить в объеме V , чтобы электрическая и магнитная индукции изменялись со скоростями $(\partial/\partial t)\mathbf{D}$ и $(\partial/\partial t)\mathbf{B}$. Можно сказать и так, что работа dA , которую нужно совершить над единичным объемом среды, чтобы изменить электрическую индукцию на $d\mathbf{D}$ и магнитную индукцию на $d\mathbf{B}$, есть

$$dA = \mathbf{E}d\mathbf{D} + \mathbf{H}d\mathbf{B}.$$

Для полей в вакууме, когда $\mathbf{D} = \epsilon_0\mathbf{E}$ и $\mathbf{B} = \mu_0\mathbf{H}$, работа dA представляет собой полный дифференциал функции

$$\omega_0 = (\epsilon_0 E^2 + \mu_0 H^2)/2,$$

которую можно интерпретировать как плотность электромагнитной энергии в вакууме.

В случае материальной среды работа dA , вообще говоря, не есть полный дифференциал какой-либо функции состояния среды, подобно тому, как работа при изменении объема тела $dA = p dV$, где p — давление, также не есть полный дифференциал функции состояния тела. Последнее, как известно, связано с тем, что давление p зависит не только от объема V , но и от других величин, например от температуры, и поэтому работа зависит от процесса, при котором она производится. Подобным же образом работа, которую нужно совершить при изменении индукций \mathbf{D} и \mathbf{B} на $d\mathbf{D}$ и $d\mathbf{B}$, зависит от процесса, при котором изменяются поля, в частности от теплового режима. Только в том случае, если напряженность электрического поля \mathbf{E} является функцией \mathbf{D} и только \mathbf{D} и аналогично напряженность магнитного поля \mathbf{H} зависит только от \mathbf{B} , работа dA является полным дифференциалом функции

$$\omega = \int (\mathbf{E}d\mathbf{D} + \mathbf{H}d\mathbf{B}),$$

где интегрирование по \mathbf{D} и \mathbf{B} производится от нулевых до заданных значений \mathbf{D} и \mathbf{B} . В этом случае величину ω можно интерпретировать как *объемную плотность внутренней энергии среды*, причем она представляет собой сумму плотностей собственно электромагнитной энергии и энергии материальной среды, обусловленной электромагнитным полем.

При линейных связях между \mathbf{E} и \mathbf{D} и \mathbf{H} и \mathbf{B}

$$\omega = 1/2 \mathbf{E}\mathbf{D} + 1/2 \mathbf{H}\mathbf{B}.$$

Первое слагаемое здесь носит название *объемной плотности электрической энергии в среде*, а второе — *объемной плотности магнитной энергии в среде*, хотя, как было сказано выше, в ω входят как собственно энергия электромагнитного поля, так и энергия вещества, обусловленная электромагнитным полем.

Для переменных и пространственно неоднородных полей dA не представляет собой полного дифференциала какой-либо функции состояния полей, так как $\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P}$, $\mu_0^{-1}\mathbf{B} = \mathbf{H} + \mathbf{J}$, а величины \mathbf{P} и \mathbf{J} зависят, как правило, не только от \mathbf{E} и \mathbf{H} , но и от производных \mathbf{E} и \mathbf{H} по времени и координатам.

Но и для постоянных и однородных полей dA будет полным дифференциалом только в исключительных случаях, а именно для неполярных диэлектриков и диамагнетиков (да и то приближенно), так как только для этих веществ \mathbf{P} и \mathbf{J} практически не зависят от температуры, в иных же случаях \mathbf{P} и \mathbf{J} зависят от температуры и соответственно работа dA зависит от процесса изменения полей.

Если, однако, к работе dA , производимой над системой, прибавить теплоту $dQ = TdS$, приобретаемую системой (S — энтропия системы, T — температура), то мы получим полный дифференциал внутренней энергии системы. Поэтому если над средой не производится никакой другой работы, кроме работы, связанной с изменением индукций \mathbf{D} и \mathbf{B} , то изменение объемной плотности внутренней энергии среды

$$dw = Tds + (\mathbf{E}d\mathbf{D} + \mathbf{H}d\mathbf{B}),$$

где s — объемная плотность энтропии среды; среда предполагается при этом термодинамически равновесной и, следовательно, поля считаются не зависящими от времени. Величина dw является в этом случае полным дифференциалом функции состояния среды $w \equiv w(\mathbf{D}, \mathbf{B}, s)$, представляющей собой объемную плотность внутренней энергии среды — энергии, содержащей как собственно энергию электромагнитного поля (постоянного), так и часть энергии среды, обусловленной электромагнитным полем.

Перейдем от объемной плотности энергии w к объемной плотности свободной энергии $F = w - Ts$. Ясно, что

$$dF = -sdT + (\mathbf{E}d\mathbf{D} + \mathbf{H}d\mathbf{B}).$$

Эта формула показывает, что для изотермического процесса второе слагаемое (т. е. dA) является полным дифференциалом объемной плотности свободной энергии среды $F = F(\mathbf{D}, \mathbf{B}, T)$. В частности, при линейных связях между \mathbf{D} и \mathbf{E} , \mathbf{B} и \mathbf{H}

$$F = 1/2 \mathbf{E}\mathbf{D} + 1/2 \mathbf{H}\mathbf{B}.$$

Здесь коэффициенты пропорциональности между индукциями и напряженностями полей могут как угодно зависеть от температуры, если же такой зависимости нет, то $F = w$. Величину F называют *объемной плотностью свободной энергии полей в среде* (подчеркнем еще раз, что в F входит как объемная плотность собственно энергии электромагнитного поля, так и часть объемной плотности свободной энергии среды, связанная с полями).

Во всех приведенных формулах \mathbf{E} и \mathbf{B} представляют собой суперпозиции полей, создаваемых внешними источниками и порождаемых молекулярными токами и связанными зарядами среды. Но часто приходится относить работу не к изменению индукций \mathbf{D} и \mathbf{B} , а к изменению только внешних полей. В этом случае также легко установить энергетический баланс в среде. Воспользуемся с этой целью тем, что энергия электрического дипольного момента \mathbf{d} во внешнем электрическом поле $\mathbf{E}^{(e)}$ равна $-\mathbf{d}\mathbf{E}^{(e)}$, а энергия магнитного момента \mathbf{m} во внешнем магнитном поле $\mathbf{H}^{(e)}$ равна $-\mu_0\mathbf{m}\mathbf{H}^{(e)}$. Поэтому над моментами \mathbf{d} и \mathbf{m} при изменении внешних полей на $d\mathbf{E}^{(e)}$ и $d\mathbf{H}^{(e)}$ совершается работа

$$da = -\mathbf{d} \cdot d\mathbf{E}^{(e)} - \mu_0\mathbf{m}d\mathbf{H}^{(e)}.$$

Соответственно этому работа внешних полей над электрическим и магнитным моментами, отнесенная к единице объема среды, составит

$$dA = - PdE^{(e)} - \mu_0 JdH^{(e)} .$$

Прибавив к этому выражению теплоту Tds , получим полный дифференциал объемной плотности энергии среды как функции переменных $T, E^{(e)}, H^{(e)}$. Отсюда далее следует, что дифференциал объемной плотности свободной энергии $F = w - Ts$ среды при наличии внешних полей есть

$$dF = - sdT - PdE^{(e)} - \mu_0 JdH^{(e)} .$$

6.6. Системы электромагнитных единиц

При изучении электрических и магнитных явлений кроме СИ пользуются и *гауссовой системой единиц*. Гауссова система представляется естественной при общем исследовании электромагнитных явлений, так как она отражает главное свойство электромагнитного поля — его единство и относительность разделения на электрическую и магнитную компоненты (в этой системе величины E, H, D, B имеют одинаковую размерность). Однако для практического использования она мало удобна, так как значения электромагнитных величин, с которыми практически приходится встречаться, т. е. значения зарядов, сил токов, напряженностей полей, емкостей, сопротивлений, индуктивностей, таковы, что в гауссовой системе они выражаются либо очень малыми, либо очень большими числами. Например, силе тока в 1 А соответствует примерно тысяча миллионов гауссовых единиц силы тока или разности потенциалов в один вольт соответствует $1/300$ гауссовой единицы потенциала. СИ не обладает этим недостатком и поэтому широко используется, особенно в технике.

Мы остановимся в этом параграфе на общем вопросе о выборе систем единиц и разясним сперва, какие величины следует выбирать в качестве основных.

По традиции основными и обычно считают три величины — длину L , время T и массу M , для которых и выбираются основные единицы. В системе СГС ими являются сантиметр, секунда и грамм, а в СИ — метр, секунда и килограмм. Единицы для других величин — производные единицы — устанавливают с помощью уравнений, выражающих те или иные физические законы из основных для данной системы единиц, и их размерности выражают через размерности L, T, M .

Но законов существует много, и поэтому производные единицы можно выбирать, вообще говоря, разными способами и, более того, число основных единиц не обязательно должно быть равно трем. Обратимся, например, к закону всемирного тяготения

$$f = G \frac{m_1 m_2}{r^2} ,$$

где m_1 и m_2 — массы двух материальных точек, находящихся на расстоянии r друг от друга, и G — гравитационная постоянная. В систе-

ме СГС $G = 6,67 \cdot 10^{-8} \text{ см}^3 \cdot \text{с}^{-2} \cdot \text{г}^{-1}$. Но так как она универсальна, то ее можно считать безразмерной и выбирать равной единице. Тогда закон всемирного тяготения примет вид

$$f = \frac{m_1 m_2}{r^2}.$$

Эта формулировка должна быть, естественно, согласована с вторым законом Ньютона

$$m dv/dt = f.$$

Приравнивая размерности силы f в обеих формулах, мы получим $MLT^{-2} = M^2L^{-2}$, откуда

$$M = L^3T^{-2},$$

т. е. размерность массы получилась производной, а основными единицами стали только две единицы — для длины и времени.

Но можно пойти еще дальше, воспользовавшись универсальностью скорости света c в вакууме. По этой причине можно считать ее безразмерной и выбрать равной единице: $c = 1$. Мы получим тогда совпадающие размерности и совпадающие единицы для длины и времени: $L = T$. Соединив требования $G = 1$ и $c = 1$, мы получим систему единиц с одной лишь основной единицей для длины. Размерность массы совпадает в этой системе с размерностью длины: $M = L$.

В квантовой физике вводится новая универсальная постоянная — постоянная Планка \hbar , в системе СГС $\hbar \simeq 10^{-27} \text{ см}^2 \cdot \text{с}^{-1} \cdot \text{г}$. Но в силу универсальности \hbar эту постоянную можно считать безразмерной и выбрать ее равной единице: $\hbar = 1$. Присоединив сюда требование $c = 1$, мы получим систему единиц с одной основной единицей, так как теперь

$$T = L, \quad M = L^{-1},$$

т. е. в качестве основной единицы может быть взята единица длины. Масса выражается в обратных длинах. Заряд безразмерен, и, например заряд электрона $e = 1/\sqrt{137}$.

Обратимся теперь к вопросу о выборе единиц для электромагнитных величин. При его решении необходимо исходить из наиболее общих формулировок основных законов электромагнетизма. Начнем с закона Кулона, который следует (как уже указывалось) записать в виде

$$F = k_1 \frac{q_1 q_2}{r^2},$$

где $q_{1,2}$ — заряды и k_1 — некоторый коэффициент, вообще говоря, размерный, при выборе которого необходимо лишь соблюдать условие

$$[k_1][q]^2 = L^3T^{-2}M.$$

Напряженность E электрического поля должна определяться как сила, действующая на единичный заряд, так что заряд q на расстоянии r создает напряженность

$$E = k_1(q/r^2).$$

Рассмотрим далее взаимодействие между токами и запишем формулу Ампера для силы взаимодействия токов I_1 и I_2 , протекающих в двух параллельных, длинных и тонких проводах, в следующем общем виде (об этом уже шла речь в § 4.1):

$$F_A = 2k_2 \frac{I_1 I_2}{d} l,$$

где l — длина проводов, d — расстояние между ними и k_2 — некоторый, вообще говоря, размерный коэффициент, аналогичный k_1 . Сила тока определяется при этом как производная от заряда по времени: $I = (d/dt)q$. Поэтому могут быть сравнены размерности k_1 и k_2 . Ясно, что

$$[k_1/k_2] = L^2 T^{-2},$$

т. е. k_1/k_2 имеет размерность квадрата скорости. Силы F и F_A могут быть тщательно измерены, поэтому величина k_1/k_2 может быть найдена экспериментально; она оказывается равной квадрату скорости света c в вакууме:

$$k_1/k_2 = c^2.$$

Введем теперь магнитную индукцию B . Она должна определяться как величина, пропорциональная (но не равная!) силе действующей согласно формуле Ампера на ток единичной силы:

$$B = 2k_2 \alpha I/d,$$

где α — некоторый, вообще говоря, размерный коэффициент.

Вектор индукции \mathbf{B} должен входить также в закон электромагнитной индукции Фарадея:

$$\oint_L \mathbf{E} d\mathbf{l} = -k_3 \frac{d}{dt} \int_S \mathbf{B} ds,$$

где k_3 — также некоторый произвольный, но размерный коэффициент.

Итак, в нашем распоряжении имеется четыре коэффициента k_1 , k_2 , α , k_3 и, выбирая их по-разному, можно получить различные системы единиц. Однако эти коэффициенты нельзя выбирать совершенно произвольно, так как они не независимы. Действительно, уравнения Максвелла имеют теперь вид

$$\text{rot } \mathbf{E} = -k_3 \frac{\partial}{\partial t} \mathbf{B}; \quad \text{div } \mathbf{E} = 4\pi k_1 \rho;$$

$$\text{rot } \mathbf{B} = 4\pi k_2 \alpha \mathbf{j} + \frac{k_2 \alpha}{k_1} \frac{\partial}{\partial t} \mathbf{E}; \quad \text{div } \mathbf{B} = 0.$$

Для полей в вакууме отсюда следует волновое уравнение

$$\Delta \mathbf{B} - k_3 \frac{k_2 \alpha}{k_1} \frac{\partial^2}{\partial t^2} \mathbf{B} = 0.$$

Но коэффициент при $(\partial^2/\partial t^2)\mathbf{B}$ должен быть равен $1/c^2$ независимо от систем единиц, т. е. $k_1/(k_2k_3\alpha) = c^2$, а так как $k_1/k_2 = c^2$, то $\alpha = 1/k_3$.

Таким образом, свободно можно распоряжаться только двумя из четырех коэффициентов k_1, k_2, k_3, α , но выбирать их следует так, чтобы получаемые системы единиц были удобными либо для практического, либо для теоретического использования. Важную роль играет также простота и даже красота написания уравнений Максвелла, о которых Больцман говорил «Уж не Боги ли начертали эти знаки!». Следует также иметь в виду, что число основных единиц (как механических, так и электромагнитных) не должно быть обязательно равно трем.

К настоящему времени было предложено пять систем единиц — электростатическая, электромагнитная, гауссова, система Хевисайда — Лоренца и рационализированная система МКС (или СИ). В первых четырех основными величинами считаются длина, время, масса с размерностями L, T, M (и единицами сантиметр, секунда, грамм); в пятой к основным величинам добавляется заряд с размерностью Q или ток с размерностью I (единицами являются метр, секунда, килограмм, кулон или ампер). Значения констант k_1, k_2, α, k_3 в этих системах приведены в следующей таблице:

Система	k_1	k_2	α	k_3
Электростатическая	1	$c^2 (L^{-2}T^2)$	1	1
Электромагнитная	$c^2 (L^2T^{-2})$	1	1	1
Гауссова	1	$c^{-2} (L^{-2}T^2)$	$c(LT)^{-1}$	$c^{-1} (L^{-1}T)$
Хевисайда — Лоренца	$1/(4\pi)$	$1/(4\pi c^2) (L^{-2}T^2)$	$c (LT)^{-1}$	$c^{-1} (L^{-1}T)$
СИ	$\frac{1}{4\pi\epsilon_0} = 9 \cdot 10^9$ $(L^3T^{-2} MQ^{-2})$	$\frac{\mu_0}{4\pi} = 10^{-7}$ (LMQ^{-2})	1	1

В скобках указаны размерности величин. В СИ полагают $k_1 = 1/(4\pi\epsilon_0)$, $k_2 = \mu_0/(4\pi)$. Поэтому ϵ_0 и μ_0 имеют размерности $L^3T^{-2}MQ^{-2}$ и LMQ^{-2} (о размерностях k_1 и k_2 говорилось выше), т. е.

$$\frac{1}{4\pi\epsilon_0} = 9 \cdot 10^9 \frac{\text{Н} \cdot \text{м}^2}{\text{Кл}^2}, \quad \frac{\mu_0}{4\pi} = 10^{-7} \frac{\text{м} \cdot \text{кг}}{\text{Кл}^2}.$$

Выпишем теперь уравнения Максвелла во всех этих системах. Для полноты изложения приведем все соотношения сразу для материальных сред. Напомним, что электрическое поле в среде характеризуется двумя векторами \mathbf{E} и \mathbf{D} , связанными в гауссовой системе соотношением $\mathbf{D} = \mathbf{E} + 4\pi\mathbf{P}$ (\mathbf{P} — поляризованность). В других системах это соотношение записывают в виде

$$\mathbf{D} = \epsilon_0\mathbf{E} + \beta\mathbf{P},$$

где β — некоторая константа, которую всегда выбирают равной либо 1, либо 4π .

Магнитное поле в среде также характеризуется двумя векторами — средней напряженностью \mathbf{B} , называемой индукцией, и намагниченностью \mathbf{J} (аналогичной \mathbf{P}). Напряженность поля \mathbf{H} определяется соотношением

$$\mathbf{H} = \mu_0^{-1}\mathbf{B} - \beta'\mathbf{J},$$

где β' — некоторая константа, которую выбирают всегда равной 1 или 4π .

Уравнения Максвелла вместе с выражением для силы Лоренца приведены в следующей таблице:

Система	ϵ_0	μ_0	\mathbf{D}, \mathbf{H}	Уравнения Максвелла	Сила Лоренца
Электростатическая	1	c^{-2}	$\mathbf{D} = \mathbf{E} + 4\pi\mathbf{P}$ $\mathbf{H} = c^2\mathbf{B} - 4\pi\mathbf{J}$	$\text{rot } \mathbf{E} = -\partial\mathbf{B}/\partial t$ $\text{rot } \mathbf{H} = 4\pi\mathbf{j} + \partial\mathbf{D}/\partial t$ $\text{div } \mathbf{D} = 4\pi\rho$ $\text{div } \mathbf{B} = 0$	$q(\mathbf{E} + [\mathbf{vB}])$
Электромагнитная	c^{-2}	1	$\mathbf{D} = c^{-2}\mathbf{E} + 4\pi\mathbf{P}$ $\mathbf{H} = \mathbf{B} - 4\pi\mathbf{J}$	$\text{rot } \mathbf{E} = -\partial\mathbf{B}/\partial t$ $\text{rot } \mathbf{H} = 4\pi\mathbf{j} + \partial\mathbf{D}/\partial t$ $\text{div } \mathbf{D} = 4\pi\rho$ $\text{div } \mathbf{B} = 0$	$q(\mathbf{E} + [\mathbf{vB}])$
Гауссова	1	1	$\mathbf{D} = \mathbf{E} + 4\pi\mathbf{P}$ $\mathbf{H} = \mathbf{B} - 4\pi\mathbf{J}$	$\text{rot } \mathbf{E} = -\frac{1}{c}\partial\mathbf{B}/\partial t$ $\text{rot } \mathbf{H} = \frac{4\pi}{c}\mathbf{j} + \frac{1}{c}\partial\mathbf{D}/\partial t$ $\text{div } \mathbf{D} = 4\pi\rho$ $\text{div } \mathbf{B} = 0$	$q\left(\mathbf{E} + \frac{1}{c}[\mathbf{vB}]\right)$
Хевисайда—Лоренца	1	1	$\mathbf{D} = \mathbf{E} + \mathbf{P}$ $\mathbf{H} = \mathbf{B} - \mathbf{J}$	$\text{rot } \mathbf{E} = -\frac{1}{c}\partial\mathbf{B}/\partial t$ $\text{rot } \mathbf{H} = \frac{1}{c}(\mathbf{j} + \partial\mathbf{D}/\partial t)$ $\text{div } \mathbf{D} = \rho$ $\text{div } \mathbf{B} = 0$	$q\left(\mathbf{E} + \frac{1}{c}[\mathbf{vB}]\right)$
СИ	$\frac{1}{4\pi} \cdot \frac{1}{9} \cdot 10^{-9}$	$4\pi \cdot 10^{-7}$	$\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P}$ $\mathbf{H} = \frac{\mathbf{B}}{\mu_0} - \mathbf{J}$	$\text{rot } \mathbf{E} = -\partial\mathbf{B}/\partial t$ $\text{rot } \mathbf{H} = \mathbf{j} + \frac{1}{c}\partial\mathbf{D}/\partial t$ $\text{div } \mathbf{D} = \rho$ $\text{div } \mathbf{B} = 0$	$q(\mathbf{E} + [\mathbf{vB}])$

Приведем еще таблицу единиц в СИ и значения их в гауссовой системе (ГС):

Величина	Размерность в СИ	Единица	Значение в единицах ГС
Длина	L	метр (м)	10^2 см
Масса	M	килограмм (кг)	10^3 г
Время	T	секунда (с)	1 с
Сила электрического тока	I	ампер (А)	$3 \cdot 10^9$ ед. ГС
Сила	LMT^{-2}	ньютон (Н)	10^5 дин
Работа, энергия	L^2MT^{-2}	джоуль (Дж)	10^7 эрг
Мощность	L^2MT^{-3}	ватт (Вт)	10^7 эрг·с ⁻¹
Электрический заряд	TI	кулон (Кл)	$3 \cdot 10^9$ ед. ГС
Плотность электрического тока	$L^{-2}I$	ампер на квадратный метр (А/м ²)	$3 \cdot 10^5$ см ⁻² ед. силы тока ГС
Плотность заряда	$L^{-3}TI$	кулон на кубический метр (Кл/м ³)	$3 \cdot 10^3$ см ⁻³ ед. заряда ГС
Электрический потенциал, э. д. с.	$L^2MT^{-3}I^{-1}$	Вольт (В)	$1/300$ ед. потенциала ГС
Напряженность электрического поля	$LMT^{-3}I^{-1}$	вольт на метр (В/м)	$\frac{1}{3} \cdot 10^{-4}$ см ⁻¹ ед. потенциала ГС
Электрическая индукция	$L^{-2}TI$	кулон на квадратный метр (Кл/м ²)	$12 \cdot 10^5$ см ⁻² ед. заряда ГС
Электрическая емкость	$L^{-2}M^{-1}T^4I^2$	фарад (Ф)	$9 \cdot 10^{11}$ см
Электрическое сопротивление	$L^2MT^{-3}I^{-2}$	ом (Ом)	$\frac{1}{9} \cdot 10^{-11}$ см ⁻¹ с
Удельное электрическое сопротивление	$L^3MT^{-3}I^{-2}$	ом·метр (Ом·м)	$\frac{1}{9} \cdot 10^{-9}$ с
Электрическая проводимость	$L^{-2}M^{-1}T^3I^2$	сименс (См)	$9 \cdot 10^{11}$ см с ⁻¹
Магнитная индукция	$MT^{-2}I^{-1}$	тесла (Тл)	10^4 Гс
Магнитный поток	$L^2MT^{-2}I^{-1}$	вебер (Вб)	10^8 Гс·см ² = 10^8 Мкс
Напряженность магнитного поля	$L^{-1}I$	ампер на метр (А/м)	$4 \pi \cdot 10^{-3}$ Э
Намагниченность	$L^{-1} I$	ампер на метр (А/м)	$\frac{1}{4\pi} \cdot 10^4$ Гс
Индуктивность	$L^2MT^{-2}I^{-2}$	генри (Гн)	10^9 см

Здесь Гс (\equiv см^{-1/2}·г^{1/2}·с⁻¹) — гаусс — единица **B**; Э (\equiv см^{-1/2} × г^{1/2}·с⁻¹) — эрстед — единица **H**; Мкс (\equiv см^{3/2}·г^{1/2}·с⁻¹) — максвелл — единица магнитного потока (в гауссовой системе).

В проводимой ниже таблице сведены основные законы и формулы электромагнетизма в ГС (величины отмечены *) и СИ и указаны соотношения между величинами X^* и X :

ГС	СИ	Соотношение между величиной X^* и соответствующей величиной X
$c \cdot \text{rot } \mathbf{E}^* = -\partial \mathbf{B}^* / \partial t$	$\text{rot } \mathbf{E} = -\partial \mathbf{B} / \partial t$	$E^* / E = (4 \pi \epsilon_0)^{1/2}$
$\text{div } \mathbf{D}^* = 4 \pi \rho^*$	$\text{div } \mathbf{D} = \rho$	$D^* / D = (4 \pi / \epsilon_0)^{1/2}$
$\text{div } \mathbf{B}^* = 0$	$\text{div } \mathbf{B} = 0$	$B^* / B = (4 \pi / \mu_0)^{1/2}$
$c \cdot \text{rot } \mathbf{H}^* = 4 \pi \mathbf{j}^* + \partial \mathbf{D}^* / \partial t$	$\text{rot } \mathbf{H} = \mathbf{j} + \partial \mathbf{D} / \partial t$	$H^* / H = (4 \pi \mu_0)^{1/2}$
$\mathbf{F} = q^* \mathbf{E}^* + q^* [\mathbf{v} \mathbf{B}^*] / c$	$\mathbf{F} = q \mathbf{E} + q [\mathbf{v} \mathbf{B}]$	$q^* / q = (4 \pi \epsilon_0)^{-1/2}$
$\omega = (\mathbf{E}^* \mathbf{D}^* + \mathbf{B}^* \mathbf{H}^*) / (8 \pi)$	$\omega = (\mathbf{E} \mathbf{D} + \mathbf{B} \mathbf{H}) / 2$	$\rho^* / \rho = (4 \pi \epsilon_0)^{-1/2}$
$\mathbf{S} = c [\mathbf{E}^* \mathbf{H}^*] / (4 \pi)$	$\mathbf{S} = [\mathbf{E} \mathbf{H}]$	$\mathbf{j}^* / \mathbf{j} = (4 \pi \epsilon_0)^{-1/2}$
$\mathbf{E}^* = -\text{grad } \varphi^* - \frac{1}{c} \frac{\partial \mathbf{A}^*}{\partial t}$	$\mathbf{E} = -\text{grad } \varphi - \partial \mathbf{A} / \partial t$	$I^* / I = (4 \pi \epsilon_0)^{-1/2}$
$\mathbf{B}^* = \text{rot } \mathbf{A}^*$	$\mathbf{B} = \text{rot } \mathbf{A}$	$\varphi^* / \varphi = (4 \pi \epsilon_0)^{1/2}$
$\epsilon \mathbf{E}^* = \mathbf{D}^*, \epsilon = 1 + 4 \pi \alpha^*$	$\epsilon_0 \mathbf{E} = \mathbf{D}, \epsilon = 1 + \alpha$	$A^* / A = (4 \pi / \mu_0)^{1/2}$
$\mathbf{E}^* = \mathbf{D}^* - 4 \pi \mathbf{P}^*$	$\epsilon_0 \mathbf{E} = \mathbf{D} - \mathbf{P}$	$P^* / P = (4 \pi \epsilon_0)^{-1/2}$
$\mathbf{B} = \mu \mathbf{H}^*; \mu = 1 + 4 \pi \chi^*$	$\mathbf{B} = \mu_0 \mu \mathbf{H}, \mu = 1 + \chi$	$J^* / J = (4 \pi / \mu_0)^{-1/2}$
$\mathbf{B}^* = \mathbf{H}^* + 4 \pi \mathbf{J}^*$	$\mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{J})$	$R^* / R = 4 \pi \epsilon_0$
$\alpha_0 = \hbar^2 / (m e^2)$	$\alpha_0 = \hbar^2 4 \pi \epsilon_0 / (m e^2)$	$\sigma^* / \sigma = (4 \pi \epsilon_0)^{-1}$
$\hbar c R_\infty = e^2 / (2 a_0)$	$\hbar c R_\infty = e^2 / (8 \pi \epsilon_0 a_0)$	$C^* / C = (4 \pi \epsilon_0)^{-1}$
$\mu_B^* = e^* \hbar / (2 m c)$	$\mu_B = e \hbar / (2 m)$	$L^* / L = 4 \pi / \mu_0$
$\gamma^* = g_L [e^* / (2 m c)]$	$\gamma = g_L [e / (2 m)]$	$\mu_B^* / \mu_B = (4 \pi / \mu_0)^{-1/2}$
$\omega_L = \mu_B^* B^* / \hbar$	$\omega_L = \mu_B B / \hbar$	$e^* / e = (4 \pi \epsilon_0)^{-1/2}$
$\alpha = e^{*2} / (\hbar c)$ (кв. эд.) [*]	$\alpha = e^2 / (4 \pi \epsilon_0 \hbar c)$ (кв. эд.)	$\alpha^* / \alpha = \chi^* / \chi = (4 \pi)^{-1}$
$r_e = e^{*2} / (m c^2)$	$r_e = e^2 / (4 \pi \epsilon_0 m c^2)$	

* Кв. эд. означает, что α есть константа связи в квантовой электродинамике.

В заключение этого параграфа приведем еще определения основных единиц в СИ.

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

Секунда определяется как время, равное 9192631770 периодам излучения, соответствующего переходу между двумя сверхтонкими уровнями основного состояния атома цезия-133.

Ампер определяется как сила неизменяющегося тока, который при прохождении по двум параллельным прямолинейным проводникам бесконечной длины и ничтожно малого сечения, расположенным на расстоянии 1 м один от другого, в вакууме вызывает между проводниками на участке длиной 1 м силу, равную $2 \cdot 10^{-7}$ Н/м.

К числу основных единиц в СИ принадлежат **моль** для количества вещества, **кандела** (кд) для силы света и **кельвин** (К) для термодинамической температуры. Они определяются следующим образом.

Моль — единица количества вещества, равная такому его количеству, в котором содержится столько же структурных элементов, сколько атомов содержится в 0,012 кг изотопа углерода—12.

Кандела — единица силы света, равная силе света, испускаемого с поверхности площадью $1/6 \cdot 10^{-6} \text{ м}^2$ в перпендикулярном направлении черным телом при температуре затвердевания платины и давлении $101\,325 \text{ Нм}^{-2}$.

Кельвин — единица термодинамической температуры, равная $1/273,16$ термодинамической температуры тройной точки воды.

ОСНОВНЫЕ ФОРМУЛЫ

Уравнение Максвелла для полей в вакууме

$$\text{rot } \mathbf{E} = - \frac{\partial \mathbf{B}}{\partial t}, \quad \text{div } \mathbf{B} = 0$$

$$\text{rot } \mathbf{B} = \varepsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t}, \quad \text{div } \mathbf{E} = 0$$

Волновое уравнение

$$\Delta f - \frac{1}{c^2} \frac{\partial^2 f}{\partial t^2} = 0,$$

$$c^2 = (\varepsilon_0 \mu_0)^{-1}$$

Плотности электрической и магнитной энергий в вакууме

$$w_e = \frac{1}{2} \mathbf{E} \mathbf{D} = \frac{1}{2} \varepsilon_0 E^2$$

$$w_m = \frac{1}{2} \mathbf{B} \mathbf{H} = \frac{1}{2\mu_0} B^2$$

Вектор плотности потока электромагнитной энергии (вектор Пойнтинга)

$$\mathbf{S} = [\mathbf{E} \mathbf{H}]$$

Импульс поля

$$\mathbf{P}_F = c^{-2} \int_V [\mathbf{E} \mathbf{H}] dV$$

Момент поля

$$\mathbf{M}_F = c^{-2} \int_V [\mathbf{r} [\mathbf{E} \mathbf{H}]] dV$$

Уравнения Максвелла—Лоренца для усредненных полей
 $\langle \mathbf{e}(\mathbf{r}, t) \rangle = \mathbf{E}(\mathbf{r}, t);$

$$\text{rot } \mathbf{E} = - \frac{\partial \mathbf{B}}{\partial t}; \quad \text{div } \mathbf{B} = 0;$$

$$\text{rot } \mathbf{H} = \mathbf{j} + \frac{\partial \mathbf{D}}{\partial t}; \quad \text{div } \mathbf{D} = \rho;$$

$\langle \mathbf{b}(\mathbf{r}, t) \rangle = \mathbf{B}(\mathbf{r}, t)$

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P}; \quad \mathbf{H} = \mu_0^{-1} \mathbf{B} - \mathbf{J}$$

Плотности поляризации тока и тока намагниченности

$$\mathbf{j}_P = \partial \mathbf{P} / \partial t, \quad \mathbf{j}_m = \text{rot } \mathbf{J}$$

Граничные условия

$$E_{1t} = E_{2t}, \quad D_{1n} = D_{2n},$$

$$H_{1t} = H_{2t}, \quad B_{1n} = B_{2n}$$

Баланс энергии в среде

$$P^{(e)} = \int_V (\mathbf{j}^2/\sigma) dV + \int_V (\mathbf{E} \mathbf{D} + \mathbf{H} \mathbf{B}) dV + \int_{\Sigma} S ds$$

Изменение внутренней энергии единичного объема среды (s — плотность энтропии)

$$d\omega = T ds + \mathbf{E} d\mathbf{D} + \mathbf{H} d\mathbf{B}$$

Плотность, свободной энергии при изотермическом процессе и линейных связях между \mathbf{D} , \mathbf{E} , и \mathbf{B} , \mathbf{H}

$$F = 1/2 (\mathbf{E} \mathbf{D} + \mathbf{H} \mathbf{B})$$



Глава 7. ФИЗИКА ПРОСТРАНСТВА — ВРЕМЕНИ

7.1. Преобразования Лоренца

В § 6.2 было показано, что в вакууме могут распространяться поперечные электромагнитные волны с вполне определенной скоростью $c = (299792,4580 \pm 0,0012)$ км/с. Этот результат представляется парадоксальным, потому что он не связан с какой-либо системой отсчета и поэтому противоречит закону сложения скоростей классической механики. В самом деле, согласно этому закону, если \mathbf{v} — скорость волны в инерциальной системе отсчета K , то в другой инерциальной системе K' скорость волны должна быть равна $\mathbf{v} - \mathbf{V}$, где \mathbf{V} — скорость системы K' относительно системы K . Скорость же электромагнитной волны этому условию не удовлетворяет, имея одинаковое значение во всех инерциальных системах отсчета.

Для того чтобы разрешить этот парадокс, нужно отказаться от молчаливо делаемого в классической механике предположения об абсолютном характере времени и считать несправедливым преобразование Галилея $\mathbf{r}' = \mathbf{r} - \mathbf{V}t$, связывающее координаты \mathbf{r} и \mathbf{r}' какого-либо события в системах отсчета K и K' .

Действительно, рассмотрим одномерное волновое уравнение

$$\frac{\partial^2 \varphi}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 \varphi}{\partial t^2} = 0,$$

решения которого имеют вид плоских волн, распространяющихся со скоростью c . Для простоты будем считать, что величина $\varphi \equiv \varphi(x, t)$ представляет собой некоторую скалярную функцию от x и t . Тогда, если это уравнение относится к инерциальной системе отсчета K , в другой инерциальной системе K' соответствующее волновое уравнение должно иметь вид

$$\frac{\partial^2 \varphi'}{\partial x'^2} - \frac{1}{c^2} \frac{\partial^2 \varphi'}{\partial t'^2} = 0,$$

причем $\varphi' \equiv \varphi'(x', t') = \varphi(x, t)$. Но такая связь между уравнениями, определяющими величины φ и φ' , невозможна, если справедливо преобразование Галилея $x' = x - Vt$, $t' = t$. Поэтому мы заменим его более общим преобразованием

$$x' = \alpha x + \beta t, \quad t' = \gamma x + \delta t,$$

где α , β , γ и δ — некоторые константы, которые могут зависеть только от относительной скорости систем отсчета V . Постараемся подобрать их таким образом, чтобы из уравнения для φ вытекало уравнение для φ' . Заметим с этой целью, что

$$\begin{aligned} \frac{\partial^2 \varphi}{\partial x^2} &= \frac{\partial^2 \varphi'}{\partial x'^2} \alpha^2 + 2 \frac{\partial^2 \varphi'}{\partial x' \partial t'} \alpha \gamma + \frac{\partial^2 \varphi'}{\partial t'^2} \gamma^2, \\ \frac{\partial^2 \varphi}{\partial t^2} &= \frac{\partial^2 \varphi'}{\partial x'^2} \beta^2 + 2 \frac{\partial^2 \varphi'}{\partial x' \partial t'} \beta \delta + \frac{\partial^2 \varphi'}{\partial t'^2} \delta^2. \end{aligned}$$

Поэтому неизменность, или, как говорят, *инвариантность*, волнового уравнения приводит к условиям

$$-\gamma^2 + \delta^2/c^2 = 1; \quad \alpha^2 - \beta^2/c^2 = 1; \quad \alpha\gamma - \beta\delta/c^2 = 0.$$

Полагая $\beta/\alpha = -V$, получим модифицированное преобразование Галилея

$$x' = \frac{x - Vt}{\sqrt{1 - V^2/c^2}}, \quad t' = \frac{t - Vx/c^2}{\sqrt{1 - V^2/c^2}},$$

которое называют преобразованием Лоренца.

Легко видеть, что V представляет собой скорость системы K' относительно системы K . Действительно, начало координат системы K' , т. е. точка $x' = 0$, движется в системе K по закону $x - Vt = 0$.

Самым удивительным в преобразовании Лоренца является то, что время t' в системе K' не совпадает с временем t в системе K .

Можно поэтому сказать, что инвариантность скорости света достигается путем отказа от абсолютного характера времени.

Чтобы разъяснить это, рассмотрим два события, происходящие в некоторой точке x в системе K в моменты времени t_1 и t_2 . Тогда в системе K' соответствующие моменты времени

$$t'_1 = \frac{t_1 - Vx/c^2}{\sqrt{1 - V^2/c^2}}, \quad t'_2 = \frac{t_2 - Vx/c^2}{\sqrt{1 - V^2/c^2}},$$

откуда следует, что

$$t_2 - t_1 = (t'_2 - t'_1) \sqrt{1 - V^2/c^2},$$

т. е. $t_2 - t_1 \neq t'_2 - t'_1$.

Физический смысл этого соотношения заключается в следующем. Рассмотрим часы U , покоящиеся в системе K в точке x , и пусть t_1

и t_2 — два показания этих часов. В системе K' часы U движутся со скоростью — V , а t'_1 и t'_2 — показания двух различных часов U'_1 и U'_2 , покоящихся, в K' в точках x'_1 и x'_2 . Тогда времена t'_1 и t'_2 покажут часы U'_1 и U'_2 , когда мимо них пролетают часы U . Таким образом, время, показываемое движущимися часами, меньше времени, показываемого покоящимися часами.

На первый взгляд кажется, что это утверждение противоречит эквивалентности систем отсчета, но в действительности это не так, потому что в системе K мы имели одни часы, а в системе K' часов двое.

Сравнивая показания часов, мы молчаливо предполагали, что все часы в системе K (и в системе K') идут синхронно. Но так как мы вынуждены были отказаться от абсолютного характера времени, то теперь следует разъяснить, как в принципе часы могут быть синхронизированы.

Проще всего часы синхронизировать, используя световые сигналы. Установим для этого в разных точках рассматриваемой системы одинаковые и одинаково идущие часы. Отправим из места, где находятся часы U_1 в момент времени t_1 , световой сигнал к часам U_2 . Если l — расстояние между часами U_1 и U_2 , то в момент прихода сигнала мы должны на часах U_2 установить время $t_2 = t_1 + l/c$.

Обратимся снова к преобразованию Лоренца и рассмотрим стержень, покоящийся в системе K' . Длина его равна $L = x_2 - x_1$, где x_2 и x_1 — координаты концов стержня. Выясним, какова будет длина стержня в системе K . Для этого нужно найти координаты x'_2 и x'_1 концов стержня в системе K' в один и тот же момент времени t' . Переписав преобразование Лоренца в виде

$$x = \frac{x' + Vt'}{\sqrt{1 - V^2/c^2}}; \quad t = \frac{t' + Vx'/c^2}{\sqrt{1 - V^2/c^2}}.$$

(оно отличается от написанного выше заменой V на $-V$), получим, полагая $t'_1 = t'_2 = t'$,

$$x'_2 - x'_1 = (x_2 - x_1) \sqrt{1 - V^2/c^2}.$$

Таким образом,

стержень имеет разную длину в разных системах отсчета. Наибольшую длину стержень имеет в той системе отсчета, в которой он покоится. Эту длину называют *собственной длиной стержня*.

Длина стержня в системе, в которой он движется, уменьшается по сравнению с собственной длиной в $(1 - V^2/c^2)^{-1/2}$ раз. Это явление называют *лоренцевым сокращением*.

Рассмотренное преобразование Лоренца отличается тем свойством, что оно не изменяет, или, как говорят, *оставляет инвариантной*, величину $x^2 - c^2t^2$, т. е. $x'^2 - c^2t'^2 = x^2 - c^2t^2$. Преобразование это не касалось координат y и z , но в общем случае мы должны рассматривать и такие преобразования, которые затрагивают все три координаты x ,

y, z . Преобразования Лоренца отличаются тем свойством, что квадратичная форма $x^2 + y^2 + z^2 - c^2 t^2$ остается при них неизменной:

$$x'^2 + y'^2 + z'^2 - c^2 t'^2 = x^2 + y^2 + z^2 - c^2 t^2.$$

Здесь x, y, z определяют пространственные координаты какого-либо события и t — момент времени, когда это событие произошло в системе K , а x', y', z', t' имеют аналогичный смысл для того же события в системе K' .

Преобразования Лоренца можно придать простой геометрический смысл. Введем с этой целью четырехмерное пространство (*4-пространство*), представляющее собой совокупность точек (x_1, x_2, x_3, x_4) , где $x_1 = x, x_2 = y, x_3 = z, x_4 = ict$ и x, y, z, t — координаты и время какого-либо события. Точки в этом пространстве называют *мировыми точками*, причем каждой мировой точке соответствует некоторое событие.

Преобразования Лоренца оставляют инвариантной величину $x_1^2 + x_2^2 + x_3^2 + x_4^2$. Поэтому можно сказать, что они соответствуют «вращениям» системы координат в рассматриваемом 4-пространстве (называемом псевдоевклидовым пространством Минковского).

Например, рассмотренный нами ранее частный случай преобразований Лоренца соответствует «повороту» в плоскости (x_1, x_4) . Четырехмерное вращение системы координат всегда можно рассматривать как совокупность движения системы отсчета в некотором направлении и поворота ее в обычном трехмерном пространстве.

Легко получить преобразование Лоренца, соответствующее движению системы K' относительно системы K в произвольном направлении со скоростью \mathbf{V} . Представим для этого вектор \mathbf{r} в виде

$$\mathbf{r} = \mathbf{r}_\perp + \mathbf{r}_\parallel = \left(\mathbf{r} - \frac{(\mathbf{r}\mathbf{V})\mathbf{V}}{V^2} \right) + \frac{(\mathbf{r}\mathbf{V})\mathbf{V}}{V^2},$$

где \mathbf{r}_\parallel и \mathbf{r}_\perp — составляющие вектора \mathbf{r} вдоль направления скорости \mathbf{V} и в перпендикулярном направлении. В соответствии с выведенными выше формулами мы можем написать

$$\begin{aligned} \mathbf{r}' &= \mathbf{r} - (\mathbf{r}\mathbf{V})\mathbf{V}/V^2 - (t - \mathbf{r}\mathbf{V}/c^2)\mathbf{V}/\sqrt{1 - V^2/c^2}, \\ t' &= (t - \mathbf{r}\mathbf{V}/c^2)/\sqrt{1 - V^2/c^2}. \end{aligned}$$

Преобразования Лоренца, связывающие различные инерциальные системы отсчета, вместе с вытекающей из них относительностью времени и пространства составляют основу теории относительности — одной из важнейших современных физических теорий. Она возникла в результате исследований Лоренца, Пуанкаре и Эйнштейна и была завершена в 1905 г. Эйнштейном, которому мы обязаны нынешним глубоким пониманием всех проблем теории относительности.

7.2. Собственное время

Рассмотрим два события: (x_1, y_1, z_1, t_1) и (x_2, y_2, z_2, t_2) . Если образовать величину

$$s_{12}^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 - c^2(t_1 - t_2)^2,$$

то она будет, очевидно, при преобразованиях Лоренца вести себя так же, как и величина $x^2 + y^2 + z^2 - c^2t^2$, т. е. будет инвариантом:

$$s_{12}^2 = s_{12}'^2,$$

где $s_{12}'^2 = (x_1' - x_2')^2 + (y_1' - y_2')^2 + (z_1' - z_2')^2 - c^2(t_1' - t_2')^2$. Величину s_{12} называют *интервалом* между рассматриваемыми событиями.

Для бесконечно близких событий в системе K интервал определяется формулой

$$ds^2 = dx^2 + dy^2 + dz^2 - c^2dt^2.$$

Перейдя в другую систему K' , имеем

$$ds'^2 = dx'^2 + dy'^2 + dz'^2 - c^2dt'^2.$$

С понятием интервала можно связать рассмотренное в предыдущем параграфе замедление хода движущихся часов.

Пусть часы движутся относительно системы K , причем движутся произвольным образом. Свяжем с часами систему, которую назовем K' и которая не будет, вообще говоря, инерциальной. Однако в течение бесконечно малых промежутков времени систему K' можно, очевидно, рассматривать как инерциальную.

Возьмем теперь бесконечно малый промежуток времени, равный dt в системе K , и найдем соответствующий ему промежуток времени dt' , отсчитываемый часами в системе K' . В этой системе часы покоятся, поэтому для них $dx' = dy' = dz' = 0$. Обращаясь к выражению для интервала и учитывая его инвариантность, имеем

$$ds^2 = dx^2 + dy^2 + dz^2 - c^2dt^2 = -c^2dt'^2.$$

Если часы движутся относительно K со скоростью v , то в написанном выражении для ds^2 можно заменить $dx^2 + dy^2 + dz^2$ на v^2dt^2 , в результате получим

$$dt' = \sqrt{1 - v^2/c^2} dt.$$

Подчеркнем, что здесь v может быть произвольной функцией времени, так как при выводе мы пользовались только тем, что система K' является инерциальной в течение времени dt .

Наши часы движутся с некоторым объектом (относительно которого они неподвижны). Поэтому естественно назвать время, показываемое этими часами, *собственным* (или истинным) *временем объекта*. Впредь мы будем обозначать его τ , так что

$$d\tau = dt \sqrt{1 - v^2/c^2};$$

$$\tau_1 - \tau_2 = \int_{t_1}^{t_2} dt \sqrt{1 - v^2/c^2}.$$

Мы видим, что

собственное время движущегося объекта всегда меньше времени, отсчитываемого какими-либо покоящимися часами.

Из инвариантности интервала следует общая формула для $d\tau$

$$d\tau = ds/(ic).$$

В заключение этого параграфа проиллюстрируем изменение хода времени на примере распада элементарной частицы, называемой мюоном. Время жизни покоящегося мюона составляет $\tau_\mu = 2,2 \cdot 10^{-6}$ с. Допустим, что мюон образуется в атмосфере на высоте 30 км и имеет скорость $v = 299\,791$ км/с. Чтобы достичь Земли, ему нужно время $t = 30/300\,000 = 10^{-4}$ с. Это время значительно больше τ_μ , тем не менее мюон успевает долететь до Земли. Происходит это потому, что время жизни мюона по отношению к Земле

$$\Delta t = \tau_\mu / \sqrt{1 - v^2/c^2} \approx 0,7 \cdot 10^{-3} \text{ с.}$$

7.3. Сложение скоростей

Приступая к выводу преобразований Лоренца, мы говорили, что инвариантность скорости света в вакууме, т. е. одинаковость ее во всех инерциальных системах отсчета, противоречит закону сложения скоростей в классической механике. Посмотрим теперь, как устраняется это противоречие и как выглядит закон сложения скоростей в теории относительности.

Скорости какой-либо частицы в системах K и K' равны соответственно

$$\mathbf{v} = d\mathbf{r}/dt, \quad \mathbf{v}' = d\mathbf{r}'/dt',$$

где дифференциалы координат и времен связаны между собой преобразованиями Лоренца

$$dx = \frac{dx' + Vdt'}{\sqrt{1 - V^2/c^2}}, \quad dy = dy', \quad dz = dz', \quad dt = \frac{dt' + Vdx'/c^2}{\sqrt{1 - V^2/c^2}}$$

(система K' предполагается движущейся со скоростью V относительно системы K). Разделив первые три равенства на четвертое, мы получим интересующие нас формулы преобразования скорости:

$$v_x = \frac{v'_x + V}{1 + v'_x V/c^2}, \quad v_y = \frac{v'_y \sqrt{1 - V^2/c^2}}{1 + v'_x V/c^2}, \quad v_z = \frac{v'_z \sqrt{1 - V^2/c^2}}{1 + v'_x V/c^2}.$$

Устремляя c к бесконечности, получим формулы сложения скоростей классической механики.

Если частица движется вдоль оси x , то

$$v = \frac{v' + V}{1 + v'V/c^2}.$$

Полагая здесь $v' = c$, получим $v = c$, т. е. скорость света c действительно обладает свойством инвариантности.

Написанному закону сложения скоростей можно придать очень изящную форму, если ввести гиперболические тангенсы скоростей (деленных на c) $\alpha = \text{th}(v/c)$, $\alpha' = \text{th}(v'/c)$, $\gamma = \text{th}(V/c)$. Мы получим тогда

$$\alpha = \alpha' + \gamma,$$

т. е. при параллельном движении складываются гиперболические тангенсы скоростей. Нетрудно убедиться, что

$$1 - v^2/c^2 = (1 - v'^2/c^2)(1 - V^2/c^2)/(1 + v'V/c^2)^2.$$

Из этой формулы видно, что если $v' = c$, то и $v = c$, причем этот вывод справедлив независимо от направления распространения света. Кроме того, если $v'^2 \leq c^2$, то и $v^2 \leq c^2$. Это значит, что скорость материальных объектов не может превосходить скорости света. Такой вывод вытекает и непосредственно из вида преобразований Лоренца, поскольку относительная скорость систем отсчета, т. е. относительная скорость материальных объектов, не может быть больше c , так как при $V > c$ преобразования Лоренца теряют смысл.

Определим, как изменяется направление скорости частицы при переходе от одной системы отсчета к другой. Считая для простоты, что $v_z = v'_z = 0$, и обозначая θ и θ' углы между \mathbf{v} и \mathbf{v}' и осью x (вдоль оси x направлена относительная скорость систем отсчета), так что $v_x = v \cos \theta$, $v_y = v \sin \theta$, $v'_x = v' \cos \theta'$, $v'_y = v' \sin \theta'$, получим из формул преобразований скоростей

$$\text{tg } \theta = v' \sqrt{1 - V^2/c^2} \sin \theta' / (v' \cos \theta' + V),$$

В наиболее интересном случае распространения света $v = v' = c$, и формула принимает вид

$$\text{tg } \theta = \frac{\sqrt{1 - V^2/c^2}}{\cos \theta' + V/c} \sin \theta'.$$

Эта формула описывает *аберрацию света*, т. е. изменение направления распространения света при переходе от одной системы отсчета к другой.

Легко получить также для аберрации света формулы

$$\sin \theta = \frac{\sqrt{1 - V^2/c^2}}{1 + \cos \theta' \cdot V/c} \sin \theta'; \quad \cos \theta = \frac{\cos \theta' + V/c}{1 + \cos \theta' \cdot V/c}.$$

При малых V/c отсюда получаем

$$\theta' - \theta = \sin \theta' \cdot V/c, \quad V/c \ll 1.$$

7.4. Преобразования Лоренца для электромагнитного поля

В § 5.2 мы отмечали относительность электрического и магнитного полей. Если, например, в системе K имеется магнитное поле \mathbf{B} и нет электрического поля, то в системе K' , движущейся относительно K со скоростью V , возникнет электрическое поле, которое при малых относительных скоростях систем равно $[V\mathbf{B}]/c$ (здесь и далее в этой главе мы обозначаем \mathbf{B} индукцию магнитного поля и используем ГС; так как речь идет о вакууме, то $\mathbf{B} = \mathbf{H}$).

Поэтому можно говорить о преобразовании электрического и магнитного полей при переходе из одной инерциальной системы в другую. Для области малых скоростей $V \ll c$ формулы преобразования были установлены в § 5.2, они имеют следующий вид:

$$\mathbf{E}' = \mathbf{E} + \frac{1}{c} [\mathbf{V}\mathbf{H}], \quad \mathbf{H}' = \mathbf{H} - \frac{1}{c} [\mathbf{V}\mathbf{E}],$$

где поля \mathbf{E} и \mathbf{H} действуют в системе K , а поля \mathbf{E}' и \mathbf{H}' — в системе K' .

Мы покажем теперь, что, используя эти формулы, можно получить общее преобразование полей, справедливое при любых скоростях V .

Вводя обозначение $\boldsymbol{\psi} = \mathbf{E} + i\mathbf{H}$, можно, очевидно, переписать преобразование полей при малых V/c в виде

$$\boldsymbol{\psi}' = \boldsymbol{\psi} - \frac{i}{c} [\mathbf{V}\boldsymbol{\psi}],$$

откуда при V , параллельном оси z ,

$$\psi'_x = \psi_x + i \frac{V}{c} \psi_y, \quad \psi'_y = \psi_y - i \frac{V}{c} \psi_x, \quad \psi'_z = \psi_z.$$

Так как $\psi'_z = \psi_z$, то речь идет о преобразовании комплексного вектора $\boldsymbol{\psi}$ с двумя компонентами: $\psi_x \equiv \psi_1$ и $\psi_y \equiv \psi_2$. Это преобразование мы запишем в матричном виде:

$$\psi_i \rightarrow \psi'_i = L_{ik}(V/c) \psi_k,$$

или сокращенно

$$\boldsymbol{\psi}' = L(V/c) \boldsymbol{\psi},$$

где L — матрица преобразования, зависящая от V/c (по дважды встречающемуся индексу производится суммирование от 1 до 2). При малых $V/c = \beta$ матрица L имеет вид

$$L(\beta) = \begin{pmatrix} 1 & i\beta \\ -i\beta & 1 \end{pmatrix}, \quad \beta \ll 1,$$

и задача заключается в том, чтобы найти матрицу $L(\beta)$ при любых β (безразмерную скорость β мы будем иногда называть просто скоростью).

С этой целью мы перейдем от системы K' , движущейся относительно K со скоростью β , к системе K'' , движущейся относительно

системы K' со скоростью β' (также направленной вдоль оси z). Таким образом мы имеем два преобразования:

$$\psi \rightarrow \psi' = L(\beta) \psi, \quad \psi' \rightarrow \psi'' = L(\beta') \psi',$$

но они эквивалентны одному преобразованию

$$\psi \rightarrow \psi'' = L(\beta'') \psi,$$

соответствующему переходу $K \rightarrow K''$, где β'' — скорость системы K'' относительно K . Поэтому должно иметь место матричное соотношение

$$L(\beta'') = L(\beta) L(\beta'),$$

где согласно закону сложения скоростей

$$\beta'' = (\beta + \beta') / (1 + \beta\beta').$$

Вспомним теперь, что если перейти от скоростей к гиперболическим тангенсам скоростей $\alpha = \text{th } \beta$, то закон сложения скоростей приобретает вид $\alpha'' = \alpha + \alpha'$. Поэтому, вводя обозначение $L(\beta) = \tilde{L}(\alpha)$, имеем

$$\tilde{L}(\alpha + \alpha') = \tilde{L}(\alpha) \tilde{L}(\alpha').$$

Мы получили функциональное уравнение для матрицы \tilde{L} , решение которого тривиально:

$$\tilde{L}(\alpha) = e^{\alpha\sigma},$$

где σ — константа (в данном случае — постоянная, т. е. не зависящая от α , двухрядная матрица). Вид ее легко найти, если учесть, что при малых α

$$\tilde{L}(\alpha) = \begin{pmatrix} 1 & i\alpha \\ -i\alpha & 1 \end{pmatrix}, \quad |\alpha| \ll 1.$$

Эту матрицу можно, очевидно, представить в виде $e^{\alpha\sigma}$, где

$$\sigma = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}.$$

Таким образом, мы нашли вид $\tilde{L}(\alpha)$ при любых α . Замечая, что

$$\sigma^{2n} = 1, \quad \sigma^{2n+1} = \sigma,$$

где n — целое число, и используя разложение $e^z = \sum z^n/n!$, можно убедиться, что

$$\tilde{L}(\alpha) = \begin{pmatrix} \text{ch } \alpha & i \text{ sh } \alpha \\ -i \text{ sh } \alpha & \text{ch } \alpha \end{pmatrix}.$$

Теперь нам нужно перейти лишь от $\tilde{L}(\alpha)$ к $L(\beta)$. В результате получим

$$L(\beta) = \frac{1}{\sqrt{1-\beta^2}} \begin{pmatrix} 1 & i\beta \\ -i\beta & 1 \end{pmatrix}, \quad \beta = V/c.$$

Вспоминая, что $\psi' = L(\beta)\psi$, где $\psi = (\mathbf{E} + i\mathbf{H})_{\perp}$ и символ \perp служит для обозначения перпендикулярных (по отношению к \mathbf{V}) составляющих полей, легко установить общие формулы преобразования Лоренца для полей:

$$\mathbf{E}'_{\perp} = \left(\mathbf{E}_{\perp} + \frac{1}{c} [\mathbf{V}\mathbf{H}_{\perp}] \right) / \sqrt{1-V^2/c^2},$$

$$\mathbf{H}'_{\perp} = \left(\mathbf{H}_{\perp} - \frac{1}{c} [\mathbf{V}\mathbf{E}_{\perp}] \right) / \sqrt{1-V^2/c^2},$$

$$\mathbf{E}'_{\parallel} = \mathbf{E}_{\parallel}, \quad \mathbf{H}'_{\parallel} = \mathbf{H}_{\parallel},$$

где \mathbf{E}_{\parallel} , \mathbf{H}_{\parallel} — составляющие полей вдоль \mathbf{V} . Мы видим, что

параллельные составляющие полей не изменяются при переходе из одной системы отсчета в другую, перпендикулярные же составляющие «перепутываются» между собой.

Из формул преобразования полей следует, что

$$E^2 - H^2 = E'^2 - H'^2, \quad \mathbf{E}\mathbf{H} = \mathbf{E}'\mathbf{H}';$$

величины $E^2 - H^2$ и $\mathbf{E}\mathbf{H}$ являются инвариантами. Здесь имеется, однако, одна тонкость. Дело в том, что мы рассматривали только непрерывные преобразования Лоренца, представляющие собой вращения 4-мира Минковского. Но в принципе возможны еще два преобразования дискретных: преобразование пространственного отражения ($\mathbf{r} \rightarrow \mathbf{r}' = -\mathbf{r}$) и преобразование обращения времени ($t \rightarrow t' = -t$). Эти преобразования также должны учитываться, но при них величина $\mathbf{E}\mathbf{H}$ изменяет свой знак, т. е. $\mathbf{E}\mathbf{H} \rightarrow \mathbf{E}'\mathbf{H}' = -\mathbf{E}\mathbf{H}$. Поэтому, если мы хотим иметь истинный инвариант решительно при всех возможных преобразованиях 4-мира, то должны в качестве него взять не $\mathbf{E}\mathbf{H}$, а $(\mathbf{E}\mathbf{H})^2$.

Используем преобразование Лоренца и определим напряженности электрического и магнитного полей равномерно движущегося электрического заряда. Пусть заряд движется в системе K со скоростью V вдоль оси z . Требуется определить создаваемые им поля \mathbf{E} и \mathbf{H} . Перейдем с этой целью в систему K' , в которой заряд покоится. В этой системе имеется, очевидно, только электрическое поле

$$\mathbf{E}' = e\mathbf{r}'/r'^3.$$

Система K движется относительно K' со скоростью $-\mathbf{V}$. Поэтому согласно формулам преобразования полей

$$\mathbf{E}_{\perp} = \mathbf{E}'_{\perp} / \sqrt{1-V^2/c^2}, \quad \mathbf{E}_{\parallel} = \mathbf{E}'_{\parallel},$$

$$\mathbf{H} = \frac{1}{c} [\mathbf{V}\mathbf{E}'_{\perp}] / \sqrt{1-V^2/c^2},$$

и остается лишь выразить радиус-вектор \mathbf{r}' от заряда к точке наблюдения (x' , y' , z') в системе K' через радиус-вектор \mathbf{r} от заряда к точке

наблюдения (x, y, z) в системе K (компоненты его равны $x-Vt, y, z$). Связь между компонентами этих векторов дается формулами

$$x' = \frac{x - Vt}{\sqrt{1 - V^2/c^2}}, \quad y' = y, \quad z' = z,$$

с помощью которых легко убедиться, что

$$\mathbf{E} = \frac{e\mathbf{r}}{r^{*3}} (1 - V^2/c^2),$$

где $r^{*2} = (x - Vt)^2 + (y^2 + z^2) (1 - V^2/c^2) = r^2(1 - V^2/c^2 \sin^2 \Theta)$ и Θ — угол между \mathbf{r} и \mathbf{V} . Таким образом, окончательно получим

$$\mathbf{E} = \frac{e\mathbf{r}}{r^3} \frac{1 - V^2/c^2}{\left(1 - \frac{V^2}{c^2} \sin^2 \Theta\right)^{3/2}}, \quad \mathbf{H} = \frac{1}{c} [\mathbf{V}\mathbf{E}].$$

Мы видим, что при заданном расстоянии от заряда напряженность электрического поля \mathbf{E} возрастает с увеличением Θ от нуля до $\pi/2$. При $\Theta = 0$, т. е. при $\mathbf{r} \parallel \mathbf{V}$, напряженность имеет наименьшее значение:

$$E_{\parallel} = \frac{e}{r^2} (1 - V^2/c^2).$$

При $\Theta = \pi/2$, т. е. при $\mathbf{r} \perp \mathbf{V}$, напряженность максимальна:

$$E_{\perp} = \frac{e}{r^2} \frac{1}{\sqrt{1 - V^2/c^2}}.$$

При $V \rightarrow c$ продольная напряженность поля стремится к нулю. Можно сказать, что

с ростом V электрическое поле движущегося заряда как бы «сплющивается» по направлению движения и при $V \rightarrow c$ становится поперечным. Магнитное поле перпендикулярно электрическому, и при $V \rightarrow c$ его напряженность становится равной напряженности электрического поля.

7.5. Импульс и энергия частицы

Изучив преобразования Лоренца, перейдем теперь к задаче о динамическом законе движения в теории относительности, или, как говорят иначе, в релятивистской динамике. Для этого нужно иметь выражения для импульса частицы, производная по времени от которого всегда должна быть равна силе, действующей на частицу.

В классической механике импульс определяется как $\mathbf{p} = m\mathbf{v}$, где \mathbf{v} — скорость частицы и m — ее масса. Но такое определение не может быть правильным в теории относительности, так как импульс может быть сколь угодно большим как в классической, так и в релятивистской механике, скорость же в теории относительности ограничена и не может быть больше c .

Поэтому выражение $\mathbf{p} = m\mathbf{v}$ должно быть изменено. Мы покажем, что структура нового выражения для \mathbf{p} может быть установлена исходя из закона сохранения импульса и требования инвариантности его относительно преобразований Лоренца.

Рассмотрим с этой целью следующий мысленный эксперимент. Два наблюдателя A и B движутся вдоль оси x с относительной скоростью V и бросают друг другу вдоль оси y с одинаковой скоростью u шары одинаковой массы. Столкновение шаров происходит вдоль оси y , так что x -компоненты скоростей обоих шаров сохраняются.

Введем системы отсчета K_A и K_B , связанные соответственно с наблюдателями A и B , и обозначим \mathbf{v}_1^A и \mathbf{v}_2^B скорости шаров в этих системах до столкновения и \mathbf{v}_3^A и \mathbf{v}_4^B — после столкновения. Тогда, используя закон сложения скоростей, легко определить компоненты этих векторов в обеих системах:

$$\begin{aligned} (\mathbf{v}_1^A)_x &= 0; & (\mathbf{v}_1^A)_y &= u; & (\mathbf{v}_3^A)_x &= 0; & (\mathbf{v}_3^A)_y &= -u'; \\ (\mathbf{v}_2^B)_x &= 0; & (\mathbf{v}_2^B)_y &= -u; & (\mathbf{v}_4^B)_x &= 0; & (\mathbf{v}_4^B)_y &= u'; \\ (\mathbf{v}_2^A)_x &= -V; & (\mathbf{v}_2^A)_y &= -u \sqrt{1 - V^2/c^2}; \\ (\mathbf{v}_1^B)_x &= V; & (\mathbf{v}_1^B)_y &= u \sqrt{1 - V^2/c^2}; \\ (\mathbf{v}_3^B)_x &= V; & (\mathbf{v}_3^B)_y &= -u' \sqrt{1 - V^2/c^2}; \\ (\mathbf{v}_4^A)_x &= -V; & (\mathbf{v}_4^A)_y &= u' \sqrt{1 - V^2/c^2}, \end{aligned}$$

где $-u'$ — y -компонента скорости первого шара в системе A после столкновения.

Будем теперь предполагать, что импульс \mathbf{p} частицы определяется формулой

$$\mathbf{p} = m\mathbf{v}f(u^2),$$

где m — масса частицы и f — некоторая функция квадрата ее скорости, и что имеет место закон сохранения импульса. Тогда ясно, что проекции суммарного импульса шаров в системе A на оси x и y до и после столкновения

$$\begin{aligned} P_x &= -mVf(V^2 + u^2(1 - V^2/c^2)), \\ P'_x &= -mVf(V^2 + u'^2(1 - V^2/c^2)), \\ P_y &= mu'f(u^2) - mu \sqrt{1 - V^2/c^2} f(V^2 + u^2(1 - V^2/c^2)), \\ P'_y &= -mu'f(u'^2) + mu' \sqrt{1 - V^2/c^2} f(V^2 + u'^2(1 - V^2/c^2)). \end{aligned}$$

Равенство x -проекций дает $u = u'$, после чего равенство $P_y = P'_y$ приобретает вид

$$u \sqrt{1 - V^2/c^2} f(V^2 + u^2(1 - V^2/c^2)) = uf(u^2).$$

Разделив это равенство на u , устремим после этого u к нулю. В результате получим $\sqrt{1 - v^2/c^2} f(v^2) = f(0)$. При малых скоростях $\mathbf{p} = m\mathbf{v}$, т. е. $f(0) = 1$. Поэтому $f(v^2) = 1/\sqrt{1 - v^2/c^2}$ и мы приходим к следующему выражению для импульса частицы:

$$\mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1 - v^2/c^2}}.$$

При $v \ll c$, как и должно быть, $\mathbf{p} = m\mathbf{v}$. Если $v \rightarrow c$, то импульс стремится к бесконечности.

Производная от импульса по времени есть сила, действующая на частицу. Из выражения для \mathbf{p} следует, что если скорость меняется только по направлению, то

$$\frac{d\mathbf{p}}{dt} = \frac{m}{\sqrt{1 - v^2/c^2}} \frac{d\mathbf{v}}{dt}.$$

Если же скорость изменяется только по модулю, то

$$\frac{d\mathbf{p}}{dt} = \frac{m}{(1 - v^2/c^2)^{3/2}} \frac{d\mathbf{v}}{dt}.$$

В первом случае сила направлена перпендикулярно скорости, а во втором — вдоль скорости. Мы видим, что отношение силы $\mathbf{F} = d\mathbf{p}/dt$ к ускорению $d\mathbf{v}/dt$ в обоих этих случаях разное, т. е. эффективная масса, определяемая как отношение силы к ускорению, различается для двух направлений силы.

Определим теперь энергию частицы в релятивистской динамике. Если частица сместилась на $d\mathbf{r}$ под действием силы \mathbf{F} , то, очевидно, ее энергия изменилась на $d\mathcal{E} = \mathbf{F}d\mathbf{r}$, и так как $\mathbf{F} = d\mathbf{p}/dt$, то

$$d\mathcal{E} = \mathbf{v}d\mathbf{p}.$$

Подставляя сюда найденное выражение для \mathbf{p} , получим

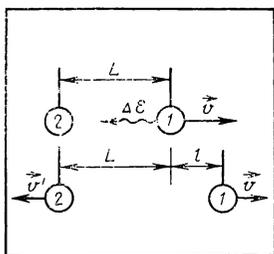
$$\begin{aligned} d\mathcal{E} &= (1 - v^2/c^2)^{-1/2} d \frac{mv^2}{2} + \\ &+ \frac{mv^2}{2} d(1 - v^2/c^2)^{-1/2} = mc^2 d(1 - v^2/c^2)^{-1/2}, \end{aligned}$$

откуда следует, что

$$\mathcal{E} = \frac{mc^2}{\sqrt{1 - v^2/c^2}} + \text{const.}$$

Вопрос теперь сводится к нахождению константы. Можно было бы, казалось, выбрать ее так, чтобы энергия обращалась в нуль для покоящейся частицы, т. е. выбрать $\text{const} = -mc^2$. В действительности, однако, нужно выбрать константу равной нулю, т. е. считать, что

$$\mathcal{E} = \frac{mc^2}{\sqrt{1 - v^2/c^2}}.$$



7.1.

К выводу соотношения Эйнштейна

При этом энергия покоящейся частицы не равна нулю, а есть

$$\mathcal{E}_0 = mc^2,$$

т. е. энергия равна произведению массы частицы на квадрат скорости света.

Такой выбор константы делает релятивистскую динамику последовательной теорией в том смысле, что импульс и энергия частицы

$$\mathbf{p} = m\mathbf{v}/\sqrt{1-v^2/c^2}, \quad \mathcal{E} = mc^2/\sqrt{1-v^2/c^2}$$

объединяются при этом в единый 4-вектор энергии — импульса. Мы вернемся

к этому вопросу в § 7.7, а пока заметим, что из соотношения $\mathcal{E}_0 = mc^2$ (оно было установлено Эйнштейном вместе с формулами для \mathbf{p} и \mathcal{E}) вытекает следующий замечательный вывод: *если частица теряет энергию $\Delta \mathcal{E}$, то ее масса одновременно уменьшается на*

$$\Delta m = \Delta \mathcal{E}/c^2.$$

Заметим, также, что из формул для энергии и импульса вытекают соотношения

$$\mathcal{E} = c \sqrt{m^2 c^2 + p^2}, \quad \mathbf{p} = \mathcal{E} \mathbf{v}/c^2.$$

При $v \ll c$ энергия имеет вид

$$\mathcal{E} = mc^2 + mv^2/2.$$

Второе слагаемое здесь представляет собой кинетическую энергию частицы в механике Ньютона, а первое слагаемое — *энергию покоя* частицы.

В формуле, связывающей \mathcal{E} и \mathbf{p} , можно положить $m = 0$. Мы получим тогда связь между энергией и импульсом безмассовой частицы

$$\mathcal{E} = c p.$$

Скорость ее, очевидно, равна c . Известным примером безмассовой частицы является фотон (квант света) — частица, связанная с электромагнитной волной.

Дадим теперь физическое доказательство соотношения $\Delta m = \Delta \mathcal{E}/c^2$, принадлежащее Эйнштейну.

Рассмотрим два тела, обладающих одинаковой массой m и находящихся на расстоянии L (рис.7.1). Пусть тело 1 излучает электромагнитную энергию $\Delta \mathcal{E}$, которая поглощается телом 2. При испускании энергии тело 1 испытывает отдачу и начинает двигаться со скоростью v . К моменту поглощения излучения телом 2 проходит время L/c и тело 1 смещается на расстояние $l = vL/c$. Чтобы найти скорость тела 1, заметим, что при излучении оно, согласно результатам § 6.2, теряет импульс $\Delta p = \Delta \mathcal{E}/c$. Одновременно тело 1 должно потерять некоторую массу Δm . Если тело достаточно массивно, то скорость его мала

по сравнению с c . Поэтому эта скорость должна быть связана с Δp соотношением $\Delta p = (m - \Delta m)v$. Подставляя сюда $\Delta p = \Delta \mathcal{E}/c$, получим

$$v = \Delta \mathcal{E}/[c(m - \Delta m)].$$

Таким образом, смещение тела l составит

$$l = L\Delta \mathcal{E}/[(m - \Delta m)c^2].$$

Учтем теперь, что центр тяжести обоих наших тел после обмена энергией, т. е. массой, не должен изменить своего положения. Это значит, что должно иметь место соотношение

$$(m - \Delta m) \left(\frac{L}{2} + l \right) = (m + \Delta m) \frac{L}{2}.$$

Подставляя сюда найденное выражение для l , мы приходим к соотношению Эйнштейна

$$\Delta m = \Delta \mathcal{E}/c^2.$$

7.6. Движение заряженной частицы в электромагнитном поле

Имея выражение для импульса частицы, можно сформулировать уравнение движения заряженной частицы, находящейся во внешнем электромагнитном поле. На частицу действует в этом случае сила (в ГС)

$$\mathbf{f} = e \left(\mathbf{E} + \frac{1}{c} [\mathbf{v}\mathbf{B}] \right),$$

где e — заряд частицы, \mathbf{v} — ее скорость и \mathbf{E} , \mathbf{B} — напряженности электрического и магнитного полей. Поэтому уравнение движения частицы имеет вид

$$\frac{d\mathbf{p}}{dt} = e \left(\mathbf{E} + \frac{1}{c} [\mathbf{v}\mathbf{B}] \right); \quad \mathbf{p} = m\mathbf{v}/\sqrt{1 - v^2/c^2}.$$

При малых скоростях ($v \ll c$) импульс $\mathbf{p} = m\mathbf{v}$ и мы получаем уравнение Ньютона с лоренцевой силой \mathbf{f} .

Определим, как изменяется во времени энергия частицы: $d\mathcal{E}/dt = \mathbf{v} \cdot d\mathbf{p}/dt$. Подставляя сюда $d\mathbf{p}/dt = \mathbf{f}$ и замечая, что $\mathbf{v}[\mathbf{v}\mathbf{B}] = 0$, получим

$$d\mathcal{E}/dt = e\mathbf{E}\mathbf{v}.$$

Мы видим, что изменение энергии частицы определяется только электрическим полем. Это объясняется тем, что изменение энергии равно работе сил, действующих на частицу, магнитная же сила перпендикулярна скорости и работы не производит.

Импульс частицы можно выразить через ее энергию:

$$\mathbf{p} = \mathcal{E}\mathbf{v}/c^2.$$

Подставляя это выражение в уравнение движения и используя формулу для $d\mathcal{E}/dt$, найдем ускорение частицы в электромагнитном поле:

$$\dot{\mathbf{v}} = \frac{e}{m} \sqrt{1 - v^2/c^2} \left\{ \mathbf{E} + \frac{1}{c} [\mathbf{vB}] - \frac{1}{c^2} \mathbf{v}(\mathbf{vE}) \right\}.$$

Если $v \ll c$, мы получим знакомую формулу классической механики $\dot{\mathbf{v}} = \mathbf{f}/m$.

Формула ускорения в релятивистской динамике, т. е. при $v \approx c$, имеет два отличия от формулы классической механики: во-первых, здесь появляется характерный множитель $\sqrt{1 - v^2/c^2}$ и, во-вторых, возникает добавочный член с напряженностью электрического поля $\mathbf{v}(\mathbf{vE})/c^2$.

Перейдем к изучению движения частицы в электромагнитном поле. Начнем с рассмотрения движения в постоянном и однородном электрическом поле. Если \mathbf{v}_0 — начальная скорость частицы, то движение частицы, очевидно, происходит в плоскости $(\mathbf{E}, \mathbf{v}_0)$. Направим ось x вдоль \mathbf{E} , тогда уравнение движения примет вид

$$\dot{p}_x = eE, \quad \dot{p}_y = 0,$$

откуда $p_x = eEt + p_{x0}$, $p_y = p_{y0}$, где p_{x0} и p_{y0} — начальные значения p_x и p_y . Если выбрать начало отсчета времени таким образом, чтобы ему соответствовало нулевое значение p_x , то

$$p_x = eEt, \quad p_y = p_0,$$

где p_0 — новое начальное значение импульса частицы.

Мы видим, что p_x неограниченно возрастает со временем. Посмотрим, как со временем изменяются скорость частицы и ее энергия. Энергия частицы \mathcal{E} связана с ее импульсом соотношением $\mathcal{E} = c\sqrt{m^2c^2 + p^2}$, поэтому в рассматриваемом случае

$$\mathcal{E} = c \sqrt{m^2c^2 + (eEt)^2 + p_0^2}.$$

Чтобы найти зависимость от времени скорости, воспользуемся соотношением $\mathbf{v} = c^2\mathbf{p}/\mathcal{E}$, откуда следует, что

$$v_x = \frac{c^2 p_x}{\mathcal{E}} = \frac{ceEt}{\sqrt{m^2c^2 + (eEt)^2 + p_0^2}},$$

$$v_y = \frac{c^2 p_y}{\mathcal{E}} = \frac{cp_0}{\sqrt{m^2c^2 + (eEt)^2 + p_0^2}}.$$

Мы видим, что

при неограниченном возрастании времени проекция скорости вдоль поля стремится к c , а перпендикулярная проекция скорости — к нулю.

Учитывая, что $v_x = dx/dt$ и $v_y = dy/dt$, мы можем найти координаты частицы как функции времени:

$$x = \frac{1}{eE} \sqrt{\varepsilon_0^2 + (ceEt)^2}; \quad y = \frac{p_0 c}{eE} \operatorname{Arsh} \frac{ceEt}{\varepsilon_0},$$

где $\varepsilon_0 = c\sqrt{m^2c^2 + p_0^2}$ (соответствующие постоянные интегрирования положены равными нулю). Отсюда легко найти траекторию частицы:

$$x = \frac{\varepsilon_0}{eE} \operatorname{ch} \frac{eEy}{p_0 c}.$$

Часть траектории, соответствующая малым y (при условии $eEy/(p_0 c) \ll 1$), имеет вид параболы:

$$x = \frac{eE}{2mv_0^2} y^2 + \operatorname{const}.$$

Перейдем теперь к изучению движения частицы в магнитном поле. Используя формулу $\mathbf{p} = \mathcal{E}\mathbf{v}/c^2$ и учитывая, что в магнитном поле энергия не изменяется, можно представить уравнения движения частицы в виде

$$\frac{1}{c^2} \mathcal{E} \frac{d\mathbf{v}}{dt} = \frac{e}{c} [\mathbf{v}\mathbf{B}]$$

или в виде

$$\dot{\mathbf{v}} = [\boldsymbol{\omega}\mathbf{v}],$$

где $\boldsymbol{\omega} = -e\mathbf{B}/\mathcal{E}$.

Это уравнение аналогично уравнению $\mathbf{v} = [\boldsymbol{\omega}\mathbf{r}]$, определяющему скорость частиц твердого тела, вращающегося с угловой скоростью $\boldsymbol{\omega}$. Поэтому можно сказать, что конец вектора скорости частицы вращается вокруг магнитного поля с угловой скоростью $\boldsymbol{\omega} = -e\mathbf{B}/\mathcal{E}$. Эту угловую скорость называют *циклотронной частотой*. В случае постоянного и однородного поля частота $\boldsymbol{\omega}$ является константой. В нерелятивистской ($v \ll c$) области $\mathcal{E} = mc^2$ и циклотронная частота приобретает знакомый нам вид

$$\boldsymbol{\omega} = -e\mathbf{B}/(mc)^2.$$

Легко видеть, что движение в плоскости, перпендикулярной \mathbf{B} , происходит по окружности радиуса

$$r = v_{\perp}/\omega = cp_{\perp}/(eB),$$

где v_{\perp} и p_{\perp} — составляющие скорости и импульса, перпендикулярные \mathbf{B} . Так как составляющая скорости частицы вдоль \mathbf{B} не изменяется, то траектория частицы есть винтовая линия с осью, направленной вдоль индукции \mathbf{B} , и с радиусом r .

7.7. 4-векторы и 4-тензоры

Как мы видели, при переходе от одной инерциальной системы отсчета к другой изменяются не только простран-

ственные и временная координаты каждого события, но претерпевают также изменения и компоненты электромагнитного поля, причем изменения происходят по вполне определенному закону — закону преобразования, связанному с тем преобразованием Лоренца, которое соответствует рассматриваемому переходу систем отсчета.

Подобным же образом любой физической величине свойствен свой закон преобразования при переходе от одной системы отсчета к другой.

Эти законы можно расклассифицировать и установить в общем виде, если ввести понятия *четырёхмерных векторов* и *тензоров* (сокращенно 4-векторов и 4-тензоров).

4-вектором называют совокупность четырех величин A_1, A_2, A_3, A_4 , если она преобразуется так же, как координаты мировой точки x_1, x_2, x_3, x_4 .

Лоренцевское преобразование координат мировой точки мы будем записывать в виде

$$x_i \rightarrow x'_i = L_{ik}x_k,$$

где L_{ik} — некоторые величины, зависящие от относительной скорости систем отсчета (подразумевается суммирование по индексу k от 1 до 4). Так как величина $x_i^2 \equiv x_1^2 + x_2^2 + x_3^2 + x_4^2$ не изменяется при преобразовании Лоренца, то матрицы L удовлетворяют соотношениям

$$L_{ik}L_{jk} = \delta_{ij},$$

где $\delta_{ij} = 1$ при $i = j$ и $\delta_{ij} = 0$ при $i \neq j$ (здесь и в дальнейшем подразумевается суммирование по дважды встречающемуся индексу). Поэтому 4-вектор A_i преобразуется по закону

$$A_i \rightarrow A'_i = L_{ik}A_k.$$

Величина $A_i A_i \equiv A_1^2 + A_2^2 + A_3^2 + A_4^2$, так же как и величина $x_i x_i$, остается инвариантной при лоренцевских преобразованиях: $A_i A_i = A'_i A'_i$. Эту величину можно записать в виде

$$A_i A_i = \mathbf{A}^2 - A_0^2,$$

где \mathbf{A} объединяет три пространственные компоненты A_1, A_2, A_3 4-вектора, а $A_0 = A_4/i$ — временная компонента 4-вектора (A_0 — вещественная величина). Таким образом,

$$\mathbf{A}^2 - A_0^2 = \mathbf{A}'^2 - A_0'^2.$$

Если взять два 4-вектора A_i и B_i , то величина $A_i B_i = \mathbf{A}\mathbf{B} - A_0 B_0$ преобразуется при лоренцевских преобразованиях, так же, как и $A_i A_i$, т. е. является инвариантом, или, как говорят иначе, *скаляром*:

$$\mathbf{A}\mathbf{B} - A_0 B_0 = \mathbf{A}'\mathbf{B}' - A_0' B_0'.$$

Эту величину называют *скалярным произведением 4-векторов* A_i и B_i

Существует и обратная теорема: если $A_i B_i$ — скаляр и A_i — 4-вектор, то и B_i — 4-вектор.

Приведем примеры 4-векторов. Так как координаты x_i представляют собой 4-вектор, то и дифференциалы координат dx_i образуют 4-вектор. Величина $dx_i dx_i = ds^2$ представляет собой инвариант (квадрат интервала). Поэтому инвариантом является и дифференциал собственного времени

$$d\tau = ds/(ic).$$

Разделив dx_i на $d\tau$, мы получим, следовательно, 4-вектор скорости:

$$u_i = dx_i/d\tau.$$

Учитывая, что $d\tau = dt\sqrt{1-v^2/c^2}$, где v — скорость частицы, можно записать u_i в виде

$$u_i \equiv \left(\frac{dr}{dt} / \sqrt{1-v^2/c^2}, \quad ic / \sqrt{1-v^2/c^2} \right).$$

Умножив u_i на массу частицы, мы получим 4-вектор импульса, или иначе 4-вектор энергии-импульса частицы:

$$p_i \equiv (mv / \sqrt{1-v^2/c^2}, \quad imc / \sqrt{1-v^2/c^2}).$$

Пространственные компоненты p_i определяют импульс \mathbf{p} , а временная компонента — энергию частицы \mathcal{E} , умноженную на i/c . Отсюда можно сразу написать, как преобразуется импульс и энергия частицы при переходе от одной инерциальной системы к другой:

$$p'_x = \frac{p_x - V\mathcal{E}/c^2}{\sqrt{1-V^2/c^2}}, \quad \mathcal{E}' = \frac{\mathcal{E} - Vp_x}{\sqrt{1-V^2/c^2}}, \quad p'_y = p_y, \quad p'_z = p_z$$

(движение системы K' по отношению к K происходит вдоль оси x со скоростью V). Покажем, что

плотность электрического тока $\mathbf{j} = \rho\mathbf{v}$ и плотность заряда ρ образуют 4-вектор:

$$(\rho v, \quad ic\rho) \equiv j_i,$$

Умножим для этого равенство $de = \rho dV$ на dx_i :

$$dedx_i = \rho dV dt \frac{dx_i}{dt}.$$

Легко убедиться, что $dVdt$ представляет собой инвариант. Действительно,

$$dV'dt' = dx'_1 dx'_2 dx'_3 dx'_4 / (ic) = \left\| \partial x'_i / \partial x_k \right\| dx_1 dx_2 dx_3 dx_4 / (ic),$$

где $\left\| \partial x'_i / \partial x_k \right\|$ — якобиан перехода от переменных x_i к переменным x'_i . Но этот якобиан, как видно из преобразований Лоренца, равен единице.

С другой стороны, $dedx_i$ представляет собой 4-вектор. Поэтому 4-вектором является и $j_i = \rho dx_i / dt \equiv (\rho v, \quad ic\rho)$. Отсюда следует, что

плотности тока и заряда преобразуются при переходе от одной инерциальной системы отсчета к другой согласно формулам

$$j'_x = \frac{j_x - V\rho}{\sqrt{1 - V^2/c^2}}; \quad \rho' = \frac{\rho - Vj_x/c^2}{\sqrt{1 - V^2/c^2}}; \quad j'_y = j_y, \quad j'_z = j_z.$$

Перейдем теперь к определению 4-тензоров. 4-тензором второго ранга называют совокупность 16 величин T_{ij} ($i, j = 1, 2, 3, 4$), преобразующуюся при лоренцевском преобразовании по закону

$$T_{ij} \rightarrow T'_{ij} = L_{ik}L_{jl}T_{kl}.$$

В частности, 4-тензор второго ранга образуют величины A_iB_j , где A_i и B_j — два произвольных 4-вектора. Так как A_iB_i представляет собой инвариант (или скаляр), то инвариантом будет и величина T_{ii} , т. е. сумма диагональных элементов матрицы T_{ij} .

Аналогичным образом можно определить тензоры более высокого ранга. Так, тензором 3-го ранга называют совокупность 4^3 величин T_{ijl} , преобразующихся при лоренцевских преобразованиях согласно закону

$$T_{ijl} \rightarrow T'_{ijl} = L_{in}L_{jm}L_{lp}T_{nmp},$$

тензором 4-го ранга — совокупность 4^4 величин T_{ijkl} , преобразующихся по закону

$$T_{ijkl} \rightarrow T'_{ijkl} = L_{ir}L_{js}L_{lt}L_{ku}T_{rstu},$$

и т. д.

Простейшими примерами тензоров являются совокупности произведений компонент 4-векторов. Так, величины $A_iB_jC_l$, где A_i, B_i, C_i — 4-векторы, образуют тензор 3-го ранга, а величины $A_iB_jC_lD_k$ — тензор 4-го ранга (D_k — 4-вектор).

Отсюда легко заключить, что если в тензоре считать два индекса одинаковыми и произвести по этим индексам суммирование, то ранг тензора понизится на две единицы. В частности, T_{iij} — тензор 1-го ранга, или 4-вектор, а T_{ijil} — тензор 2-го ранга. Понижение ранга тензора, производимое таким образом, называют *операцией сокращения*.

Если T_{ij} и A_i — 4-тензор 2-го ранга и 4-вектор, то совокупность величин $T_{ij}A_l$ — 4-тензор 3-го ранга, а совокупность величин $B_i = T_{ij}A_j$ — 4-вектор. Наоборот, если известно, что A_i и B_i — 4-векторы, то связывающие их величины T_{ij} образуют 4-тензор 2-го ранга. Отсюда следует, что, если компоненты 4-вектора A_i зависят от координат x_i , т. е. $A_i = A_i(x_i)$ представляет собой некоторое 4-векторное поле, то производные $\partial A_i / \partial x_k$ образуют 4-тензор 2-го ранга. Действительно, $dA_i = dx_k \partial A_i / \partial x_k$, а так как dx_i и dA_i — 4-векторы, то $\partial A_i / \partial x_k$ — 4-тензор.

Важнейшим примером 4-тензоров является *тензор электромагнитного поля*

$$F_{ik} = \begin{pmatrix} 0 & B_z & -B_y & -iE_x \\ -B_z & 0 & B_x & -iE_y \\ B_y & -B_x & 0 & -iE_z \\ iE_x & iE_y & iE_z & 0 \end{pmatrix},$$

объединяющий компоненты векторов электрического и магнитного полей \mathbf{E} и \mathbf{B} (напомним, что в вакууме $\mathbf{B} = \mathbf{H}$). То, что величины F_{ik} образуют 4-тензор, вытекает из закона преобразования электромагнитного поля и определения тензора.

Обратим внимание на то обстоятельство, что ни \mathbf{E} , ни \mathbf{B} не являются 4-векторами и только написанная совокупность их компонент представляет собой 4-тензор 2-го ранга. Тензор этот антисимметричен, т. е. $F_{ik} = -F_{ki}$, и имеет поэтому шесть отличных от нуля компонент.

Вместе с тем следует иметь в виду, что если речь идет только о пространственных вращениях, то и \mathbf{E} , \mathbf{B} — векторы (точнее, 3-векторы).

Наконец, следует иметь в виду своеобразие вектора индукции магнитного поля: при обращении времени ($t \rightarrow -t$) он меняет свой знак ($\mathbf{B} \rightarrow -\mathbf{B}$), в то время как напряженность электрического поля при этом не изменяется; при преобразовании отражения пространства ($\mathbf{r} \rightarrow -\mathbf{r}$), каждый истинный вектор (например, вектор скорости \mathbf{v} и вектор электрического поля \mathbf{E}) изменяет свой знак, вектор же \mathbf{B} знака не меняет, поэтому его называют не вектором, а *псевдовектором*.

Антисимметричный тензор поля F_{ik} всегда можно представить в виде

$$F_{ik} = \frac{\partial A_k}{\partial x_i} - \frac{\partial A_i}{\partial x_k},$$

где A_i — некоторый 4-вектор. Его называют *4-вектором потенциала электромагнитного поля*. Обозначая его пространственные компоненты через \mathbf{A} , а временную — через $i\phi$, можно записать магнитную индукцию и напряженность электрического поля в виде

$$\mathbf{B} = \text{rot } \mathbf{A}, \quad \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} - \text{grad } \phi.$$

Таким образом, \mathbf{A} — это уже знакомый нам векторный потенциал магнитного поля. В статическом случае, когда поля не зависят от времени, $\mathbf{E} = -\text{grad } \phi$, т. е. ϕ в этом случае есть электростатический потенциал.

Теперь мы должны понять, в чем заключается физический смысл введения понятий 4-векторов и 4-тензоров, т. е. выяснить, зачем они нужны. Ответ гласит, что все физические величины могут быть объединены в такие совокупности, которые являются либо 4-векторами, либо 4-тензорами различных рангов, либо, наконец, с п и н о р а м и, т. е. величинами, квадратичные комбинации которых образуют 4-векторы.

Например, импульс \mathbf{p} и энергия \mathcal{E} частицы образуют 4-вектор импульса p_i , векторы \mathbf{E} и \mathbf{B} образуют 4-тензор электромагнитного поля F_{ik} и т. д. Поэтому можно сказать, что 4-векторы и 4-тензоры (а также

4-спиноры) суть те основные математические объекты, которые соответствуют различным физическим величинам.

Установив, вектором или тензором является та или иная величина, мы тем самым задаем закон преобразования величины при лоренцевских преобразованиях, а так как иных математических объектов, связанных с преобразованиями Лоренца, как доказывается в математике, не существует, то мы получаем возможность расклассифицировать все физические величины по закону их преобразования при лоренцевских преобразованиях.

Но дело этим не исчерпывается. Так как все инерциальные системы отсчета эквивалентны, то законы природы должны выглядеть одинаково во всех системах отсчета. Если выбрать какую-либо определенную систему отсчета, то любой физический закон может быть сформулирован в этой системе в виде равенства $T = 0$, где T — некоторая величина или совокупность величин. Так как законы природы должны иметь одинаковую формулировку во всех системах отсчета, то аналогичное равенство должно иметь место и в других системах отсчета. Так как все физические величины, как мы только что разъяснили, имеют векторную или тензорную природу, то величина T должна быть 4-вектором или 4-тензором (или 4-спинором), но в этом случае инвариантность закона природы достигается автоматически; действительно, из линейности закона преобразования тензора (и вектора) вытекает, что если компоненты тензора равны нулю в какой-либо одной системе отсчета, то они равны нулю и в любой другой системе отсчета, т. е. соотношение $T = 0$ влечет за собой соотношение $T' = 0$.

Иными словами, введя понятие 4-векторов и 4-тензоров, мы должны придать каждому закону природы четырехмерную формулировку — векторную или тензорную, и тогда мы можем быть убеждены в том, что эта формулировка инвариантна относительно преобразований Лоренца. Только такая формулировка законов природы (ее называют *ковариантной*) может быть правильной, ибо она удовлетворяет требованию эквивалентности всех систем отсчета.

Обратное утверждение, однако, не верно, т. е. не любое тензорное равенство $T = 0$ выражает собой истинный закон природы.

Посмотрим теперь, как обстоит дело с законами электромагнетизма, т. е. с уравнениями Максвелла и с законом движения частицы в релятивистской и классической динамике. Начнем с уравнений Максвелла.

Вторая пара уравнений Максвелла

$$\operatorname{rot} \mathbf{B} = \frac{4\pi}{c} \mathbf{j} + \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t}; \quad \operatorname{div} \mathbf{E} = 4\pi\rho$$

может быть с помощью тензора поля F_{ih} записана в виде

$$\frac{\partial F_{ih}}{\partial x_h} = -\frac{4\pi}{c} j_i, \quad i = 1, 2, 3, 4,$$

где $j_i \equiv (\mathbf{j}, ic\rho)$ — плотность 4-тока. Таким образом, второй паре уравнений Максвелла может быть придана четырехмерная форма.

Первая пара уравнений Максвелла

$$\operatorname{rot} \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t}, \quad \operatorname{div} \mathbf{B} = 0$$

может быть записана в виде

$$\frac{\partial F_{ik}}{\partial x_l} + \frac{\partial F_{kl}}{\partial x_i} + \frac{\partial F_{li}}{\partial x_k} = 0,$$

где i, k, l принимают значения 1, 2, 3, 4. Но левая часть равенства представляет собой тензор 3-го ранга, т. е. первой паре уравнений Максвелла также может быть придана четырехмерная форма.

Мы видим, что вторая пара уравнений Максвелла выражает обращение в нуль некоторого 4-вектора, а первая пара — обращение в нуль некоторого 4-тензора 3-го ранга. Но обращение в нуль 4-тензора (или 4-вектора) не зависит от выбора системы отсчета.

Мы доказали, таким образом, инвариантность уравнений Максвелла относительно преобразований Лоренца.

Перейдем теперь к рассмотрению уравнений движения заряженной частицы в электромагнитном поле. Если \mathbf{p} и \mathcal{E} — импульс и энергия частицы, то, как мы видели в § 7.6,

$$\frac{d\mathbf{p}}{dt} = e \left\{ \mathbf{E} + \frac{1}{c} [\mathbf{vB}] \right\}, \quad \frac{d\mathcal{E}}{dt} = e\mathbf{E}\mathbf{v}.$$

Этим уравнениям можно придать четырехмерную форму. Действительно, вводя 4-вектор импульса $p_i \equiv [\mathbf{p}, \mathcal{E}/(ic)]$ и 4-вектор скорости частицы $u_i = dx_i/d\tau$ и вспоминая определение тензора поля, можно переписать уравнения движения в виде

$$\frac{dp_i}{d\tau} = eF_{ik}u_k; \quad i, k = 1, 2, 3, 4,$$

а так как $F_{ik}u_k$ представляет собой 4-вектор, то написанное уравнение выражает равенство нулю 4-вектора $dp_i/d\tau - eF_{ik}u_k$. Поэтому это уравнение инвариантно относительно преобразований Лоренца. Тем самым доказана инвариантность исходных уравнений движения в релятивистской динамике относительно преобразований Лоренца. Заметим, однако, что такой инвариантностью не обладают уравнения движения классической механики Ньютона, которые инвариантны только относительно преобразований Галилея, но не инвариантны относительно преобразований Лоренца. Именно по этой причине они и не являются правильными. Точнее можно сказать, что они правильны только при малых скоростях $v \ll c$, в то время как уравнения релятивистской динамики Эйнштейна справедливы при любых скоростях.

7.8. Эффект Доплера

Преобразование 4-векторов [находит важное применение в теории эффекта Доплера.

Рассмотрим электромагнитное поле, имеющее вид плоской монохроматической волны, когда все полевые величины $f(\mathbf{r}, t)$ зависят от координат и времени, как $\exp[i(\mathbf{kr} - \omega t)]$, т. е. имеют вид $f(\mathbf{r}, t) = Fe^{i\varphi}$,

где F — амплитуда поля, \mathbf{k} и ω — волновой вектор и частота волны и $\varphi = \mathbf{k}\mathbf{r} - \omega t$ — фаза. В случае поля в вакууме, который мы будем рассматривать здесь, $\omega = ck$.

При преобразованиях Лоренца амплитуда F преобразуется так же, как и само поле f . Поэтому фаза волны φ должна быть инвариантом, или скаляром.

Отсюда можно получить важные заключения о преобразовании величин \mathbf{k} и ω при переходе от одной системы отсчета к другой. Действительно, запишем фазу в виде

$$\varphi = k_j x_j,$$

где $x_j \equiv (\mathbf{r}, ict)$ и k_j объединяет четыре величины \mathbf{k} и $i\omega/c$. Так как φ — скаляр и x_j — 4-вектор, то k_j представляет собой 4-вектор. Поскольку он объединяет волновой вектор \mathbf{k} и частоту ω , его называют *волновым 4-вектором*.

Отсюда следует, что при преобразованиях Лоренца волновой вектор \mathbf{k} преобразуется как пространственный радиус-вектор \mathbf{r} какого-либо события, а частота ω — как время t .

Пусть в системе отсчета K_0 волновой вектор и частота равны \mathbf{k}_0 и ω_0 . Тогда в системе K , движущейся относительно K_0 со скоростью \mathbf{V} ,

$$\mathbf{k} = \mathbf{k}_0 + \frac{\omega_0}{c} \frac{\mathbf{V}}{V} \left(\cos \alpha \sqrt{1 - V^2/c^2} - \cos \alpha + \frac{V}{c} \right) / \left(1 - \frac{V}{c} \cos \alpha \right);$$

$$\omega = \omega_0 \sqrt{1 - V^2/c^2} / \left(1 - \frac{V}{c} \cos \alpha \right),$$

где α — угол (в системе K) между волновым вектором \mathbf{k} и скоростью \mathbf{V} . Из первой формулы следует

$$\cos \alpha_0 = (\cos \alpha - V/c) / (1 - V \cos \alpha / c),$$

где α_0 — угол (в системе K_0) между волновым вектором \mathbf{k}_0 и скоростью \mathbf{V} . Эта формула определяет изменение направления распространения света при переходе из одной инерциальной системы отсчета в другую. Это явление — так называемая *абберация света* — было уже нами рассмотрено в § 7.3.

Вторая формула определяет изменение частоты света при переходе из одной системы отсчета в другую. Это явление называют *эффектом Доплера*. Если $V \ll c$, то изменение частоты определяется формулой

$$\omega \cong \omega_0 (1 + V \cos \alpha / c).$$

7.9. Принцип эквивалентности

В предыдущих параграфах мы всегда пользовались только инерциальными системами отсчета. Но существует и бесчисленное множество неинерциальных систем отсчета, т. е. систем, движущихся с ускорением относительно инерциальных систем, и ими также можно пользоваться при описании механических движений (что

мы, кстати, и делаем, используя в качестве системы отсчета Землю, которая не является инерциальной системой, так как она вращается вокруг своей оси, вокруг Солнца и вокруг центра нашей Галактики).

Правда, инерциальные системы имеют некоторое преимущество перед неинерциальными системами — свободное тело движется в них прямолинейно и равномерно, чего нельзя сказать о движении свободного тела в неинерциальных системах.

Простота движения в инерциальных системах связана с тем, что в неинерциальных системах возникают специфические силы, связанные с неинерциальностью системы, например центробежная сила инерции или сила Кориолиса, которых нет в инерциальных системах отсчета.

Важнейшей особенностью сил инерции является то, что ускорения, вызываемые ими, не зависят от массы движущегося тела. Благодаря этому возникает аналогия, которая переходит в глубокую физическую связь между неинерциальными системами и гравитационными полями.

Чтобы разъяснить эту аналогию, рассмотрим неинерциальную систему, движущуюся прямолинейно и равномерно ускоренно относительно инерциальной системы. В ней возникнет сила инерции, равная $-ma$, где a — ускорение системы и m — масса тела (какого угодно), находящегося в системе. Если теперь рассмотреть постоянное и однородное гравитационное поле с ускорением g , то тело испытывает в нем действие силы mg . Поэтому если $g = -a$, то движение тела (при наличии каких угодно добавочных сил), в неинерциальной системе происходит точно так же, как и движение тела в инерциальной системе при наличии в ней гравитационного поля g .

Эту *неотличимость силы инерции, возникающей в неинерциальной системе, движущейся прямолинейно и равномерно ускоренно, от однородного гравитационного поля* называют *принципом эквивалентности*. Примером, иллюстрирующим принцип эквивалентности, может служить движение в лифте. Если лифт свободно падает, что на все находящиеся в нем тела сила тяжести не действует, т. е. возникнет состояние *невесомости*; напротив, если лифт ускоренно движется вверх, то сила инерции складывается с силой тяжести, т. е. сила тяжести в лифте как бы увеличится.

Однако принцип эквивалентности, является более универсальным, чем кажется на первый взгляд. Дело в том, что

не только механические движения, но и все другие физические процессы протекают совершенно одинаково как в неинерциальной системе отсчета, движущейся с постоянным ускорением, так и в инерциальной системе отсчета, в которой действует однородное гравитационное поле.

Отсюда можно заключить, что невозможно последовательно пользоваться неинерциальными системами отсчета, пренебрегая гравитационными силами.

Если в классической механике Ньютона ввести в уравнение движения силы инерции, то после этого можно забыть о том, что мы используем неинерциальную систему отсчета: после учета сил инерции неинерциальная система не отличается от инерциальной.

Посмотрим теперь, столь ли безобиден переход от инерциальных систем к неинерциальным в теории относительности. Рассмотрим с этой целью диск, который может вращаться вокруг оси, перпендикулярной его плоскости и проходящей через центр диска (ось z). Имея масштабную линейку размера a , сделаем на периферии диска и одном из его радиусов отметки на равном расстоянии a одна от другой. Отношение чисел отметок на периферии и по радиусу при малом a близко к 2π .

Пусть теперь диск начинает равномерно вращаться вокруг оси z и наблюдатель, находящийся на диске, производит измерение длины окружности диска и его радиуса своей прежней масштабной линейкой. Приложив линейку вдоль окружности, наблюдатель обнаружит, что она укладывается вдоль периферии большее число раз, т. е. он получит больше отметок, чем в том случае, когда диск покоился. Число же отметок вдоль радиуса не изменится. Произойдет это вследствие лоренцевского сокращения длины линейки, если она направлена вдоль движения. Так как лоренцевское сокращение определяется величиной $\sqrt{1-v^2/c^2}$, где $v = \omega R$ — скорость периферии диска (R — его радиус, ω — угловая скорость вращения), то наблюдатель, находящийся на вращающемся диске, придет к выводу, что отношение длины окружности (периферии диска) к ее радиусу равно не 2π , как было в случае покоящегося диска, а равно $2\pi/\sqrt{1-v^2/c^2}$.

Вращающийся диск представляет собой неинерциальную систему отсчета. Поэтому мы приходим к выводу, что

в неинерциальных системах пространство не является евклидовым.

Но неинерциальные системы эквивалентны некоторым гравитационным полям. Поэтому

в гравитационном поле пространство перестает быть евклидовым. Это значит, что геометрические свойства пространства не могут быть чем-то абсолютным даже в смысле специальной теории относительности, а должны зависеть от гравитационных полей.

С другой стороны, гравитационные поля создаются материей. Поэтому мы приходим к выводу, что геометрия пространства зависит, а точнее говоря, определяется материей. Иными словами, геометрия пространства теряет свое самостоятельное существование и становится как бы частью физики.

Пространство, как мы знаем, теснейшим образом связано с временем. Поэтому можно сказать, что

не только свойства пространства, но и свойства времени определяются в конечном счете материей — ее характером и структурой, распределением в пространстве и состоянием движения.

В этом заключается основная идея общей теории относительности — великой физической теории, созданной Эйнштейном в 1915 г.

7.10. Связь метрики пространства — времени с материей

Прежде чем формулировать закон, устанавливающий связь между свойствами пространства — времени и состоянием материи, разъясним, как обобщается понятие интервала на случай неинерциальных систем.

Воспользуемся для этого принципом эквивалентности, согласно которому в каждой мировой точке, точнее говоря, в небольшой окрестности каждой мировой точки P , можно ввести такую систему отсчета, в которой не действуют гравитационные силы. Эта система, которую можно назвать *свободно падающей* (будем называть ее K_0), является локальной, а не глобальной, т. е. в каждой мировой точке существует своя такая система. Ее можно считать инерциальной, и поэтому можно определить в ней квадрат интервала по известной нам формуле

$$ds^2 = dX_1^2 + dX_2^2 + dX_3^2 + dX_4^2,$$

где X_i обозначают координаты мировой точки в этой системе ($X_4 = = iX_0$, X_0 вещественно). Дифференциалы координат определяются с помощью обычных эталонных масштабов и часов, используемых в специальной теории относительности.

Перейдем теперь от K_0 к другой произвольной системе отсчета, вообще говоря, неинерциальной. В ней мировая точка P определяется некоторыми четырьмя координатами x_1, x_2, x_3, x_4 ($x_4 = ix_0$; x_0 вещественно), которые могут быть выбраны совершенно произвольно, что соответствует произвольности выбранной неинерциальной системы. Но существует функциональная связь между координатами X_i и x_i : $X_i = X_i(x_i)$. Отсюда следует, что $dX_i = dx_j \partial X_i / \partial x_j$, поэтому квадрат интервала ds^2 может быть выражен через дифференциалы новых координат dx_i :

$$ds^2 = g_{ih} dx_i dx_h,$$

где $g_{ih} = \frac{\partial X_i}{\partial x_j} \frac{\partial X_i}{\partial x_h}$ (здесь и далее по дважды встречающемуся индексу подразумевается суммирование от 1 до 4).

Координаты x_i можно, как мы говорили, выбирать произвольным образом. Поэтому можно от координат x_i перейти к новым координатам x'_i , что соответствует новой неинерциальной системе; при этом величина ds^2 останется неизменной, т. е. инвариантной:

$$ds^2 = g'_{ih} dx'_i dx'_h,$$

где g'_{ih} определяется через X_i и x'_i , так же как g_{ih} — через X_i и x_i .

Величины g_{ih} зависят от координат исходной точки P , т. е. от X_i или x_i . Они образуют тензор второго ранга, который называют *метрическим тензором*. Используя терминологию дифференциальной геометрии, говорят, что тензор g_{ih} определяет метрику пространства — времени.

На основании результатов предыдущего параграфа можно сказать, что при наличии гравитационного поля метрика пространства неевклидова. Более того, в общем случае произвольного переменного гравитационного поля метрика пространства не только неевклидова, но еще и меняется со временем, т. е. со временем меняются соотношения между различными расстояниями, а следовательно, и взаимное расположение внесенных в поле тел не остается неизменным. Отсюда следует, что

в общем случае переменного гравитационного поля теряет смысл понятие системы отсчета, которым мы пользовались в специальной теории относительности, где под системой отсчета понималась некоторая совокупность покоящихся друг относительно друга тел и совокупность часов, одинаково идущих в пределах системы отсчета. Теперь мы должны пользоваться, как уже говорилось при разъяснении понятия интервала, в каждой мировой точке своей локальной системой отсчета, имеющей свойства системы отсчета специальной теории относительности. Иными словами, теперь мы должны иметь бесконечно много локальных систем отсчета.

В специальной теории относительности использовались только инерциальные системы отсчета, причем основным требованием теории было требование инвариантности законов природы при переходе от одной системы отсчета к другой. Иными словами, требовалось, чтобы все уравнения физики были инвариантными относительно преобразований Лоренца. В общей теории относительности допускаются любые четырехмерные системы координат, и при этом требуется, чтобы законы природы формулировались единым образом для всех четырехмерных систем координат, т. е. требуется, чтобы общая формулировка законов природы не зависела от выбора системы координат. Именно по этой причине эту теорию и называют общей теорией относительности.

Не следует, однако, думать, что различные системы координат физически эквивалентны. Напротив,

физические явления могут протекать в разных системах отсчета различно.

Рассмотрим подробнее, как происходит движение частицы в разных системах координат. В специальной теории относительности в отсутствие электромагнитного поля 4-скорость частицы $u_i = dx_i/dt$ постоянна. Чтобы выяснить, как модифицируется этот результат в общей теории относительности, нужно связать его со структурой интервала ds или дифференциала собственного времени $dt = ds/(ic)$. Введем с этой целью понятие *действия*

$$J = \int_1^2 ds,$$

где интегрирование совершается по произвольной мировой линии, соединяющей мировые точки 1 и 2. Если точки закрепить, то интеграл зависит только от мировой линии, или, как говорят, является *функционалом мировой линии*.

Представим себе теперь, что мы подвергли эту линию бесконечно малой вариации. Действие тогда изменится и его вариация

$$\delta J = \int_1^2 \delta ds.$$

Замечая, что

$$\delta ds = \delta(ds^2)/(2ds) = dx_i \delta dx_i / ds,$$

и учитывая, что $\delta dx_i = d\delta x_i$, имеем, интегрируя по частям,

$$\delta J = \int_1^2 \frac{dx_i}{ds} d\delta x_i = \frac{dx_i}{ds} \delta x_i \Big|_1^2 - \int_1^2 d \frac{dx_i}{ds} \delta x_i.$$

Первое слагаемое здесь равно нулю, так как точки 1 и 2 закреплены, и, следовательно, $(\delta x_i)_1 = (\delta x_i)_2 = 0$. Таким образом,

$$\delta J = - \int_1^2 \frac{d^2 x_i}{ds^2} \delta x_i ds.$$

Постоянство u_i означает, что $d^2 x_i / ds^2 = 0$ (напомним, что $dt = ds/(ic)$). Поэтому, чтобы получить постоянство 4-скорости для свободной частицы, нужно считать, что *действие для истинного движения экстремально*, т. е. вариация его по отношению к мировой линии, соответствующей действительному движению частицы, равна нулю:

$$\delta J = 0.$$

Перейдем теперь к общей теории относительности. Так как в каждой мировой точке локально справедлива специальная теория относительности, то для установления закона движения частицы в отсутствие электромагнитного поля (и других отличных от гравитационного полей) мы можем воспользоваться свойством экстремальности действия для истинного движения, или, как говорят иначе, в а р и а ц и о н н ы м п р и н ц и п о м:

$$\delta J = \delta \int_1^2 ds = 0,$$

где теперь $ds^2 = g_{ih} dx_i dx_h$.

Не приводя здесь выкладок, сформулируем сразу ответ. Уравнения $d^2 x_i / dt^2 = 0$ заменяются теперь уравнениями

$$\frac{d^2 x_i}{dt^2} = - \Gamma_{kl}^i \frac{dx_k}{dt} \frac{dx_l}{dt},$$

где

$$\Gamma_{kl}^i = \frac{1}{2} g^{im} \left(\frac{\partial g_{mk}}{\partial x_l} + \frac{\partial g_{ml}}{\partial x_k} - \frac{\partial g_{kl}}{\partial x_m} \right)$$

и $g_{ih} g^{hl} = \delta_{il}$. Это и есть уравнение движения в гравитационном поле.

Производная $d^2 x_i / dt^2$ может рассматриваться как 4-ускорение частицы, а следовательно, $-m \Gamma_{kl}^i \frac{dx_k}{dt} \frac{dx_l}{dt}$ — как 4-сила, дей-

ствующая в гравитационном поле. Величины Γ_{hl}^i носят название *символов Кристоффеля*.

Метрический тензор g_{ik} играет, следовательно, роль потенциалов поля, а Γ_{hl}^i — напряженности поля.

Заметим, что

соответствующим выбором системы координат можно всегда обратить в нуль все символы Кристоффеля в любой заданной мировой точке. Такой выбор означает, что мы перешли в окрестности рассматриваемой точки к локально инерциальной системе отсчета, в которой не действует гравитационное поле. Возможность такого выбора системы координат и выражает математически принцип эквивалентности.

Мировые линии, определяемые вариационным принципом, носят название *геодезических линий* — 4-пространства с метрикой g_{ik} .

Геодезическая линия в «искривленном» пространстве заменяет прямую в обычном плоском пространстве. В теории гравитации Ньютона движение частицы в обычном плоском пространстве происходит по сложной траектории, которая, например, в однородном поле имеет вид параболы. Именно тот факт, что частица движется не по прямой, указывает на то, что на частицу действует сила, в данном случае гравитационная.

В общей теории относительности движение происходит по геодезической линии в искривленном пространстве—времени, и мы не говорим ни о какой специальной силе, а считаем лишь, что гравитация изменила метрику пространства—времени. В этом смысле

тяготение в теории Эйнштейна является такой же «кажущейся» силой, как центробежная сила и сила Кориолиса в теории Ньютона.

Рассмотрим в качестве примера движение частицы в том случае, когда метрический тензор g_{ik} мало отличается от единичного тензора и имеет вид

$$g_{11} = g_{22} = g_{33} = 1; \quad g_{44} = 1 + h,$$

где h — малая величина, зависящая от координат, $|h| \ll 1$, так что квадрат интервала определяется формулой

$$ds^2 = dx_1^2 + dx_2^2 + dx_3^2 + (1 + h) dx_4^2 = dr^2 - (c^2 + 2\varphi) dt^2,$$

где $dx_4 = icdt$ и $\varphi = c^2 h/2$.

Полагая $dr = vdt$, где v — скорость частицы, получим для дифференциала собственного времени выражение

$$d\tau = ds/(ic) = \sqrt{1 - \frac{v^2}{c^2} + h} dt \approx \left(1 - \frac{v^2}{2c^2} + \frac{h}{2}\right) dt.$$

Поэтому вариационный принцип приобретает вид

$$\frac{1}{ic} \delta J \equiv \delta \int_1^2 d\tau = \delta \int_1^2 \left(1 - \frac{v^2}{2c^2} + \frac{h}{2}\right) dt = 0.$$

Отсюда следует, что

$$\int_1^2 \left(-\mathbf{v} \frac{d}{dt} \delta \mathbf{r} + \frac{\partial \varphi}{\partial \mathbf{r}} \delta \mathbf{r} \right) dt = 0.$$

Интегрируя первое слагаемое по частям и учитывая, что вариация обращается в нуль на границах 1 и 2, получим

$$\int_1^2 \left(\dot{\mathbf{v}} + \frac{\partial \varphi}{\partial \mathbf{r}} \right) \delta \mathbf{r} dt = 0.$$

Так как вариация $\delta \mathbf{r}$ произвольна, то выражение в скобках должно обращаться в нуль, т. е.

$$\dot{\mathbf{v}} = -\text{grad } \varphi.$$

Таким образом, мы получили уравнение движения Ньютона, в котором гравитационная сила равна $-m \text{grad } \varphi$ (m — масса частицы), т. е. φ представляет собой потенциал гравитационного поля. Мы видим, что в слабом гравитационном поле

$$g_{44} = 1 + 2\varphi/c^2.$$

До сих пор мы считали метрический тензор g_{ik} заданной функцией координат x_i . Между тем сама эта функция, т. е. метрика пространства—времени, определяется, как мы уже говорили, материей — ее структурой, распределением и движением.

Суть дела заключается в том, что оказываются связанными между собой два тензора: *тензор кривизны* пространства—времени, определяемый метрическим тензором g_{ik} , и *тензор энергии — импульса* материи.

Тензор кривизны R_{ik} 4-пространства определяется следующей универсальной формулой:

$$R_{ik} = \frac{\partial \Gamma_{ik}^l}{\partial x_l} - \frac{\partial \Gamma_{il}^l}{\partial x_k} + \Gamma_{ik}^l \Gamma_{lm}^m - \Gamma_{il}^m \Gamma_{km}^l,$$

где Γ_{kl}^i — знакомые уже нам символы Кристоффеля. В них входят компоненты g_{ik} и их первые производные по координатам. Поэтому в тензор кривизны входят как первые, так и вторые производные от метрического тензора по координатам. Так как тензор R_{ik} определяется только тензором g_{ik} , то можно сказать, что тензор кривизны представляет собой чисто геометрическое понятие.

Тензор энергии—импульса материи T_{ik} не является столь универсальным, как тензор кривизны. Дело в том, что его структура зависит от конкретных свойств материи, «живущей» в пространстве—времени, и для него не может быть написано такое же общее выражение, как для тензора R_{ik} . В простейшем случае, когда материю можно описывать чисто гидродинамически как некоторую сплошную среду, тензор T_{ik} определяется формулой

$$T_{ik} = (p + \varepsilon) u_i u_k + p g_{ik}$$

где p — давление, ε — плотность энергии среды и $u_i = dx_i/dt$ — 4-вектор скорости.

Сформулируем теперь общий закон, устанавливающий связь между метрикой пространства—времени и состоянием материи. Закон этот состоит в том, что *тензор кривизны пространства—времени является линейной функцией тензора энергии—импульса материи*. Математическая его формулировка такова:

$$R_{ik} = \frac{8\pi G}{c^4} \left(T_{ik} - \frac{1}{2} g_{ik} T \right),$$

где $T = T_{ik} g^{ik}$ (g^{ik} — тензор, обратный g_{ik} , т. е. $g_{ik} g^{kj} = \delta_{ij}$) и G — *гравитационная постоянная*. Это та же постоянная, которая входит в закон всемирного тяготения Ньютона:

$$f = -G \frac{m_1 m_2}{r^2},$$

где m_1 и m_2 — массы двух тел, находящихся на расстоянии r друг от друга.

Закон, устанавливающий связь между тензором кривизны и тензором энергии — импульса, был открыт Эйнштейном в 1915 г. Написанная его формулировка носит название *уравнения Эйнштейна*.

Из этого закона в нерелятивистском случае вытекает закон тяготения Ньютона. В этом случае $g_{11} = g_{22} = g_{33} = 1$, $g_{44} = 1 + 2\varphi/c^2$, где φ , как мы видели выше, гравитационный потенциал. Вычисление компонент тензора кривизны с этими значениями g_{ik} приводит к результату

$$R_{44} = \Delta\varphi/c^2$$

(все остальные компоненты R_{ik} равны нулю).

Из компонент тензора энергии — импульса в нерелятивистском случае (и при нулевом давлении p) отлична от нуля только компонента $T_{44} = -\mu c^2$, где μ — плотность массы. Поэтому уравнение Эйнштейна приобретает вид

$$\Delta\varphi = 4\pi G\mu.$$

Но это и есть уравнение теории тяготения Ньютона для гравитационного потенциала. Оно аналогично уравнению Пуассона для электростатического потенциала, и так же, как из последнего вытекает закон Кулона, так из этого уравнения вытекает закон всемирного тяготения Ньютона.

В заключение этого параграфа остановимся еще на метрике пространства—времени в том случае, когда гравитационное поле обладает центральной симметрией, т. е. создается телом с центрально-симметричным распределением вещества (тело может быть любым). В этом случае уравнение Эйнштейна допускает точное решение и приводит к следующему выражению для квадрата интервала вне тела:

$$ds^2 = dr^2/(1 - r_g/r) + r^2(\sin^2\theta d\varphi^2 + d\theta^2) - (1 - r_g/r)c^2 dt^2,$$

где $r_g = 2GM/c^2$ (M — масса тела). Эту формулу называют метрикой Шварцшильда по имени автора, получившего ее в 1916 г. При $r \gg r_g$ эта метрика переходит в метрику специальной теории относительности (метрика Минковского). Величину r_g называют *гравитационным радиусом* тела. Для Земли он равен 0,887 см, для Солнца — 2,95 км, для Галактики — $4,13 \cdot 10^{11}$ км.

7.11. Красное смещение

Общая теория относительности предсказывает ряд замечательных явлений, из которых мы рассмотрим здесь два — *красное смещение* и *отклонение света в гравитационном поле*. Первое из них состоит в том, что

частота света, распространяющегося в гравитационном поле, не одинакова вдоль его пути — она больше там, где поле сильнее. Поэтому, например, свет, пришедший к нам с какой-либо звезды, имеет меньшую частоту, чем на звезде.

Второе явление заключается в том, что

луч света искривляется в гравитационном поле массивного тела.

Чтобы разъяснить красное смещение, установим сперва связь между временной координатой x_4 , или, как мы будем говорить, *мировым временем* $t = x_4/(ic)$, и собственным, или истинным, временем τ какого-либо тела. Для этого нужно рассмотреть два бесконечно близких события, происходящих в месте нахождения тела. Интервал между ними ds (с точностью до множителя $1/(ic)$) и будет тогда определять промежуток $d\tau$ истинного, или собственного, времени между этими событиями для рассматриваемого тела. Полагая в выражении для $ds^2 = g_{ih} dx_i dx_h$ дифференциалы пространственных координат равными нулю ($dx_1 = dx_2 = dx_3 = 0$), получим следующую связь между дифференциалами истинного ($d\tau$) и мирового (dt) времени:

$$d\tau = \sqrt{g_{44}} dt,$$

откуда

$$\tau = \int \sqrt{g_{44}} dt,$$

где g_{44} зависит, вообще говоря, как от t , так и от места нахождения тела.

В статическом гравитационном поле g_{44} от времени не зависит и

$$\tau = \sqrt{g_{44}} t.$$

Выше мы видели, что в случае слабого гравитационного поля $g_{44} = 1 + 2\varphi/c^2$, где φ — потенциал поля. Если при этом φ не зависит от времени, то

$$\tau = (1 + \varphi/c^2)t, \quad |\varphi/c^2| \ll 1.$$

Эта формула показывает, что собственное, или истинное, время течет быстрее в тех местах пространства, где больше гравитационный потенциал. Если $\varphi < 0$ (что обычно имеет место), то собственное время течет медленнее там, где больше модуль потенциала. Из связи между τ и t вытекает следующий замечательный вывод:

если из двух одинаковых часов одни находились некоторое время в гравитационном поле $\varphi < 0$, то после этого они окажутся отставшими по сравнению с часами, в гравитационном поле не находившимися.

Рассмотрим теперь распространение света в постоянном гравитационном поле. Зависимость от времени электромагнитного поля световой волны, распространяющейся в постоянном гравитационном поле, в простейшем случае определяется множителем $\cos[\omega_0 t + \chi(\mathbf{r})]$, где ω_0 — частота, соответствующая мировому времени t , и $\chi(\mathbf{r})$ — фаза, зависящая от пространственных координат. Существенно, что в статическом гравитационном поле частота ω_0 , измеренная по мировому времени, есть константа, не зависящая от пространственных координат.

Однако зависимость поля световой волны от мирового времени не представляет непосредственного физического интереса, так как наблюдатель, находящийся в некоторой точке пространства $\mathbf{r} = \text{const}$, анализирует поле световой волны с точки зрения не мирового времени t , а собственного времени τ . Поэтому в выражении $\cos(\omega_0 t + \chi)$ нужно выразить t через τ . Мы получим тогда $\cos(\omega\tau + \chi)$, где

$$\omega = \omega_0 t / \tau = \omega_0 / \sqrt{g_{44}}$$

или в случае слабого гравитационного поля

$$\omega = \omega_0 (1 - \varphi/c^2)$$

Эта формула определяет частоту света ω , измеренную в собственном времени в точке пространства, где потенциал гравитационного поля равен φ . Она показывает, что частота, измеренная в собственном времени, не одинакова вдоль светового луча (в отличие от постоянной частоты ω_0 !); частота больше в тех местах, где больше модуль гравитационного потенциала φ . Иными словами,

частота ω увеличивается при приближении луча к создающим гравитационное поле телам и уменьшается при удалении от них.

Если свет испускается атомом в месте, где гравитационный потенциал равен φ_1 , и поглощается в месте, где гравитационный потенциал равен φ_2 , то частоты света (измеренные в собственном времени) в этих местах различны и составляют соответственно

$$\omega_1 = \omega_0 (1 - \varphi_1/c^2), \quad \omega_2 = \omega_0 (1 - \varphi_2/c^2)$$

Исключая отсюда ω_0 и учитывая, что $|\varphi_{1,2}|/c^2 \ll 1$, получим

$$\omega_2 = \omega_1 \left(1 + \frac{\varphi_1 - \varphi_2}{c^2} \right),$$

$$(\omega_2 - \omega_1)/\omega_1 = (\varphi_1 - \varphi_2)/c^2.$$

В частности, если свет «рождается» на Солнце или звезде с частотой ω_1 и приходит на Землю, то частота пришедшего света ω_2 меньше, чем ω_1 , так как $\varphi_1 - \varphi_2 < 0$.

Мы до сих пор говорили об изменении частоты луча света при движении его в гравитационном поле, так что речь у нас шла об одном излучающем атоме.

Пусть теперь у нас имеются два одинаковых излучающих атома — один на Солнце или звезде, а другой на Земле. Тогда их спектры излучения, наблюдаемые соответственно на Солнце и на Земле, выглядят совершенно одинаково. Но когда световой луч, испущенный атомом, находящимся на Солнце, достигнет Земли, то его частота не будет уже совпадать с частотой света, испущенного тождественным атомом, находящимся на Земле. Частота пришедшей волны будет меньше частоты волны, испущенной земным атомом. Разность частот — пришедшей и земной — определится формулой

$$\omega/\omega = (\varphi_1 - \varphi_2)/c^2,$$

где φ_1 и φ_2 — гравитационные потенциалы на Солнце и Земле.

Так как $\varphi_1 - \varphi_2 < 0$, то $\Delta\omega < 0$, т. е. смещение частоты происходит в сторону уменьшения частот. Поэтому это явление и носит название красного смещения.

Красное смещение допускает простую физическую интерпретацию, основанную на корпускулярной картине излучения. Именно: будем представлять себе фотон как частицу с кинетической энергией $\mathcal{E} = \hbar\omega$ и массой $m = \mathcal{E}/c^2$. При движении такой частицы в гравитационном поле с потенциалом φ , которое мы будем предполагать слабым, к кинетической энергии прибавляется потенциальная энергия $m\varphi$, так что полная энергия частицы становится равной $\mathcal{E} + m\varphi = \hbar\omega(1 + \varphi/c^2)$. При движении фотона полная его энергия сохраняется, т. е.

$$\hbar\omega_1(1 + \varphi_1/c^2) = \hbar\omega_2(1 + \varphi_2/c^2),$$

где ω_1 и ω_2 — частота фотона в точках пространства с гравитационными потенциалами φ_1 и φ_2 . Так как $|\varphi|/c^2 \ll 1$, то отсюда следует наша прежняя формула для смещения частоты $\Delta\omega = \omega_2 - \omega_1$:

$$\Delta\omega/\omega = (\varphi_1 - \varphi_2)/c^2.$$

Если гравитационное поле создается телом с центрально-симметричным распределением массы, то на расстоянии r от центра тела потенциал поля

$$\varphi = -GM/r,$$

где M — масса тела и G — гравитационная постоянная. Пусть фотон с частотой ω_1 рождается на поверхности тела $r = R$. Частоту фотона на бесконечном расстоянии от тела обозначим ω_2 . Тогда для относительного смещения частоты мы получим формулу

$$\Delta\omega/\omega = -r_g/(2R),$$

где $r_g = 2GM/c^2$ — гравитационный радиус тела.

Формула для красного смещения прекрасно подтверждается на опыте. Например, для луча света, поднимающегося от поверхности Земли на высоту 20 м, формула для относительного изменения частоты в гравитационном поле Земли совпадает с экспериментальным результатом с точностью до 1%. Красное смещение для лучей, идущих от Солнца, измерено с точностью до 6%.

7.12. Отклонение луча света в гравитационном поле

Перейдем теперь к рассмотрению второго эффекта — отклонения светового луча в гравитационном поле. Качественно этот эффект можно понять, если, так же как и в предыдущем параграфе, исходить из представления о световом луче как совокупности частиц — фотонов, обладающих массой $m = \hbar\omega/c^2$. Пусть, например, такая частица движется в направлении оси x в гравитационном поле, создаваемом шарообразным телом массы M и радиуса R (рис. 7.2). Отклонение частицы вдоль оси y определяется в механике Ньютона уравнением

$$\frac{d^2y}{dt^2} = -\frac{GM}{r^3}y, \quad r = \sqrt{x^2 + y^2}.$$

Считая отклонение малым, можно в правой части положить $y = R$. Учитывая еще, что для фотона $x = ct$, получим

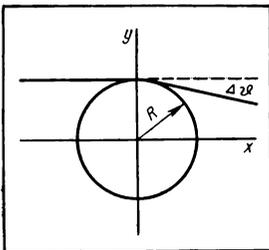
$$\frac{d^2y}{dx^2} = -\frac{GMR}{(R^2 + x^2)^{3/2}}.$$

Интегрируя это уравнение, найдем

$$\frac{dy}{dx} = -\frac{GMx}{c^2R\sqrt{R^2 + x^2}} + \text{const.}$$

При $x = -\infty$ имеем $dy/dx = 0$, поэтому $\text{const} = -GM/(c^2R)$. Отсюда можно найти угол отклонения фотона:

$$\Delta\theta = -\left.\frac{dy}{dx}\right|_{x=\infty} = \frac{2GM}{c^2R}.$$



7.2. Отклонение луча света в гравитационном поле

Истинное отклонение луча, однако, в два раза больше найденного нами в классической теории Ньютона.

Покажем, как находится отклонение луча в гравитационном поле согласно общей теории относительности. Напомним сперва, что если два события лежат на одном и том же световом луче, то в специальной теории относительности интервал между этими событиями равен нулю. В общей теории относительности это свойство сохраняется локально, т. е. для двух бесконечно близких событий, лежащих

на одном и том же световом луче, интервал равен нулю; иными словами, для мировых точек, лежащих на световом луче, выполняется условие

$$ds^2 = g_{ik} dx_i dx_k = 0.$$

Воспользуемся теперь для ds^2 метрикой Шварцшильда (см. § 7.10). Обозначая, как и ранее, гравитационный потенциал тела $\varphi = -GM/r$, получим из условия $ds^2 = 0$ при $|\varphi|/c^2 \ll 1$ соотношение

$$c^2 dt^2 = dr^2 (1 - 2\varphi/c^2) - dr^2 2\varphi/c^2,$$

где $dr^2 = dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2)$ (θ и φ — азимутальный и полярный углы).

Нам нужно теперь найти форму светового луча. Это можно сделать с помощью принципа Ферма (см. § 20.3), согласно которому свет распространяется таким образом, чтобы время прохождения его из одной точки в другую было минимальным.

Это значит, что вариация интеграла $\int_C \frac{dt}{c}$, взятого вдоль контура C , соединяющего две данные точки, т. е. изменение его значения при изменении контура C , равна нулю в том случае, когда C совпадает с лучом света:

$$\delta \int_C dt = 0.$$

Подставляя сюда вместо dt значение, даваемое условием $ds^2 = 0$ при использовании метрики Шварцшильда, мы получим следующее вариационное условие для определения формы луча:

$$\delta \int \sqrt{dr^2 (1 - 2\varphi/c^2) - dr^2 2\varphi/c^2} = 0.$$

Пусть луч распространяется в плоскости (x, y) (рис. 7.2). Тогда

$$dr^2 = dx^2 + dy^2, \quad dr = (xdx + ydy)/r, \quad r^2 = x^2 + y^2$$

и вариационное условие приобретает вид

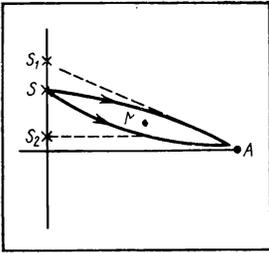
$$\delta \int \sqrt{(dx^2 + dy^2) (1 - 2\varphi/c^2) - (xdx + ydy)^2 2\varphi/(r^2 c^2)} = 0.$$

Учтем, наконец, что свет распространяется вблизи оси x , т. е. угол наклона луча к оси x мал: $|dy/dx| \ll 1$. Вариационное условие при этом сильно упрощается:

$$\delta \int \left[\frac{1}{2} (y')^2 - \frac{\varphi}{c^2} \left(1 + \frac{x^2}{r^2} \right) \right] dx = 0,$$

где $y' = dy/dx$. Переписав его в виде

$$\delta \int L dx = 0, \quad L = \frac{1}{2} (y')^2 - \frac{\varphi}{c^2} \left(1 + \frac{x^2}{r^2} \right),$$



7.3.

Идея гравитационной линзы

легко убедиться, [что из этого условия вытекает уравнение

$$\frac{d}{dx} \frac{\partial L}{\partial y'} - \frac{\partial L}{\partial y} = 0$$

(его называют уравнением Лагранжа — Эйлера). Подставив сюда выражение для L , получим

$$y'' = - \frac{GM}{c^2} \left(\frac{1}{r^3} + \frac{3x^2}{r^5} \right) y.$$

Считая, что форма луча мало отличается от прямой $y = R$ (рис. 7.2), можно справа заменить y на R . В результате мы получим следующую формулу для результирующего изменения величины y' при изменении x от $-\infty$ до ∞ :

$$\Delta y' = - \frac{GMR}{c^2} \int_{-\infty}^{\infty} \frac{4x^2 + R^2}{\sqrt{(x^2 + R^2)^5}} dx.$$

Эта величина и определяет угол отклонения луча $\Delta\theta = |\Delta y'|$. Выполнив интегрирование, получим

$$\Delta\theta = 4GM/(c^2R) = 2r_g/R.$$

Подстановка значений R и r_g для Солнца дает теоретическое значение $\Delta\theta \approx 1,75''$. Экспериментально получено $\Delta\theta = (1,73 \pm 0,05)''$.

Отклонение световых лучей в гравитационном поле лежит в основе эффекта гравитационной линзы. Если, например, свет от источника S (рис. 7.3) распространяется в точку наблюдения A и на пути света находится массивное тело M , то в случае точечного тела наблюдатель в A увидит два изображения источника: S_1 и S_2 . Этот эффект связывается с открытием уникальной пары квазаров — мощных радиосисточников QSO0957 + 561 A, B.

ОСНОВНЫЕ ФОРМУЛЫ

Преобразования Лоренца

$$x' = \frac{x - Vt}{\sqrt{1 - V^2/c^2}},$$

$$t' = \frac{t - Vx/c^2}{\sqrt{1 - V^2/c^2}}$$

Инвариантность интервала

$$ds^2 = ds'^2, \quad ds^2 = dx^2 + dy^2 + dz^2 - c^2 dt^2$$

Собственное время

$$d\tau = dt \sqrt{1 - V^2/c^2}$$

Преобразование скорости

$$v_x = \frac{v'_x + V}{1 + v'_x V / c^2},$$

$$v_y = \frac{v'_y \sqrt{1 - V^2/c^2}}{1 + v'_x V / c^2},$$

$$v_z = \frac{v'_z \sqrt{1 - V^2/c^2}}{1 + v'_x V / c^2}$$

Аберрация света

$$\operatorname{tg} \theta = \frac{\sqrt{1 - V^2/c^2}}{\frac{V}{c} + \cos \theta'} \sin \theta'$$

Преобразование электромагнитных полей

$$\mathbf{E}'_{\perp} = \left(\mathbf{E}_{\perp} + \frac{1}{c} [\mathbf{V} \mathbf{H}_{\perp}] \right) / \sqrt{1 - V^2/c^2},$$

$$\mathbf{H}'_{\perp} = \left(\mathbf{H}_{\perp} - \frac{1}{c} [\mathbf{V} \mathbf{E}_{\perp}] \right) / \sqrt{1 - V^2/c^2},$$

$$\mathbf{E}'_{\parallel} = \mathbf{E}_{\parallel}, \quad \mathbf{H}'_{\parallel} = \mathbf{H}_{\parallel}$$

Инварианты поля

$$E^2 - H^2 = E'^2 - H'^2,$$

$$(\mathbf{E} \mathbf{H})^2 = (\mathbf{E}' \mathbf{H}')^2$$

Поля равномерно движущегося заряда

$$\mathbf{E} = \frac{er}{r^3} \frac{1 - V^2/c^2}{\left(1 - \frac{V^2}{c^2} \sin^2 \theta\right)^{3/2}},$$

$$\mathbf{H} = \frac{1}{c} [\mathbf{V} \mathbf{E}]$$

Импульс и энергия частицы

$$\mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1 - v^2/c^2}},$$

$$\mathcal{E} = \frac{mc^2}{\sqrt{1 - v^2/c^2}} = c \sqrt{m^2 c^2 + p^2}$$

Уравнение движения заряженной частицы в электромагнитном поле

$$\frac{d\mathbf{p}}{dt} = e \left(\mathbf{E} + \frac{1}{c} [\mathbf{v} \mathbf{H}] \right),$$

$$\frac{d\mathcal{E}}{dt} = e\mathbf{v} \mathbf{E}$$

Эффект Доплера

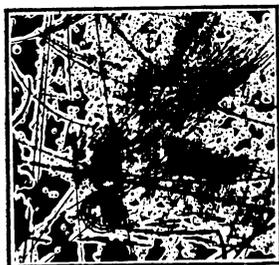
$$\omega = \omega_0 \frac{\sqrt{1 - V^2/c^2}}{1 - (V/c) \cos \theta}$$

Красное смещение

$$\frac{\Delta \omega}{\omega} = \frac{\varphi_1 - \varphi_2}{c^2}$$

Угол отклонения света в гравитационном поле

$$\Delta \theta = \frac{2r_g}{R}, \quad r_g = \frac{2GM}{c^2}$$

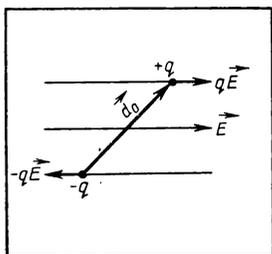


II. ЭЛЕКТРИЧЕСКИЕ И МАГНИТНЫЕ СВОЙСТВА ВЕЩЕСТВА

Глава 8. ДИЭЛЕКТРИКИ

8.1. Полярные и неполярные молекулы

Уменьшение поля в диэлектрике по сравнению с вакуумом (при тех же сторонних зарядах) связано, как мы уже не раз говорили, со смещением зарядов в молекулах диэлектрика под действием поля. Такое смещение зарядов приводит к возникновению у молекул направленного вдоль поля дипольного момента. При этом следует различать два принципиально различных типа молекул. Молекулы одного типа сами по себе в отсутствие поля не обладают дипольным моментом, но дипольный момент возникает у них под действием поля, которое как бы растягивает молекулу и раздвигает ее заряды — положительные по полю и отрицательные против поля. Такие молекулы называют *неполярными*. При не очень сильных полях дипольный момент неполярной молекулы \mathbf{d} оказывается пропорциональным напряженности электрического поля: $\mathbf{d} = \epsilon_0 \alpha' \mathbf{E}$, где α' — некоторая величина, характеризующая молекулу; ее называют *коэффициентом поляризуемости (поляризуемостью)* молекулы.



Молекулы другого типа обладают сами по себе, т. е. даже в отсутствие внешнего электрического поля, дипольным моментом. Такие молекулы, обладающие собственным, или, как говорят иначе, *спонтанным*, дипольным моментом, называют *полярными*. Если полярную молекулу с дипольным моментом \mathbf{d}_0 поместить в электрическое поле \mathbf{E} , то на нее действует пара сил (рис. 8.1) с моментом

$$\mathbf{K} = [\mathbf{d}_0 \mathbf{E}].$$

8.1.

Ориентирующее действие поля на молекулу полярного диэлектрика

Эта пара вращает диполь, стремясь расположить его вдоль поля. Таким образом,

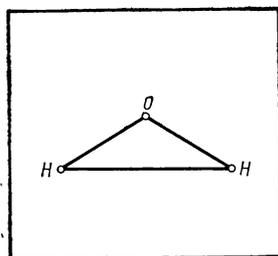
действие поля в случае диэлектрика с полярными молекулами (такой диэлектрик называют *полярным*) является прежде всего ориентирующим — поле стремится ориентировать дипольные моменты отдельных молекул вдоль поля. Кроме того, конечно, полярные молекулы, так же как и неполярные молекулы, могут деформироваться под действием внешнего электрического поля. Но этот эффект является второстепенным по сравнению с ориентационным.

Хотя молекулы полярного диэлектрика и обладают спонтанным дипольным моментом, направления дипольных моментов отдельных молекул у полярных газов, жидкостей и ряда твердых диэлектриков распределены в отсутствие поля хаотически. Поэтому сумма дипольных моментов всех молекул в отсутствие поля равна нулю.

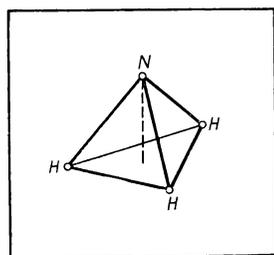
Принадлежность молекулы к тому или другому типу определяется ее структурой и тесно связана с симметрией молекулы. Это иллюстрируют рис. 8.2, 8.3, на которых показано расположение ионов в молекулах воды H_2O и аммиака NH_3 . Обе эти молекулы принадлежат к полярным, так как в них есть избранное направление, вдоль которого и располагается вектор d_0 . Дипольный момент молекулы воды направлен по биссектрисе угла при атоме кислорода (угол этот равен 108°) и составляет $6,2 \cdot 10^{-30}$ Кл·м; дипольный момент молекулы аммиака лежит вдоль высоты пирамиды (пунктир на рис. 8.3) и равен $5,1 \cdot 10^{-30}$ Кл·м. Такого же порядка дипольные моменты большинства полярных молекул, так как размер молекулы $\sim 10^{-10}$ м и единичный заряд иона $1,6 \cdot 10^{-19}$ Кл.

Примеры более симметричных молекул показаны на рис. 8.4. Эти молекулы неполярны, поскольку в них нет избранного направления, вдоль которого мог бы быть расположен собственный дипольный момент. При включении внешнего поля симметрия молекулы нарушается и молекула приобретает дипольный момент.

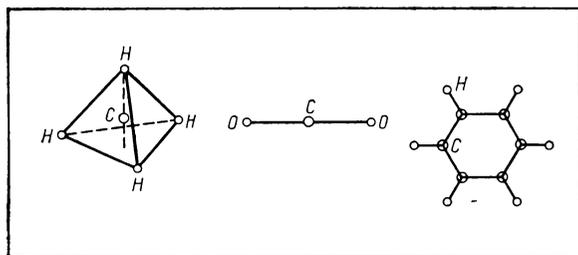
Рассмотрим газ, состоящий из одинаковых неполярных молекул.



8.2.
Молекула воды



8.3.
Молекула аммиака



8.4.
Примеры неполярных молекул

В поле напряженностью E у молекулы возникает дипольный момент $\mathbf{d} = \epsilon_0 \alpha' \mathbf{E}$. Поэтому если n — плотность молекул, то вектор поляризации газа $\mathbf{P} = n\mathbf{d} = \epsilon_0 \alpha \mathbf{E}$, где $\alpha = n\alpha'$ — абсолютная диэлектрическая восприимчивость газа. Как мы видели в гл. 2, диэлектрическая проницаемость ϵ связана с восприимчивостью соотношением $\epsilon = 1 + \alpha$. Мы видим, что из макроскопических величин в выражение для диэлектрической проницаемости неполярного газа входит только плотность газа.

Рассмотрим теперь газ, состоящий из одинаковых полярных молекул. В этом случае, как мы видели, электрическое поле стремится ориентировать собственные дипольные моменты молекул по полю. Однако этой тенденции препятствует тепловое движение молекул. Чтобы учесть оба эффекта — ориентирующее действие поля и разориентирующее действие теплового движения, — нужно найти распределение молекул по направлениям их собственных дипольных моментов.

Для решения этой задачи напомним предварительно, что если газ находится в некотором силовом поле, то в состоянии статистического равновесия пространственное распределение его молекул определяется общей формулой Больцмана

$$n(\mathbf{r}) = n_0 e^{-U(\mathbf{r})/(kT)}.$$

Здесь $n(\mathbf{r})$ — плотность молекул в точке \mathbf{r} ; $U(\mathbf{r})$ — потенциальная энергия молекулы в этой точке; n_0 — плотность молекул в точке, где $U = 0$; T — температура газа; k — постоянная Больцмана. Формулой Больцмана можно воспользоваться, чтобы решить интересующую нас задачу. При этом под U следует понимать потенциальную энергию диполя в поле (см. § 1.3):

$$U = -\mathbf{d}_0 \mathbf{E} = -d_0 E \cos \vartheta,$$

где \mathbf{d}_0 — собственный дипольный момент молекулы и ϑ — угол между дипольным моментом \mathbf{d}_0 и полем \mathbf{E} . Вместо же локальной плотности молекул $n(r)$ следует теперь говорить о числе молекул в единичном телесном угле около данного направления \mathbf{d}_0 . Обозначив dN число молекул полярного газа, дипольные моменты которых лежат в элементе телесного угла $d\Omega$, и учитывая, что $d\Omega = 2\pi \sin \vartheta d\vartheta$, можно, следовательно, написать

$$dN = C \exp [d_0 E \cos \vartheta / (kT)] 2\pi \sin \vartheta d\vartheta,$$

где C — константа, определяемая из условия

$$C \int_0^\pi \exp [d_0 E \cos \vartheta / (kT)] 2\pi \sin \vartheta d\vartheta = N$$

(N — общее число молекул газа).

Определим теперь поляризованность \mathbf{P} газа. Так как пространственное распределение дипольных моментов зависит только от угла ϑ между векторами \mathbf{d}_0 и \mathbf{E} , то составляющие вектора \mathbf{d}_0 , перпендику-

лярные полю для отдельных молекул, взаимно компенсируются. Поэтому вектор \mathbf{P} не имеет составляющей, перпендикулярной полю, параллельная же полю его составляющая $P = nd_0 \langle \cos \vartheta \rangle$, где n — плотность диполей и $\langle \cos \vartheta \rangle$ — среднее значение \cos по всем диполям:

$$\langle \cos \vartheta \rangle = \frac{1}{N} \int \cos \vartheta dN;$$

$$\langle \cos \vartheta \rangle = \frac{2\pi}{N} \int_0^\pi \exp\left(-\frac{d_0 E \cos \vartheta}{kT}\right) \cos \vartheta \sin \vartheta d\vartheta.$$

Обычно $d_0 E / (kT) \ll 1$, поэтому

$$dN = -\frac{N}{2} \left(1 + \frac{d_0 E}{kT} \cos \vartheta\right) d(\cos \vartheta);$$

$$\langle \cos \vartheta \rangle = \frac{1}{2} \int_{-1}^{+1} \left(1 + \frac{d_0 E}{kT} \cos \vartheta\right) \cos \vartheta d(\cos \vartheta) = \frac{d_0 E}{3kT}.$$

Так как $\mathbf{P} = \epsilon_0 \alpha \mathbf{E}$, то для полярного газа

$$\alpha = \frac{nd_0^2}{3\epsilon_0 kT}, \quad \epsilon = 1 + \frac{nd_0^2}{3kT\epsilon_0}.$$

Обратим внимание на то, что диэлектрическая проницаемость газа полярных молекул в отличие от диэлектрической проницаемости газа неполярных молекул зависит от температуры. Поэтому если экспериментально установлено, что диэлектрическая проницаемость диэлектрика зависит от температуры, то это значит, что диэлектрик построен из полярных молекул.

8.2. Диэлектрическая проницаемость плотных газов и жидкостей

В предыдущем параграфе мы определили диэлектрическую проницаемость газов, связав ее с поляризуемостью отдельной молекулы. Решающим было при этом предположение, что поле, действующее на отдельную молекулу, или, как мы будем говорить, локальное поле, совпадает со средним макроскопическим полем, существующим в диэлектрике. Между тем эти величины отличаются друг от друга и различие между ними тем больше, чем больше плотность частиц.

Для того чтобы понять различие между локальным и средним полем, напомним, как мы находили силу, действующую на обкладку плоского конденсатора. На первый взгляд кажется, что эта сила (отнесенная к единичной площади) равна произведению поверхностной плотности заряда σ на напряженность поля $E = \sigma/\epsilon_0$. В действительности же поверхностная плотность силы

$$f = 1/2 \sigma E,$$

так как сила, действующая на заряд обкладки, определяется не общим полем, а только полем, создаваемым второй обкладкой (в месте нахождения заряда). Это поле, имеющее смысл локального поля (по принятой выше терминологии), равно $\sigma/(2\epsilon_0)$, т. е. $E/2$.

Исходя из этой аналогии, можно сказать, что локальное поле в некоторой точке диэлектрика представляет собой поле в небольшой полости, сделанной в диэлектрике около этой точки. В случае газа или жидкости, т. е. изотропных диэлектриков, этой полости естественно придать сферическую форму, напряженность же поля в такой полости, как было показано в § 2.2,

$$\mathbf{E}_l = \frac{3\epsilon}{2\epsilon + 1} \mathbf{E},$$

где \mathbf{E} — напряженность поля вдали от полости (или, что то же самое, напряженность поля, существовавшего в рассматриваемой точке до того, как была вырезана полость). Вспоминая, что $\epsilon = 1 + \alpha$, можно переписать выражение для локального поля \mathbf{E}_l в виде

$$\mathbf{E}_l = E + \frac{\alpha}{3 + 2\alpha} E = \mathbf{E} + \frac{\epsilon_0^{-1}}{3 + 2\alpha} \mathbf{P},$$

где $\mathbf{P} = \epsilon_0 \alpha \mathbf{E}$ — поляризованность. Если $\alpha \ll 1$, то эта формула приобретает вид

$$\mathbf{E}_l = \mathbf{E} + \mathbf{P}/(3\epsilon_0).$$

Она носит название формулы Лоренца для локального поля.

Под действием локального поля отдельная молекула приобретает направленный вдоль поля дипольный момент $\mathbf{d} = \alpha' \epsilon_0 \mathbf{E}_l$, где α' — коэффициент поляризуемости молекулы, зависящий только от свойств молекулы. Если n — плотность молекул, то поляризованность $\mathbf{P} = \alpha_g \epsilon_0 \mathbf{E}_l$, где $\alpha_g = n \alpha'$. С другой стороны, $\mathbf{P} = \alpha \epsilon_0 \mathbf{E}$. Поэтому

$$\alpha_g \left(E + \frac{\alpha}{3 + 2\alpha} E \right) = \alpha E$$

и, следовательно,

$$\alpha = \frac{3}{4} \left(\alpha' n - 1 + \sqrt{1 + \frac{2}{3} \alpha' n + \alpha'^2 n^2} \right).$$

Эта формула (ее называют формулой Онзагера) определяет электрическую восприимчивость α как функцию плотности частиц n при заданном параметре α' — коэффициенте поляризуемости отдельной молекулы.

Величину α' можно найти, если предварительно экспериментально определить электрическую восприимчивость разреженного газа, так как в этом случае диэлектрическая восприимчивость газа равна $\alpha' n$ (к такому результату при малых $\alpha' n$ приводит и формула Онзагера). Поэтому, зная плотность газа, можно вычислить α' . Смысл формулы Онзагера состоит в том, что с помощью этой же величины α'

можно вычислить диэлектрическую восприимчивость плотного газа, а также жидкости (если, конечно, при переходе в конденсированное состояние не изменяется поляризуемость молекулы). Результат такого подсчета приводит, например, к значению 2,76 для диэлектрической проницаемости жидкого CS_2 , тогда как экспериментальное значение этой величины составляет 2,64 (при этом плотность жидкого CS_2 превышает плотность газа в 380 раз).

С точностью до квадратичных по $\alpha'n$ членов формула Онзагера дает $\alpha \approx \alpha'n (1 + \alpha'n/3)$. Последнее выражение (также с точностью до квадратичных членов) можно переписать в виде

$$\alpha = \frac{\alpha'n}{1 - \alpha'n/3}.$$

Эта формула носит название формулы Лоренца — Лоренца. Выразив с помощью нее $\alpha'n$ через α и заменяя α на $(\epsilon - 1)$, получим

$$\frac{\epsilon - 1}{\epsilon + 2} = \alpha'n/3.$$

(формула Клаузиуса — Моссотти).

При использовании двух этих формул следует иметь в виду, что они справедливы только в том случае, когда $\alpha'n/3$ мало по сравнению с единицей. Если $\alpha'n/3 = 1$, то по формуле Лоренца—Лоренца α обращается в бесконечность, тогда как в действительности, согласно формуле Онзагера, $\alpha = 3(1 + \sqrt{3})/2$.

8.3. Диэлектрическая проницаемость кристаллов

Газы и жидкости представляют собой *изотропные среды*. Это значит, что все их свойства, в частности электрические свойства, одинаковы во всех направлениях. Напротив, твердые тела (точнее говоря, монокристаллы) отличаются *анизотропией* в своих свойствах. Иными словами, свойства монокристалла различны в разных направлениях. В частности, поляризуемость кристалла, вообще говоря, не одинакова в разных направлениях. Анизотропия в свойствах монокристалла связана, естественно, с характером его внутренней структуры, т. е. наличием кристаллической решетки.

Анизотропия в поляризуемости кристалла проявляется, в частности, в том, что, если вырезать из монокристалла две пластинки, различно ориентированные относительно кристаллографических осей, и поместить сперва одну, а потом другую между обкладками конденсатора, то емкость конденсатора в обоих случаях, вообще говоря, не одинакова. Отсюда следует, что диэлектрическая проницаемость, а следовательно, и диэлектрическая восприимчивость кристалла различны в разных направлениях. Исключение составляют кристаллы кубической системы, которые ведут себя как изотропные тела.

Если взять, например, кристалл берилла $(\text{SiO}_3)_6\text{Al}_2\text{Be}_3$, принадлежащего к гексагональной системе, то его диэлектрическая прони-

цаемость равна 7,44 и 7,85 в зависимости от того, параллельно или перпендикулярно электрическое поле оси 6-го порядка. Для апатита диэлектрическая проницаемость при поле, направленном вдоль оси, равна 7,4, а при поле, перпендикулярном оси, равна 9,5. Особенно велика разница в значениях диэлектрической проницаемости в разных направлениях у рутила TiO_2 , принадлежащего к тетрагональной системе. Если поле направлено вдоль оси (4-го порядка), то диэлектрическая проницаемость равна 173, если же поле направлено перпендикулярно этой оси, то диэлектрическая проницаемость составляет 89.

Анизотропия поляризуемости характерна для монокристалла. В поликристаллах, состоящих из огромного числа маленьких кристаллитов, произвольно ориентированных друг относительно друга, происходит усреднение поляризуемости, и такие кристаллы ведут себя в электрическом поле как изотропные тела, т. е. как газы и жидкости.

Для газов и жидкостей поляризованность $\mathbf{P} = \epsilon_0 \alpha \mathbf{E}$, как мы видели, пропорциональна и параллельна напряженности поля \mathbf{E} , где α — диэлектрическая восприимчивость. Для диэлектрических монокристаллов эта формула, вообще говоря, неприменима, так как она не учитывает эффекта анизотропии. Кроме того, ряд монокристаллов обладает поляризованностью, не обращающейся в нуль при $E = 0$ и сложно, нелинейным образом зависящей от поля. Эти эффекты мы рассмотрим подробнее в следующем параграфе, здесь же остановимся только на эффекте анизотропии.

Если поляризованность монокристалла \mathbf{P} равна нулю при $E = 0$, то при достаточно слабом поле она является линейной функцией поля. Необходимо при этом иметь в виду, что направление вектора поляризации в монокристалле может и не совпадать с направлением поля. Однако, сколь бы ни была сложна симметрия кристалла, всегда в любом кристалле существуют по крайней мере три взаимно перпендикулярных направления таких, что если направление поля совпадает с одним из них, то такое же направление имеет и вектор поляризации. Это значит, что если оси координат X, Y, Z выбраны вдоль этих направлений (их называют *главными направлениями поляризуемости*) и электрическое поле \mathbf{E} направлено вдоль оси X , т. е. $E_X = E, E_Y = E_Z = 0$, то у вектора поляризации \mathbf{P} отлична от нуля только проекция P_X на ось X , причем с E_X она связана соотношением

$$P_X = \epsilon_0 \alpha_1 E_X,$$

где α_1 — некоторая характерная для кристалла величина, не зависящая от поля (поле считается слабым) и представляющая собой диэлектрическую восприимчивость кристалла вдоль оси X . Аналогичным образом, если поле направлено вдоль оси Y или оси Z , вектор поляризации направлен вдоль оси Y или вдоль оси Z , причем в первом случае $P_Y = \epsilon_0 \alpha_2 E_Y$, а во втором $P_Z = \epsilon_0 \alpha_3 E_Z$, где α_2 и α_3 — диэлектрические восприимчивости кристалла вдоль осей Y и Z . Величины $\alpha_1, \alpha_2, \alpha_3$ называют *главными значениями диэлектрической восприимчивости* (поляризуемости) кристалла.

Рассмотрим теперь тот случай, когда направление электрического поля \mathbf{E} не совпадает ни с одним из главных направлений поляризуе-

мости. Чтобы найти поляризованность, в этом случае представим поле в виде суперпозиции

$$\mathbf{E} = iE_X + jE_Y + kE_Z,$$

где i, j, k — орты осей X, Y, Z ; E_X, E_Y, E_Z — проекции поля на эти оси, и учтем линейность связи между P и E . Вследствие этой линейности проекции вектора поляризации на оси координат X, Y, Z составляют $P_X = \epsilon_0 \alpha_1 E_X$; $P_Y = \epsilon_0 \alpha_2 E_Y$, $P_Z = \epsilon_0 \alpha_3 E_Z$, т. е. вектор P можно представить в виде

$$\mathbf{P} = \epsilon_0 (i\alpha_1 E_X + j\alpha_2 E_Y + k\alpha_3 E_Z).$$

Ясно, что если величины $\alpha_1, \alpha_2, \alpha_3$ не одинаковы, то вектор P не параллелен вектору E , как это имеет место в случае изотропной среды.

Анизотропия поляризуемости приводит к анизотропии диэлектрической проницаемости, так как электрическая индукция D связана с поляризованностью P и напряженностью электрического поля E соотношением $D = \epsilon_0 E + P$, независимо от того, изотропно тело или анизотропно. Из этого соотношения следует, что в слабых полях между векторами D и E имеет место линейная связь, но эти векторы, вообще говоря, не параллельны. Если ввести *главные значения диэлектрической проницаемости*

$$\epsilon_1 = 1 + \alpha_1, \quad \epsilon_2 = 1 + \alpha_2, \quad \epsilon_3 = 1 + \alpha_3,$$

то проекции векторов D и E на главные направления будут связаны между собой соотношениями

$$D_X = \epsilon_1 \epsilon_0 E_X, \quad D_Y = \epsilon_2 \epsilon_0 E_Y, \quad D_Z = \epsilon_3 \epsilon_0 E_Z.$$

Величины $\alpha_1, \alpha_2, \alpha_3$ всегда положительны, поэтому главные значения диэлектрической проницаемости всегда больше единицы.

У кристаллов с малосимметричной решеткой, принадлежащих к *триклинной, моноклинной и ромбической* системам, все три главных значения диэлектрической проницаемости различны. Эти кристаллы носят название *двухосных*. При этом у кристаллов ромбической системы главные направления определяются кристаллографически, у кристаллов же триклинной системы главные направления не связаны с какими-либо кристаллографическими направлениями.

У кристаллов *тетрагональной, ромбоэдрической и гексагональной* систем два из трех главных значений диэлектрической проницаемости совпадают (такие кристаллы называют *одноосными*); при этом одно из главных направлений совпадает с осью симметрии (соответственно 4, 3, 6-го порядка), а два других можно выбрать произвольно в плоскости, перпендикулярной этой оси.

У кристаллов, принадлежащих *кубической* системе, все три главных значения диэлектрической проницаемости совпадают. Поэтому такие кристаллы ведут себя как изотропные среды; в частности, главные направления можно выбирать произвольно.

Подчеркнем еще раз, что поликристаллы в отличие от монокристаллов ведут себя в электростатическом поле как изотропные среды.

8.4. Кристаллы со спонтанной поляризованностью

В предыдущем параграфе мы уже упоминали, что кристалл может отличаться не только анизотропией поляризуемости, но обладать также поляризованностью в отсутствие поля. Дело в том, что монокристалл можно рассматривать как единую гигантскую молекулу. Поэтому не удивительно, что, так же как и молекула, кристалл в целом может быть полярным, т. е. может обладать дипольным моментом в отсутствие внешнего электрического поля. Момент этот будет макроскопическим, т. е. он пропорционален объему кристалла. Таким образом, в этом случае в отсутствие электрического поля кристалл обладает отличной от нуля поляризованностью. Ее называют *собственной* (или *спонтанной*) *поляризованностью*.

Для того чтобы кристалл обладал спонтанной поляризованностью, в нем должно существовать некоторое избранное направление, которое не должно изменяться при всех преобразованиях симметрии решетки кристалла. В частности, в числе элементов симметрии не должно быть центра симметрии, так как в этом случае каждое направление эквивалентно прямо противоположному. В числе элементов симметрии могут быть лишь одна ось и проходящие через нее плоскости симметрии. Это условие является, конечно, только необходимым, но отнюдь не достаточным условием наличия у кристалла спонтанной поляризации. Кристаллы со спонтанной поляризацией носят общее название *пирозлектриков*.

У одних кристаллов со спонтанной поляризацией спонтанный момент существует при любой температуре, т. е. во всей температурной области существования кристаллического состояния, а у других — только в определенном интервале температур. Кристаллы первого типа называют собственно пирозлектриками, а второго — *сегнетоэлектриками*. К числу пирозлектриков относится, например, турмалин, а к числу сегнетоэлектриков — сегнетова соль, титанат бария и др. У пирозлектриков спонтанный дипольный момент наблюдается обычно после их нагревания, так как при этом они освобождаются от осевших на них ионов (по этой причине и применяется приставка «пиро»). Название сегнетоэлектрика происходит от сегнетовой соли, которая была первым исследованным диэлектриком этого рода.

Спонтанный дипольный момент у пирозлектриков обусловлен либо смещением подрешеток ионов, либо упорядочением атомных групп, обладающих дипольным моментом. Смещение подрешеток имеет место, например, у кристаллов титаната бария BaTiO_3 . Структура титаната бария кубическая: ионы Ba^{2+} расположены в углах куба, ионы Ti^{4+} — в центре куба, а O^{2-} — в центрах граней (такого рода структура носит название «структуры перовскита» — это от названия минерала перовскита CaTiO_3). При температурах ниже 391 К, когда кристалл титаната бария имеет спонтанную поляризованность, решетка кристалла несколько деформирована, а именно растянута вдоль ребра куба; в этом направлении смещается ион титана, и в этом направлении образуется дипольный момент.

У сегнетоэлектриков ионы (или атомные группы, обладающие дипольным моментом), ответственные за возникновение спонтанной поляризации, смещаются сравнительно легко, тогда как у пьезоэлектриков они закреплены жестко. По этой причине у сегнетоэлектриков спонтанная поляризованность существует в определенном интервале температур, а у пьезоэлектриков — при любой температуре. Выше некоторой температуры (ее называют *точкой* или *температурой Кюри*) спонтанный дипольный момент сегнетоэлектрика исчезает. При этом происходит также изменение симметрии кристаллической решетки, так что ниже и выше температуры Кюри мы имеем дело с двумя различными фазами вещества. Их называют соответственно *сегнетоэлектрической* (или *ферроэлектрической*) и *параэлектрической* фазами.

Сегнетоэлектрический фазовый переход может быть как первого, так и второго рода. В случае фазового перехода первого рода выделяется либо поглощается теплота, а в случае фазового перехода второго рода в точке Кюри терпит скачок теплоемкость сегнетоэлектрика. Диэлектрическая проницаемость сегнетоэлектрика при приближении к точке Кюри резко возрастает. У пьезоэлектриков фазовых переходов таких типов не происходит.

В области высоких температур (выше температуры Кюри) спонтанный дипольный момент всегда отсутствует. Но он может отсутствовать и в области низких температур, так что может существовать лишь некоторый интервал температур, в котором устойчива сегнетоэлектрическая фаза. (В этом случае обе граничные температуры называют температурами Кюри — *верхней* и *нижней*.)

Так как ионы, ответственные за возникновение спонтанной поляризации сегнетоэлектриков, могут легко смещаться, то механические напряжения приводят к появлению у сегнетоэлектриков электрической поляризованности. Это явление называют *пьезоэлектрическим эффектом*. Существует и обратный эффект — изменение формы или размеров кристалла под действием внешнего электрического поля. Заметим, что этот эффект следует отличать от электрострикции у обычных диэлектриков, т. е. от изменения их формы и объема под влиянием внешнего поля. Электрострикция пропорциональна квадрату напряженности электрического поля, тогда как обратный пьезоэлектрический эффект пропорционален первой степени напряженности поля.

Все сегнетоэлектрики являются пьезоэлектриками, но пьезоэлектриками могут быть и не сегнетоэлектрики. К числу последних относятся, например, кварц.

Упомянем еще об одном классе веществ, обладающих спонтанной поляризованностью. Это так называемые *электреты*, которые получают при затвердевании смесей некоторых органических смол в сильном электрическом поле (а также другими способами). Электрическое поле ориентирует дипольные моменты некоторых из молекул этих веществ, а при затвердевании смолы моменты остаются ориентированными после выключения поля. Поляризованность, полученная таким образом, может сохраняться в течение нескольких лет, но, по-видимому, состояние поляризации электретов является метастабильным, стабильным же является неполяризованное состояние.

Момент силы, вращающей диполь

$$K = [d_0 E]$$

Потенциальная энергия диполя в электрическом поле

$$U = -d_0 E$$

Диэлектрическая проницаемость полярного газа

$$\epsilon = 1 + \frac{nd_0^2}{3\epsilon_0 kT}$$

Диэлектрическая восприимчивость жидкости или плотного газа (формула Онзагера)

$$\alpha = \frac{3}{4} \left(\alpha' n - 1 + \sqrt{1 + \frac{2}{3} \alpha' n + \alpha'^2 n^2} \right)$$

Вектор поляризации в кристалле

$$P = i\alpha_1 E_X + j\alpha_2 E_Y + k\alpha_3 E_Z$$



Глава 9. МЕТАЛЛЫ И ПОЛУПРОВОДНИКИ

9.1. Вырожденный электронный газ

Формулируя законы постоянного тока, мы ввели новые величины — удельную электрическую проводимость и электродвижущую силу, но не разъяснили пока их физической природы. Теперь мы приступим к этой задаче и начнем с проводимости металлов.

Как мы уже говорили, электрический заряд в металлах переносят электроны, которые могут свободно перемещаться от одного атома (правильнее — иона) к другому. Такое перемещение возможно благодаря перекрытию электронных оболочек отдельных атомов. Поэтому естественно исходить из представления об идеальном или почти идеальном газе электронов в металле. Этот газ, однако, должен существенно отличаться от обычного идеального газа молекул. Наиболее отчетливо это видно на примере теплоемкости металла. Если для простоты считать, что каждый атом металла вносит по одному электрону в электронный газ, то 1 г-атом металла должен содержать $6,02 \cdot 10^{23}$ ионов и столько же электронов. Молярная теплоемкость, обусловленная ионами, равна, как известно, $3R$, где R — газовая постоянная (з а к о н Д ю л о н г а — П т и). Молярная же теплоемкость электронов, если считать их обычным газом, равна $\frac{3}{2}R$. Поэтому суммарная теплоемкость металла должна составить $\frac{9}{2}R$. С другой стороны, если взять диэлектрик, в котором нет электронного газа, то его теплоемкость равна $3R$. Таким образом, теплоемкость металла должна быть больше теплоемкости диэлектрика на $\frac{3}{2}R$. Опыт, однако, находится

в резком противоречии с этим заключением: в области применимости закон Дюлонга—Пти, т. е. при достаточно высоких температурах ($T \gg T_D$, T_D — так называемая *температура Дебая*, см. ниже), молярные теплоемкости металлов и диэлектриков практически не отличаются друг от друга. Это значит, что электронный газ не вносит сколько-нибудь заметного вклада в теплоемкость металла, или, выражаясь точнее, вклад электронного газа значительно меньше $\frac{3}{2}R$. По этой причине электронный газ в металле называют *вырожденным газом*.

В чем же причина вырождения электронного газа? Она тесно связана с характером статистики электронов. Чтобы разъяснить, что это значит, напомним предварительно, что при выводе распределения Максвелла для молекул обычного газа вводится понятие *фазового пространства* молекулы, точки которого — *фазовые точки* — определяются шестью величинами — тремя координатами x , y , z молекулы и тремя проекциями ее импульса p_x , p_y , p_z . Объем фазового пространства пропорционален произведению обычного объема на объем импульсного пространства. В декартовых координатах $\Delta x \Delta y \Delta z$ и $\Delta p_x \Delta p_y \Delta p_z$ — элементы этих объемов. Объем шарового слоя радиуса r и толщины Δr в импульсном пространстве равен $4\pi r^2 \Delta r$.

При движении молекулы ее радиус-вектор $\mathbf{r}(x, y, z)$ и импульс $\mathbf{p}(p_x, p_y, p_z)$ изменяются со временем, т. е. фазовая точка описывает в фазовом пространстве некоторую кривую — *фазовую траекторию*. В зависимости от начальных условий (т. е. значений x , y , z , p_x , p_y , p_z в начальный момент времени $t = 0$) мы получим различные фазовые траектории. Каждая фазовая траектория соответствует определенному динамическому состоянию молекулы, а так как, очевидно, через каждую фазовую точку проходит только одна фазовая траектория, то каждой фазовой точке соответствует определенное динамическое состояние молекулы.

Рассмотрим теперь не одну, а много одинаковых молекул, почти не взаимодействующих друг с другом, т. е. их идеальный газ. Тогда наборы динамических состояний у различных молекул тождественны, но возникает вопрос: коррелируют ли как-либо между собой динамические состояния у различных молекул? Более просто вопрос этот формулируется так: произвольно ли число молекул, могущих находиться в заданном динамическом состоянии? При выводе распределения Максвелла вопрос этот, собственно говоря, не ставится и молчаливо предполагается, что число молекул, находящихся в заданном состоянии, может быть произвольным. Однако это предположение не является логически необходимым: ведь можно представить себе ситуации, когда оно не выполняется. Такая ситуация, в частности, и имеет место для газа электронов. При этом важно иметь в виду, что электрон обладает собственным моментом импульса (его называют *спином*) и что если мы хотим измерить проекцию спина на какую-либо ось, то всегда получим только два возможных значения. Поэтому в число переменных, определяющих динамическое состояние электрона, должна быть включена еще проекция его спина. Если сбобщить таким образом понятие динамического состояния электрона, то окажется, что

в каждом динамическом состоянии может находиться не больше одного электрона, т. е. состояние может быть либо не занято, либо в нем может находиться только один электрон и никак не более.

Это замечательное положение, носящее название принципа Паули, и определяет характер статистики электронов. Не одни электроны подчиняются принципу Паули. Этому же принципу подчиняются и протоны, и нейтроны, и вообще все частицы, спин которых равен полуцелому числу квантовых постоянных \hbar (см. ниже). Все такие частицы называют *фермионами*, и о них говорят, что они подчиняются статистике Ферми — Дирака.

Возможна и другая ситуация, когда число частиц, могущих находиться в данном состоянии, произвольно. Она осуществляется, если спин частиц равен целому числу \hbar . Такие частицы называют *бозонами*, и о них говорят, что они подчиняются статистике Бозе — Эйнштейна. Примерами бозонов являются фотоны, α -частицы, дейтроны, пионы.

Чтобы выяснить, к какому распределению электронов в состоянии статистического равновесия приводит принцип Паули, разобьем фазовое пространство Γ электрона на элементарные ячейки одинакового объема Γ_0 и будем сопоставлять каждой такой ячейке одно кинематическое состояние электрона, т. е. два состояния, отличающихся проекцией спина электрона. Для того чтобы такой метод сопоставления состояниям элементарных ячеек был разумным, нужно, очевидно, брать ячейки как можно меньшего объема. Но здесь имеется ограничение, вносимое квантовой механикой: оказывается, что объем ячейки Γ_0 нельзя взять меньше чем $(2\pi\hbar)^3$, где \hbar — квантовая постоянная, равная примерно 10^{-34} Дж·с, или 10^{-27} эрг·с (величину $h = 2\pi\hbar$ называют *постоянной Планка*). Взяв такое значение Γ_0 , мы получим, что в фазовом объеме Γ содержится ячеек

$$G = \Gamma / (2\pi\hbar)^3.$$

Об электронах, состояниям которых соответствуют определенные фазовые ячейки, говорят, что они находятся в этих ячейках. Если взять электроны с малым разбросом импульса и разделить их число на число соответствующих фазовых ячеек, то мы получим *функцию распределения* электронов. Задача состоит теперь в нахождении этой функции.

С понятием функции распределения нам придется неоднократно встречаться, поясним поэтому это понятие подробнее. Рассмотрим малый объем в фазовом пространстве $\Delta\Gamma = (2\pi\hbar)^{-3} \Delta x \Delta y \Delta z \Delta p_x \Delta p_y \Delta p_z$ вблизи точки с импульсом \mathbf{p} и радиусом-вектором \mathbf{r} . Число частиц $\Delta N = f(\mathbf{p}, \mathbf{r}) \Delta\Gamma$ в этом объеме пропорционально $\Delta\Gamma$. Коэффициент $f(\mathbf{p}, \mathbf{r})$ в этом соотношении и называется *функцией распределения*, точнее, *функцией распределения частиц в фазовом пространстве*.

Рассмотрим сперва состояние при $T = 0$. Энергия системы электронов должна быть при этом минимально возможной. Если бы электроны не подчинялись принципу Паули, то минимуму энергии соответствовал бы импульс, равный нулю, т. е. все электроны находились бы в одной ячейке, соответствующей нулевому импульсу. Но в действительности справедлив принцип Паули, и чтобы получить мини-

мум энергии, нужно плотно заполнить ячейки, т. е. разместить по одному электрону (с данной проекцией спина) в ячейку, а ячейки взять с минимально возможным для этого импульсом. Пусть при этом окажутся заполненными ячейки со всеми импульсами вплоть до некоторого импульса p_F . Тогда мы должны взять объем сферы радиуса p_F , т. е. $\frac{4}{3}\pi p_F^3$, умножить его на объем газа и разделить на $(2\pi\hbar)^3$. Таким образом, мы получим, что число нужных нам ячеек равно $\frac{4}{3}\pi p_F^3 V / (2\pi\hbar)^3$. Поместив в каждую ячейку по два электрона (с разными проекциями спина), найдем общее число электронов N , соответствующее такой плотной упаковке:

$$N = 2 \frac{4\pi}{3} \frac{p_F^3 V}{(2\pi\hbar)^3}.$$

Отсюда следует, что

$$p_F = (3\pi^2)^{1/3} \hbar n^{1/3},$$

где n — плотность электронного газа. Таким образом,

при $T=0$ распределение электронов, соответствующее минимуму энергии, таково: все состояния электронов с импульсами, заключенными между нулем и p_F , заняты, а состояния с $p > p_F$ свободны. Величину p_F принято называть *граничным импульсом Ферми*.

Импульсу p_F соответствует граничная энергия электрона (*граничная энергия Ферми*)

$$\epsilon_F = \frac{p_F^2}{2m_e} = \frac{(3\pi^2)^{2/3}}{2m_e} \hbar^2 n^{2/3},$$

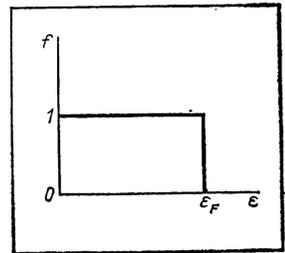
где m_e — масса электрона. Мы видим, что граничный импульс пропорционален плотности электронов в степени $1/3$, а граничная энергия — плотности в степени $2/3$.

Найденное распределение электронов, очевидно, изотропно в пространстве импульсов, т. е. зависит только от модуля, а не от направления импульса. Поэтому мы можем также рассматривать функцию распределения f как функцию только от энергии электрона $\epsilon = p^2/(2m_e)$. Она, очевидно, имеет вид ступеньки, резко обрывающейся при энергии, равной ϵ_F (рис. 9.1). Высота ступеньки равна единице (функция распределения относится к определенной проекции спина).

Найдем полную энергию электронного газа при $T = 0$. По определению функции распределения эту энергию можно представить в виде

$$W = 2 \int \epsilon f \frac{V 4\pi p^2 dp}{(2\pi\hbar)^3},$$

где V — объем, занимаемый электронами, и множитель 2 перед интегралом соответствует двум возможным проекциям спина электрона. Чтобы вычислить этот интеграл, энергию ϵ



9.1. Распределение Ферми для нуля температур

будем считать независимой переменной. Так как $\varepsilon = p^2/(2m_e)$, то $d\varepsilon = pdp/m_e$; поэтому

$$W = \frac{8\sqrt{2}\pi m_e^{3/2}}{(2\pi\hbar)^3} V \int_0^{\varepsilon_F} f\varepsilon^{3/2} d\varepsilon.$$

Учитывая ступенчатый характер функции распределения, можно заменить f на единицу и интегрировать в пределах от $\varepsilon = 0$ до $\varepsilon = \varepsilon_F$. В результате мы получим

$$W = \frac{2\sqrt{2}}{5\pi^2} \frac{m_e^{3/2}}{\hbar^3} V \varepsilon^{5/2}.$$

Вспомнив, что $\varepsilon_F = (3\pi^2)^{2/3} \hbar^2 n^{2/3} / (2m_e)$, можно переписать выражение для энергии электронного газа при $T = 0$ в форме

$$W = \frac{3^{5/3} \pi^{4/3}}{10} \frac{\hbar^2 n^{5/3} V}{m_e}.$$

Мы видим, что энергия электронного газа при $T = 0$ пропорциональна плотности электронов в степени $5/3$.

9.2. Распределение Ферми

Найдем теперь функцию распределения электронов в состоянии статистического равновесия при температуре, не равной нулю. Для этого нужно, так же как и при выводе распределения Максвелла, рассмотреть столкновения электронов либо друг с другом, либо, что проще, с посторонними молекулами, примешанными к электронному газу и находящимися в состоянии статистического равновесия. Функция распределения таких молекул f_M пропорциональна $\exp[-\varepsilon_M/(kT)]$, где ε_M — энергия молекулы, T — температура, k — постоянная Больцмана.

Пусть электрон, обладающий энергией ε , сталкивается с молекулой, обладающей энергией ε_M . В результате столкновения энергии частиц изменятся, но это изменение не будет однозначным, так что можно лишь говорить о вероятности того или иного изменения. Обозначим $\omega(\varepsilon, \varepsilon_M; \varepsilon', \varepsilon'_M)$ вероятность того, что за 1 с в результате столкновения энергии частиц стали равными соответственно ε' и ε'_M . Ясно, что, каково бы ни было столкновение, должен выполняться закон сохранения энергии, т. е. $\varepsilon + \varepsilon_M = \varepsilon' + \varepsilon'_M$.

Найдем, как изменится за 1 с из-за столкновений с молекулами газа число электронов с энергией ε . Изменение числа частиц равно, очевидно, разности числа электронов, приходящих за 1 с в состоянии с энергией ε , и числа электронов, уходящих из этого состояния (также за 1 с). Если бы электроны не подчинялись принципу Паули, то число, уходящих из состояния с энергией ε электронов вследствие столкновений с молекулами, имеющими энергию ε_M , определялось бы ве-

личной $\sum_{\epsilon', \epsilon'_M} \omega(\epsilon, \epsilon_M; \epsilon', \epsilon'_M) f(\epsilon) f_M(\epsilon_M)$, где $f(\epsilon)$ — функция распределения

электронов. Но в действительности имеет место принцип Паули, поэтому электрон сможет перейти в состояние с энергией ϵ' только в том случае, если это состояние свободно. Это значит, что написанное выражение должно быть умножено еще на $1 - f(\epsilon')$. Таким образом, число электронов, уходящих за 1 с из состояния с энергией ϵ , равно

$$\sum_{\epsilon', \epsilon'_M, \epsilon_M} \omega(\epsilon, \epsilon_M; \epsilon', \epsilon'_M) f(\epsilon) f_M(\epsilon_M) [1 - f(\epsilon')].$$

Аналогичным образом определяется число электронов, приходящих за 1 с в состояние с энергией ϵ :

$$\sum_{\epsilon', \epsilon'_M, \epsilon_M} \omega(\epsilon', \epsilon'_M; \epsilon, \epsilon_M) f(\epsilon') f_M(\epsilon'_M) [1 - f(\epsilon)].$$

Входящая сюда величина $\omega(\epsilon', \epsilon'_M; \epsilon, \epsilon_M)$ представляет собой вероятность перехода за 1 с частиц (электрона и молекулы) из состояний ϵ', ϵ'_M в состояния ϵ, ϵ_M . Так как законы механики не изменяются при обращении времени, то вероятности переходов прямого $\epsilon, \epsilon_M \rightarrow \epsilon', \epsilon'_M$ и обратного $\epsilon', \epsilon'_M \rightarrow \epsilon, \epsilon_M$ равны между собой: $\omega(\epsilon, \epsilon_M; \epsilon', \epsilon'_M) = \omega(\epsilon', \epsilon'_M; \epsilon, \epsilon_M)$. Поэтому окончательно для изменения числа электронов, обусловленного столкновениями с молекулами, мы получим

$$\left(\frac{df}{dt}\right)^{\text{coll}} = \sum_{\epsilon', \epsilon'_M, \epsilon_M} \omega(\epsilon, \epsilon_M; \epsilon', \epsilon'_M) \{f(\epsilon') [1 - f(\epsilon)] f_M(\epsilon'_M) - f(\epsilon) [1 - f(\epsilon')] f_M(\epsilon_M)\}.$$

В состоянии статистического равновесия эта величина должна быть равна нулю, т. е. $f(\epsilon') [1 - f(\epsilon)] f_M(\epsilon'_M) = f(\epsilon) [1 - f(\epsilon')] f_M(\epsilon_M)$. Подставляя сюда максвелловское распределение для молекул, получим

$$\frac{1 - f(\epsilon)}{f(\epsilon)} e^{\epsilon_M/(kT)} = \frac{1 - f(\epsilon')}{f(\epsilon')} e^{\epsilon'_M/(kT)}.$$

Но $\epsilon_M - \epsilon'_M = \epsilon' - \epsilon$, поэтому

$$\frac{1 - f(\epsilon)}{f(\epsilon)} e^{-\epsilon/(kT)} = \frac{1 - f(\epsilon')}{f(\epsilon')} e^{-\epsilon'/(kT)}.$$

Первая часть этого равенства зависит только от ϵ' , а левая — от ϵ . С другой стороны, ϵ и ϵ' никак не связаны между собой, поэтому следует считать, что и правая и левая части равенства представляют собой одну и ту же константу, не зависящую от энергии электрона. Эту константу можно записать в виде $e^{-\mu/kT}$, где μ — некоторая другая константа. Ее называют *химическим потенциалом* электронов. Итак,

$$[1 - f(\epsilon)] e^{-\epsilon/(kT)} / f(\epsilon) = e^{-\mu/(kT)},$$

откуда

$$f(\epsilon) = \frac{1}{e^{(\epsilon-\mu)/(kT)} + 1}.$$

Эта формула и определяет распределение по энергиям газа электронов, находящегося в состоянии статистического равновесия. Оно носит название **распределения Ферми**.

Напомним, что функция $f(\epsilon)$ дает распределение электронов с заданной проекцией спина. Если мы хотим определить число электронов dN в объеме V , энергия которых заключена между ϵ и $\epsilon + d\epsilon$, то функцию распределения Ферми $f(\epsilon)$ надо умножить на

$$2 \frac{4\pi p^2 dp V}{(2\pi\hbar)^3} = \frac{V\sqrt{2} m_e^{3/2}}{\pi^2 \hbar^3} \epsilon^{1/2} d\epsilon V.$$

Мы получим

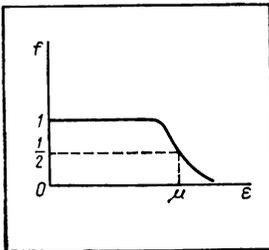
$$dN \frac{V\sqrt{2} m_e^{3/2}}{\pi^2 \hbar^3} V \frac{\epsilon^{1/2} d\epsilon}{e^{(\epsilon-\mu)/(kT)} + 1}.$$

Интегрируя это выражение по энергии, найдем полное число электронов:

$$N = \frac{V\sqrt{2} m_e^{3/2}}{\pi^2 \hbar^3} V \int_0^\infty \frac{\epsilon^{1/2} d\epsilon}{e^{(\epsilon-\mu)/(kT)} + 1}.$$

Эта формула связывает константу μ с плотностью электронов $n = N/V$.

Выясним теперь, каковы особенности распределения Ферми. Будем считать, что температура достаточно мала, так что $kT \ll \mu$. Тогда если энергия электрона ϵ меньше μ , причем настолько, что $(\mu - \epsilon) > kT$, то экспонента $e^{(\epsilon-\mu)/(kT)}$, входящая в знаменатель распределения Ферми, мала и распределение электронов при $T \neq 0$ практически не отличается от распределения при $T = 0$. Если же энергия $\epsilon < \mu$, причем настолько, что $(\epsilon - \mu) \gg kT$, то экспонента очень велика и функция распределения близка к нулю. Поэтому распределение имеет вид, изображенный на рис. 9.2. Это распределение представляет собой ступеньку, но не резко обрывающуюся, как при $T = 0$, а «размытую» с шириной «размытости» порядка kT . Мы видим, что если $kT \ll \mu$, то распределение при $T \neq 0$ отличается от распределения при $T = 0$ только в «зоне размытости» $\sim kT$. При этом величина μ мало отличается от ϵ_F и условие $kT \ll \mu$ может быть переписано в виде $T \ll T_F$, где



9.2. Распределение Ферми для ненулевой температуры

$$T_F = \frac{\epsilon_F}{k} = \frac{(3\pi^2)^{2/3}}{2m_e k} \hbar^2 n^{2/3}.$$

Величину T_F называют *температурой вырождения* электронного газа. Она тем выше, чем больше плотность электронов. В металлах $n \sim 10^{24} \text{ см}^{-3}$, $p_F/\hbar \sim 1,2 \cdot 10^8 \text{ см}^{-1}$ и, следовательно, $\epsilon_F \sim 5,6 \text{ эВ}$ и $T_F \sim 6 \cdot 10^4 \text{ К}$. Поэтому при всех температурах, вплоть до температуры плавления металла, выполняется условие $T \ll T_F$.

Если $T \sim T_F$, то ширина размытости сравнима с граничной энергией ϵ_F . Если $T \ll T_F$, то μ мало отличается от ϵ_F , а именно μ немного меньше ϵ_F — на величину порядка $(kT)^2/\epsilon_F$. Если $T \gg T_F$, то распределение Ферми переходит в распределение Максвелла. (Такие ситуации могут реализовываться в полупроводниках и полуметаллах, где плотность электронов много меньше, чем в металлах.)

Найдем теперь энергию электронного газа при $T \neq 0$. Она определяется согласно общей формуле

$$W = 2 \int \epsilon f(\epsilon) \frac{4\pi p^2 dp V}{(2\pi\hbar)^3} = \frac{\sqrt{2} m_e^{3/2}}{\pi^2 \hbar^3} V \int_0^\infty \frac{\epsilon^{3/2} d\epsilon}{e^{(\epsilon-\mu)/(kT)} + 1}.$$

Мы не будем вычислять этого интеграла, а выясним лишь, как зависит энергия от температуры при $T \ll T_F$. Для этого будем исходить из упрощенной функции распределения, изображенной на рис. 9.3. Ее можно, очевидно, в свою очередь, представить как полусумму двух ступенек, характерных для $T = 0$: одной, соответствующей граничной энергии $\epsilon_F - kT$, и другой, соответствующей граничной энергии $\epsilon_F + kT$. Поэтому энергию при $T \neq 0$ можно приближенно оценить как

$$W \sim \frac{\sqrt{2} m_e^{3/2}}{5\pi^2 \hbar^3} \left[(\epsilon_F - kT)^{5/2} + (\epsilon_F + kT)^{5/2} \right].$$

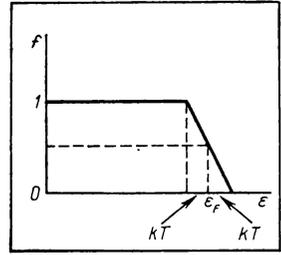
Так как $kT \ll \epsilon_F$, то мы получим

$$W = \frac{2\sqrt{2}}{5\pi^2} \frac{m_e^{3/2}}{\hbar^3} \epsilon_F^{5/2} V \left[1 + \xi \left(\frac{kT}{\epsilon_F} \right)^2 \right],$$

где ξ — число порядка единицы. Итак,

$$W \approx W_0 \left[1 + \xi \left(\frac{kT}{\epsilon_F} \right)^2 \right],$$

где W_0 — энергия при $T = 0$. Мы видим, что при $T \ll T_F$ энергия вырожденного электронного газа практически не зависит от температуры, так как относительное увеличение энергии электронного газа при нагревании по сравнению с энергией при $T = 0$ составляет по порядку величины $(T/T_F)^2$. Отсюда следует, что средняя квадра-



9.3.

К выводу энергии, электронного газа при ненулевой температуре

тичная или средняя тепловая скорость электрона $\langle v \rangle$ практически не зависит от температуры. Эта скорость определяется как $\langle v \rangle = \sqrt{\frac{2W}{mN}}$ и равна приближенно $\langle v \rangle \sim \sqrt{\frac{\epsilon_F}{m}}$.

Найдем, наконец, теплоемкость C электронного газа. Она определяется известной формулой $C = dW/dT$, откуда

$$C = 2\xi W_0 \frac{T}{T_F^2}.$$

Отнесенная к одному электрону теплоемкость равна по порядку величины

$$c \sim k \frac{T}{T_F}.$$

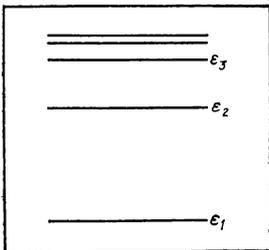
Если бы электронный газ вел себя как обычный максвелловский газ, то мы имели бы $c = \frac{3}{2} k$; в действительности же, как мы видим, теплоемкость уменьшается по сравнению с этой величиной в T_F/T раз. Происходит это благодаря принципу Паули, которому подчиняются электроны.

Таким образом, принцип Паули объясняет специфические тепловые свойства, электронного газа при $T \ll T_F$, или, как говорят, вырожденного электронного газа.

9.3. Металлы, диэлектрики и полупроводники

Принцип Паули объясняет не только особенности тепловых свойств электронного газа, но и сам факт разделения кристаллов на металлы, диэлектрики и полупроводники. Чтобы разъяснить, в чем заключается принципиальное различие между этими типами кристаллов и какую роль играет при этом принцип Паули, попытаемся сперва понять, что происходит при объединении отдельных атомов в единый кристалл.

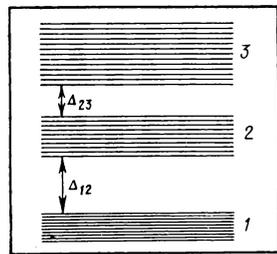
Если взять отдельный атом, то движение электронов в нем происходит таким образом, что их энергия может принимать лишь вполне определенные значения, иными словами, атом характеризуется определенным дискретным энергетическим спектром. Схематически он изображен на рис. 9.4, на котором возможным значениям энергии атома $\epsilon_1, \epsilon_2, \dots$ соответствуют различные горизонтальные линии.



9.4. Энергетический спектр отдельного атома

Пусть теперь отдельные атомы объединяются в кристалл. Тогда ядра атомов расположатся в узлах некоторой решетки, электронные же оболочки атомов в большей или меньшей степени перекроются, так что электроны, строго говоря, станут коллективизированными, т. е. принадлежатими всей решетке в целом. При этом изменится также ха-

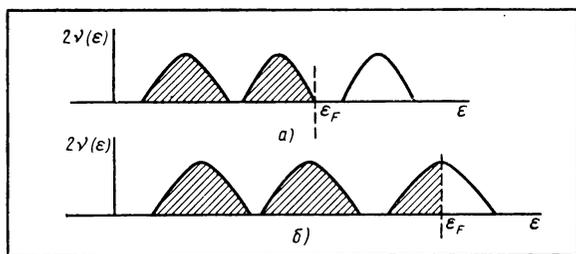
рактически энергетического спектра. Если говорить сперва об одном электроне и исходить из схемы, изображенной на рис. 9.4, то из каждого энергетического уровня ϵ_i ($i = 1, 2, \dots$) возникнет целая полоса уровней, содержащая столько их, сколько всего есть электронов, т. е. сколько всего атомов объединяется в кристалл. Иными словами, мы получим схему, изображенную на рис. 9.5.



9.5. Энергетический спектр кристалла

Здесь цифры 1, 2, ... служат для обозначения отдельных энергетических зон, или полос, возникающих из уровней энергии $\epsilon_1, \epsilon_2, \dots$ отдельного атома. Горизонтальные линии в пределах каждой зоны обозначают возможные энергетические уровни отдельного электрона в кристалле. Число уровней в каждой зоне равно числу электронов в кристалле. Схема эта является упрощенной, в действительности ситуация значительно сложнее, так как возможно перекрытие отдельных энергетических полос; мы не будем, однако, рассматривать здесь этих усложнений. Главным является наличие энергетических зон, разделенных конечными интервалами $\Delta_{12}, \Delta_{23}, \dots$, и тот факт, что в пределах каждой зоны спектр практически непрерывен, так как число уровней в каждой зоне очень велико. Зоны же разделены запрещенными интервалами (т. е. электрон не может обладать энергией, попадающей в один из этих интервалов).

До сих пор мы говорили об энергетическом спектре одного электрона. Но в кристалле очень много электронов — по одному или по несколько на каждый атом. Если отвлечься от взаимодействия между электронами, то возникает задача — разместить электроны по допустимым энергетическим уровням отдельного электрона. Тут вступает в действие принцип Паули, согласно которому на каждом уровне может находиться не более одного электрона с определенной ориентацией спина. Если известно число уровней энергии ΔN , приходящихся на малый интервал энергий $\Delta \epsilon$, $\Delta N = \nu(\epsilon)\Delta \epsilon$, то легко понять, сколько электронов может иметь энергию, в этом интервале: число таких электронов не может превышать $2\nu(\epsilon)\Delta \epsilon$. График функции $2\nu(\epsilon)$, которую называют плотностью электронных состояний, схематически изображен на рис. 9.6. Если мы хотим, чтобы энергия кристалла была минимально возможной, то мы должны сперва заполнить самую низкую зону, затем следующую более высокую, и т. д. до тех пор, пока не исчерпаем все электроны, т. е. пока не доберемся до фермиевской граничной энергии ϵ_F , определяющейся, как мы видели, плотностью электронов в рассматриваемом кристалле. Ясно, что при таком постепенном размещении электронов могут представиться два случая: либо все электроны полностью заполнят несколько зон, остальные же зоны будут свободными (рис. 9.6, а), либо электронов окажется недостаточно для полного заполнения, так что в последней из заполняемых зон останутся незанятыми уровни (рис. 9.6, б). В первом случае фермиевский уровень $\epsilon = \epsilon_F$ совпадает с краем энергетической зоны, во



9.6. Плотность электронных состояний кристаллического диэлектрика (а) и металла (б)

втором — лежит где-то внутри одной из зон (на рисунках электронные состояния с $\epsilon < \epsilon_F$ заштрихованы).

Легко видеть, что кристалл является изолятором, т. е. диэлектриком, если зоны заполнены полностью, или же проводником, т. е. металлом, если одна из зон заполнена не полностью (но в ней есть электроны!). Действительно, если к кристаллу приложить слабое электрическое поле, то под влиянием поля электроны будут ускоряться и переходить из состояния с меньшей в состояние с большей энергией. Но в первом случае такой переход невозможен: ведь для того, чтобы электрону перейти в какое-либо состояние, это состояние согласно принципу Паули должно быть свободным, свободных же состояний, нет — все они заняты. Напротив, во втором случае переход возможен — в зоне есть свободные места. Поэтому в этом случае протекание тока возможно при сколь угодно слабом электрическом поле, т. е. кристалл является проводником. На этом основании не полностью заполненную зону называют *зоной проводимости*, а находящиеся в ней электроны — *электронами проводимости*. Подчеркнем, что в металле ток может протекать при сколь угодно малом электрическом поле, поскольку в нем имеются незанятые уровни в зоне проводимости. Диэлектрики не обладают этим свойством. Однако и в них могут протекать токи, но не при слабых, а очень сильных полях, когда становится возможным переход электронов в зону проводимости. Явление это (оно носит название *электрического пробоя диэлектрика*) носит сложный характер, и мы не будем здесь его рассматривать.

Таким образом, различие между диэлектриками и металлами заключается не в степени связи электрона — в первом случае большей, а во втором — меньшей, как может показаться на первый взгляд. Дело также не в перекрытии электронных оболочек — и у металлов, и у диэлектриков электронные оболочки перекрываются, причем особенно значительно это перекрытие для внешних электронов. Сущность различия между металлами и диэлектриками связана с характером энергетического спектра тела в сочетании с принципом Паули. Ясно, что если бы электроны не подчинялись принципу Паули, то различия между металлами и диэлектриками не было бы и все тела были бы проводниками. Так было бы, например, если бы атомы были построены из ядер и отрицательных пионов, так как пионы принадлежат к бозонам — частицам, не подчиняющимся принципу Паули.

Как мы видели, в диэлектрике существует энергетическая щель между полностью заполненной и незаполненной зонами. Но пред-

ставим себе, что ширина этой щели мала. Тогда электроны легко перебросить из занятой в свободную зону: стоит только нагреть кристалл, и мы получим проводник с двумя зонами проводимости. Такие кристаллы с малой шириной щели Δ носят название *полупроводников* (точнее, *собственных полупроводников*). Число электронов проводимости в них пропорционально $\exp[-\Delta/(kT)]$. Эта величина быстро растет с повышением температуры, при низких же температурах и больших Δ она близка к нулю. Таким образом, в полупроводниках число электронов проводимости существенно зависит от температуры, тогда как в металлах число электронов проводимости от температуры не зависит. (Приведем значения Δ для некоторых полупроводников: 0,7 эВ для Ge, 0,2 эВ для InSb.)

Важно, что в собственном полупроводнике при отличной от нуля температуре под влиянием электрического поля движутся не только электроны, попавшие в незаполненную зону, но и электроны в зоне, которая при $T = 0$ была полностью заполнена (ее называют валентной зоной). Так как число освободившихся состояний в этой зоне мало, то удобнее говорить не о движении электронов валентной зоны, а о движении в обратном направлении самых свободных состояний. Эти свободные состояния называются *дырками*; как и электроны, они подчиняются принципу Паули. Поскольку мы считаем, что дырки движутся в электрическом поле в направлении, противоположном направлению движения электронов, им следует приписать заряд, равный по модулю и противоположный по знаку заряду электрона. Ясно, что в собственном полупроводнике концентрации электронов проводимости и дырок совпадают (*условие электронейтральности*), а уровень Ферми проходит внутри запрещенной полосы между валентной зоной и зоной проводимости.

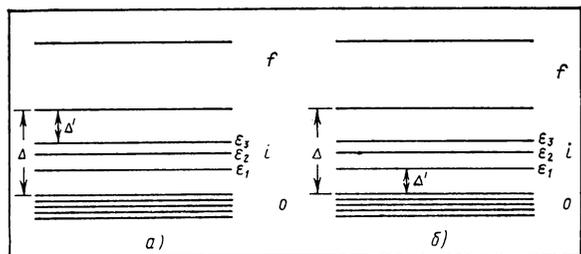
Заметим, что, строго говоря, переход электронов из заполненной в свободную зону принципиально возможен и в диэлектрике, но здесь интервал энергии между зонами столь велик (обычно он составляет несколько электрон-вольт, т. е. десятков тысяч кельвин), что вплоть до температуры плавления диэлектрика величина $\exp[-\Delta/(kT)]$ ничтожно мала.

Однако в диэлектрике с большим Δ может возникнуть проводимость, если диэлектрик содержит достаточно примесных атомов (например, тысячные доли процента). В этом случае электронные уровни примесных атомов (так называемые *примесные уровни*) могут попасть между занятой и свободной зонами диэлектрика. На рис. 9.7, а пред-

9.7.

Схема уровней энергии электронов в примесных полупроводниках донорного (а) и акцепторного (б) типов:

о — занятые, f — свободные зоны, l — примесные уровни



ставлена ситуация, когда примесные уровни ϵ_1 , ϵ_2 , ϵ_3 заняты и электроны с них могут переходить в свободную зону. Число перешедших электронов пропорционально $c \cdot \exp[-\Delta'/(kT)]$, где Δ' — интервал энергии между верхним занятым примесным уровнем и нижним уровнем свободной зоны и c — концентрация примесных атомов. Если $\Delta' \ll \Delta$, то эта величина значительно превышает $\exp[-\Delta/(kT)]$ даже при малых c . Попадая в незанятую энергетическую зону диэлектрика, электроны примесных атомов становятся электронами проводимости. Кристалл перестает быть диэлектриком и становится *примесным полупроводником донорного типа* (n -типа), или *электронным полупроводником*. Другая возможная ситуация представлена на рис. 9.7, б. Здесь примесные уровни ϵ_1 , ϵ_2 , ϵ_3 не заняты и электроны могут переходить на них из заполненной зоны, так что в этой зоне создается дырка. Такой кристалл называют *примесным полупроводником акцепторного типа* (p -типа) или *дырочным полупроводником*.

Примесные полупроводники получают добавляя примесные атомы не только в диэлектрики, но и в собственные полупроводники. В этом случае в полупроводнике присутствуют как электроны проводимости, так и дырки, но концентрации частиц одного сорта превышают концентрацию частиц другого сорта (в зависимости от того, какого типа примесь добавлена).

Приведем значения Δ' для некоторых примесных полупроводников: $\sim 0,01$ эВ для германия с примесью фосфора (донорный полупроводник) и для германия с примесью бора (акцепторный полупроводник), $\sim 0,045$ эВ для кремния с теми же примесями.

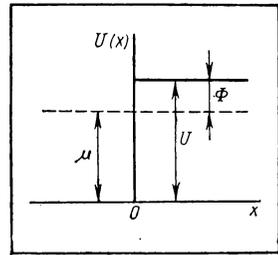
Электронов проводимости и дырок в полупроводниках значительно меньше, чем свободных электронов в металлах, и газы электронов и дырок в полупроводниках, как правило, вырождены. Поэтому, в частности, средняя квадратическая скорость электрона и дырки в полупроводнике пропорциональна \sqrt{T} , как это имеет место для обычных молекул (напомним, что в металлах средняя квадратическая скорость электрона проводимости от температуры практически не зависит). Бывают, однако, случаи (в несобственных проводниках), когда электроны или дырки вырождены; о таких полупроводниках говорят как о *вырожденных*.

9.4. Термоэлектронная эмиссия

Электроны проводимости в металле образуют, как мы разъясняли в предыдущих параграфах, вырожденный газ частиц, подчиняющихся принципу Паули (такой газ называется *ферми-газом*). Этот газ удерживается в металле атомными силами связи, т. е. силами притяжения электронов к атомным ядрам. Но подобно тому как молекулы могут испаряться из жидкости, так и электроны могут покидать металл. Этот эффект называют *термоэлектронной эмиссией*. Испарение жидкости возрастает с повышением температуры. Подобным образом растет с температурой и термоэлектронная эмиссия.

Легко найти плотность электронов, «испаряющихся» с поверхности металла. Для этого электроны в металле и электроны, покинувшие

металл, следует рассматривать как единый газ частиц, находящийся в состоянии статистического равновесия. Ясно, что потенциальная энергия электрона внутри металла меньше, чем снаружи (иначе электрон вообще бы не находился внутри металла). Таким образом, возникает задача о нахождении распределения электронов не только по импульсам, но и по пространственным координатам. Это можно сделать, вспомнив, что распределение обычных молекул по скорости и координатам (*распределение Максвелла—Больцмана*) получается из распределения Максвелла по скоростям, заменой кинетической энергии молекулы на ее полную энергию, т. е. на сумму ее кинетической и потенциальной энергий. Аналогичная ситуация имеет место и для фермиевского газа. Нам нужно лишь знать зависимость потенциальной энергии электрона $U(x)$ от расстояния x до поверхности металла. Будем исходить из простейшей зависимости, изображенной на рис. 9.8: здесь области внутри металла ($x < 0$) соответствует прямоугольная потенциальная яма глубины $U > 0$. На рисунке отмечен также уровень химического потенциала μ . Так как газ электронов внутри металла вырожден, то значение μ практически совпадает с максимумом энергии электрона, причем все уровни (считая от дна ямы) вплоть до μ заняты. Поэтому легче всего удалить электрон из металла в том случае, когда его энергия близка к μ — для этого нужно затратить минимальную работу, равную $\Phi = U - \mu$. Эту величину называют *работой выхода*.



9.8.
К выводу формулы для тока термоэлектронной эмиссии

Обратимся теперь к электронам вне металла. Так как они находятся в равновесии с электронами внутри металла, то химический потенциал электронного газа вне металла не отличается от химического потенциала электронного газа внутри металла. (Напомним, что в условиях равновесия во всех точках физической системы одинаковы температура, давление и химические потенциалы частиц каждого сорта.) Поэтому функция распределения электронов вне металла (в вакууме) имеет вид

$$f_{\text{вак}} = \left\{ e^{(U + \epsilon_h - \mu)/(kT)} + 1 \right\}^{-1},$$

где $\epsilon_h = p^2/(2m_e)$ — кинетическая энергия электрона с импульсом p вне металла (m_e — масса электрона). Вспомнив, что $U - \mu = \Phi$, и предположив, что $\Phi \gg kT$, в знаменателе распределения Ферми можно отбросить единицу. Таким образом, вне металла распределение электронов является максвелловским:

$$f_{\text{вак}} \approx e^{-\Phi/(kT)} e^{-\epsilon_h/(kT)} \quad (kT \gg \Phi).$$

Для того чтобы найти плотность электронов $n_{\text{вак}}$ вне металла, нужно умножить $f_{\text{вак}}$ на $2 \cdot 4\pi p^2 dp / (2\pi\hbar)^3$ и проинтегрировать по p :

$$n_{\text{вак}} = e^{-\Phi/(kT)} \frac{1}{\pi^2 \hbar^3} \int_0^{\infty} e^{-p^2/(2m_e kT)} p^2 dp.$$

Входящий сюда интеграл равен $\sqrt{\pi/2}(m_e kT)^{3/2}$, так что

$$n_{\text{вак}} = \left(\frac{\pi}{2}\right)^{1/2} \frac{(m_e kT)^{3/2}}{\pi^2 \hbar^3} e^{-\Phi/(kT)}.$$

Мы видим, что плотность электронов, «испаряющихся» из металла, зависит от температуры, причем зависимость эта в основном экспоненциальная. С повышением температуры плотность электронов очень быстро возрастает.

Зная функцию распределения электронов вне металла, можно определить электронный ток эмиссии. Для этого нужно определить поток электронов, уходящих от границы металла в сторону вакуума, т. е. таких электронов, для которых проекция скорости v_x положительна. Отсюда следует, что плотность тока эмиссии j_s (т. е. ток с единичной поверхности металла) определяется формулой

$$j_s = 2e \int_{(v_x > 0)} v_x f_{\text{вак}} \frac{d^3 p}{(2\pi \hbar)^3},$$

где e — заряд электрона и $d^3 p = dp_x dp_y dp_z$. Интеграл этот распадается на три интеграла по p_x , p_y , p_z , а так как входящие сюда интегралы по p_y и p_z одинаковы и равны $\sqrt{2\pi m kT}$, а интеграл по p_x равен kT , то окончательно получим

$$j_s = \frac{em}{2\pi^2 \hbar^3} (kT)^2 e^{-\Phi/(kT)}.$$

Эту формулу называют формулой Ричардсона — Дешмана. Обычно ее записывают в виде

$$j_s = CT^2 e^{-\Phi/(kT)}, \quad C \simeq 1,204 \cdot 10^6 \text{ А} \cdot \text{м}^{-2} \cdot \text{К}^{-2},$$

где T — температура. Мы видим, что ток эмиссии, так же как и плотность испущенных электронов, зависит от температуры в основном экспоненциально, быстро возрастающая с ее увеличением. Поэтому, чтобы получить большие токи эмиссии, нужно металл сильно нагреть. Большой ток эмиссии (при одной и той же температуре) дает тот металл, у которого работа выхода меньше. Значения работы выхода для разных металлов таковы: 4,52 эВ для вольфрама, 4,07 эВ для тантала, 6,27 эВ для платины, 2,24 эВ для кальция, 1,81 эВ для цезия. Для того чтобы лучше почувствовать, как влияет температура на плотность тока эмиссии, приведем данные для вольфрама. При температуре 1000 К плотность тока эмиссии в нем составляет $10^{-19} \text{ А} \cdot \text{м}^{-2}$; при температуре 2000 К она становится равной $10^{-7} \text{ А} \cdot \text{м}^{-2}$, а при температуре 3000 К достигает значения $1,4 \cdot 10^{-3} \text{ А} \cdot \text{м}^{-2}$.

Обратим внимание на то, что температурная зависимость тока эмиссии оказывается следствием только того факта, что вне кристалла

электронный газ невырожден; поэтому формула Ричардсона—Дешмана справедлива и для полупроводников. Мало того, может быть доказано, что хотя в полупроводнике нет электронов с $\epsilon = \mu$, тем не менее работа выхода, как и в металле, равна $\Phi = U - \mu$.

Электроны могут покинуть и холодный металл, но для этого нужно либо приложить к металлу сильное электрическое поле (направленное в сторону металла), либо осветить металл светом достаточно малой длины волны (см. ниже); наконец, электроны можно выбивать из металла, бомбардируя его пучком ионов. Что касается вырывания электронов полем (так называемое «холодное» вырывание), то зависимость плотности электронов $n_{\text{вак}}$ вне металла от напряженности поля E определяется соотношением $n_{\text{вак}} \sim \exp[-\gamma\Phi^{3/2}/(|e|E)]$, где Φ — работа выхода и $\gamma = 4\sqrt{2m}/(3\hbar)$. Плотность электронов $n_{\text{вак}}$ не зависит от температуры, но в сильнейшей степени зависит от напряженности поля, резко возрастая с ее увеличением. Можно сказать, что напряженность поля играет здесь такую же роль, как температура в термоэлектронной эмиссии.

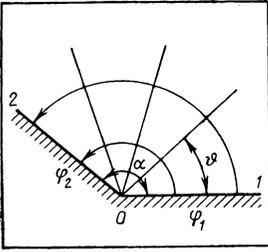
Вырывание из металла электронов светом (так называемый *фотоэффект*) возможно только в том случае, когда частота света ω превосходит некоторый предел. Именно: необходимо, чтобы выполнялось условие $\hbar\omega > \Phi$. При этом кинетическая энергия освобожденного электрона

$$mv^2/2 = \hbar\omega - \Phi.$$

Это соотношение (его называют с о о т н о ш е н и е м Э й н ш т е й н а) показывает, что свет в фотоэффекте ведет себя как газ частиц, причем энергия такой частицы — фотона — равна $\hbar\omega$. Если энергия фотона меньше работы выхода Φ электрона, то вылет невозможен, если же $\hbar\omega > \Phi$, то вырывание возможно и написанное соотношение представляет собой не что иное, как закон сохранения энергии.

9.5. Контактная разность потенциалов

Если привести в соприкосновение два различных проводника (далее для определенности будем говорить о металлах), то они будут обмениваться электронами и в результате придут в состояние статистического равновесия. При этом электроны обоих металлов можно рассматривать как единый ферми-газ с одинаковыми температурой и химическим потенциалом. Различаться будут лишь энергии электронов и, в частности, глубины потенциальных ям U_1 и U_2 . Это значит, что на границе обоих металлов существует скачок $\Delta U = U_1 - U_2$ потенциальной энергии электрона. Потенциальную энергию можно представить как произведение заряда частицы e на потенциал ϕ , поэтому на границе двух металлов имеется скачок потенциала $\phi_1 - \phi_2 = \frac{1}{e}(U_1 - U_2)$. Выше мы связали потенциальную энергию U с работой выхода $\Phi = U - \mu$, а так как химические потенциалы обоих металлов одинаковы, то скачок потенциала определяется разностью работ выхода Φ_1 и Φ_2 для обоих металлов:



9.9.
К выводу формулы для напряженности поля, обусловленного контактной разностью потенциалов

$$\Phi_1 - \Phi_2 = (\Phi_1 - \Phi_2)/e_0$$

Разность потенциалов между соприкасающимися металлами называют контактной разностью потенциалов.

Мы видим, что контактная разность потенциалов пропорциональна разности работ выхода.

Физический смысл этого соотношения очень прост. Пусть например, $\Phi_1 > \Phi_2$. Тогда электронам легче перейти из металла 2 в металл 1, чем обратно. В металле 1 образуется избыточный заряд, между металлами возникнет поле, направленное от металла 2 к металлу 1, оно препятствует переходу электронов из металла 2 в металл 1. Равновесие будет достигнуто тогда, когда контактная разность потенциалов станет равной $(\Phi_1 - \Phi_2)/e_0$.

Вычислим напряженность поля, обусловленного контактной разностью потенциалов. Рассмотрим простейший случай двух соприкасающихся плоских металлических поверхностей с потенциалами Φ_1 и Φ_2 (рис. 9.9, на котором изображен разрез в плоскости, перпендикулярной линии соприкосновения, она проходит через точку O, причем α — угол между плоскостями). Ясно, что в этом случае эквипотенциальные поверхности — это плоскости, проходящие через линию соприкосновения металлов, а силовые линии — дуги окружностей с центром в точке O. Потенциал на плоскости, образующей угол θ с плоскостью 1, составляет $\varphi = \Phi_1 + \frac{\theta}{\alpha}(\Phi_2 - \Phi_1)$, а напряженность

поля на расстоянии r от O есть $E = (\Phi_2 - \Phi_1)/\alpha r$. Она убывает обратно пропорционально расстоянию до точки O. Умножив E на ϵ_0 , найдем плотность поверхностных зарядов на поверхностях проводников 1 и 2.

Заметим, что так как металл представляет собой не изотропное, а анизотропное тело, свойства которого не одинаковы в разных направлениях, то работа выхода электронов может зависеть от ориентации поверхности металла по отношению к кристаллографическим осям его кристаллической решетки. Вследствие этого в одном и том же кристалле возможно существование контактных разностей потенциалов между различными его гранями. Поэтому в металлических монокристаллах у их ребер может существовать электрическое поле.

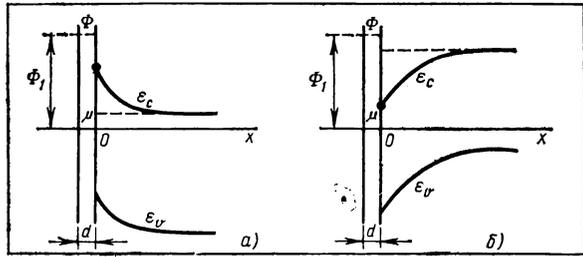
Если взять не два, а целый ряд попарно соприкасающихся металлов 1, 2, ..., то между соседними металлами возникнут контактные разности потенциалов $\Phi_{12} = (\Phi_1 - \Phi_2)/e_0$, $\Phi_{23} = (\Phi_2 - \Phi_3)/e_0$, ..., где Φ_1, Φ_2 — работы выхода для металлов 1, 2, Отсюда легко найти разность потенциалов между крайними металлами — 1 и n :

$$\Phi_{1n} = \Phi_{12} + \Phi_{23} + \dots + \Phi_{(n-1)n} = (\Phi_1 - \Phi_n)/e_0$$

Мы видим, что она определяется только крайними металлами. Поэтому если замкнуть ряд металлов в одно кольцо, то в этом случае Φ_n равно Φ_1 и сумма всех скачков потенциала обратится в нуль. Это значит, что

9.10.

Искривление энергетических зон в полупроводнике вблизи контакта полупроводник — металл



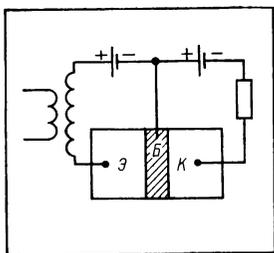
с помощью одних лишь металлов, находящихся в состоянии статистического равновесия, нельзя создать источника э. д. с. Это и неудивительно, так как в противном случае мы получили бы противоречие с законом сохранения энергии. Однако источник э. д. с. можно создать, если спай металлов держать при разных температурах, так как в этом случае система проводников становится неравновесной.

Посмотрим, что произойдет при плотном контакте двух проводников 1 и 2, когда зазор между ними $d \sim 10^{-9}$ м. Очевидно, в зазоре создается поле с напряженностью $E = \varphi_{12}/d = (\Phi_1 - \Phi_2)/(ed)$, которому отвечает поверхностная плотность заряда $\sigma = \epsilon_0(\Phi_1 - \Phi_2)/(ed)$. Это значит, что вблизи контакта в одном проводнике возникнет обогащенный, а в другом — обедненный электронами слой толщины $L \sim \sigma/(en)$, где n — плотность носителей зарядов (электронов, дырок). Наличие этих слоев приводит к асимметрии контакта: сопротивление контакта становится неодинаковым для токов, текущих через контакт в противоположных направлениях.

Пусть один или оба контактирующих проводника относятся к полупроводникам с плотностью носителей $n \sim 10^{16} \div 10^{21}$ м⁻³. Обычно $|\Phi_1 - \Phi_2| \sim 1$ эВ, поэтому в полупроводнике $L \sim 5 \cdot 10^{-5} \div 5$ м. Иными словами, $L \gg d$, так что большая часть падения напряжения из-за контактной разности потенциалов приходится не на зазор, а на сам полупроводник. В результате происходит так называемое *искривление зон*: дно зоны проводимости ϵ_c и верхний край валентной зоны ϵ_v зависят от расстояния x до контакта. Эта зависимость изображена на рис. 9.10. Рис. 9.10, а относится к контакту с металлом донорного полупроводника (у которого работа выхода меньше, чем у металла), а рис. 9.10, б — к контакту металл—акцепторный полупроводник (у таких полупроводников, как правило, работа выхода больше, чем у металла).

Из-за больших размеров обогащенных и обедненных слоев в полупроводниках на эти слои легко можно влиять с помощью внешнего поля. Именно поэтому контактные явления в полупроводниках широко используются в первую очередь в твердотельной электронике (что же касается металлов, где $n \sim 10^{29}$ м⁻³, то в этом случае L определяется из квантовой механики и оказывается порядка атомных расстояний, так что контакты между металлами с этой точки зрения интереса не представляют).

Простейший прибор, работающий на основе полупроводниковых контактов, — это *полупроводниковый выпрямитель*. Пусть в кристалле



9.11.
Плоский транзистор

имеются пространственные области с электронной (n -область) и с дырочной (p -область) проводимостями. Если приложить к p -области отрицательное (относительно n -области) напряжение, то оно будет «вытягивать» дырки из n -области, а электроны — из p -области, так что возникнет ток. Если же, наоборот, приложить к p -области положительное напряжение, то pn -контакт «запрется» — ток в цепи будет отсутствовать. На этом принципе основано выпрямляющее действие полупроводникового кристалла.

Примесный полупроводник с двумя pn -контактами может усиливать слабые сигналы — его называют *транзистором* или кристаллическим триодом (рис. 9.11). Пусть для определенности первый и третий участки кристалла (эмиттер Э и коллектор К) обладают p -проводимостью, а средний участок (база Б) — n -проводимостью. Тогда при приложенном постоянном напряжении (рис. 9.11) контакт ЭБ заперт. «Отпирающий» его слабый переменный сигнал вызовет сильный импульс тока в коллекторной цепи; коэффициент усиления при этом может достигать многих сотен.

9.6. Формула Друде

Перейдем теперь к изучению электропроводности металлов и полупроводников. Свяжем прежде всего плотность электрического тока \mathbf{j} с функцией распределения электронов $f(\mathbf{p})$. По определению функции распределения плотности электронов, импульс которых лежит в интервале $(\mathbf{p}, \mathbf{p} + d\mathbf{p})$, равна $2f(\mathbf{p})d^3p/(2\pi\hbar)^3$, где $d^3p = dp_x dp_y dp_z$ — элемент объема в импульсном пространстве и множитель 2 учитывает две возможные ориентации спина электрона. Умножив эту величину на $e\mathbf{v}$, где e — заряд и \mathbf{v} — скорость электрона, мы найдем заряд, переносимый этими электронами за 1 с. Проинтегрировав полученное выражение по d^3p , определим плотность тока:

$$\mathbf{j} = 2e \int \mathbf{v} f(\mathbf{p}) \frac{d^3p}{(2\pi\hbar)^3}.$$

Эту формулу можно записать также в виде

$$\mathbf{j} = 2e \int \mathbf{v} f(\mathbf{p}) \nu(\epsilon) d\epsilon \frac{d\Omega}{4\pi},$$

где $\nu(\epsilon)d\epsilon d\Omega/(4\pi)$ — число состояний электрона в интервале энергий $(\epsilon, \epsilon + d\epsilon)$ и интервале телесных углов $d\Omega$ импульса \mathbf{p} .

В состоянии статистического равновесия распределение электронов изотропно, т. е. $f(\mathbf{p})$ зависит лишь от модуля импульса, а не от его направления. Поэтому в этом случае $f(-\mathbf{p}) = f(\mathbf{p})$, а так как скорость меняет свой знак вместе с импульсом, то интеграл, определяющий плотность тока, обращается в нуль. Такая ситуация имеет место в отсутствие поля; если же к проводнику приложено электрическое поле,

то возникает выделенное направление; функция распределения становится несимметричной, т. е. $f(\mathbf{p})$ отличается от $f(-\mathbf{p})$ и вектор плотности тока не равен нулю.

Отвлечемся сперва от того, что электроны находятся в кристаллической решетке, и будем рассматривать их как газ свободных частиц. Тогда скорость электрона \mathbf{v} пропорциональна его импульсу \mathbf{p} т. е. $\mathbf{v} = \mathbf{p}/m_e$, где m_e — масса электрона, и поэтому суммарный импульс электронов, отнесенный к единичному объему,

$$\mathbf{P} = 2 \int \mathbf{p} f(\mathbf{p}) \frac{d^3p}{(2\pi\hbar)^3} = 2m_e \int \mathbf{v} f(\mathbf{p}) \frac{d^3p}{(2\pi\hbar)^3}.$$

Сравнение этой формулы с формулой для плотности тока показывает, что $\mathbf{j} = (e/m_e)\mathbf{P}$, т. е. плотность тока пропорциональна плотности импульса. Отсюда вытекает важный вывод относительно изолированного газа электронов. Для такого газа сохраняется импульс, причем независимо от того, как взаимодействуют электроны друг с другом, — лишь бы они не взаимодействовали с внешним миром. Но в таком случае сохраняется и плотность тока, пропорциональная импульсу. Поэтому в системах отсчета, где импульс отличен от нуля, отлична от нуля и плотность тока, хотя к газу и не приложено внешнее электрическое поле. Это значит, что

изолированный газ электронов, как угодно взаимодействующих между собой (но не с окружающими телами), не обладает электрическим сопротивлением.

Если такой газ поместить в постоянное внешнее поле \mathbf{E} , то суммарный импульс электронов \mathbf{P} уже не будет сохраняться, а будет изменяться согласно уравнению $\dot{\mathbf{P}} = en\mathbf{E}$, где n — плотность электронов (внутренние силы взаимодействия между электронами в это соотношение не входят в силу закона равенства действия и противодействия). Отсюда следует, что

$$\frac{d}{dt} \mathbf{j} = \frac{e^2 n}{m_e} \mathbf{E}.$$

Это соотношение показывает, что

ток в изолированном газе электронов при включении постоянного электрического поля неограниченно возрастает, причем поле может быть сколь угодно слабым.

Этот вывод связан с двумя сделанными предположениями: во-первых, отношение e/m_e для всех частиц газа должно быть одинаковым и, во-вторых, газ должен быть замкнутой системой.

Чтобы проиллюстрировать и разъяснить смысл полученного результата, рассмотрим полностью ионизированную плазму (т. е. газ, состоящий из электронов и ионов с суммарным зарядом, равным нулю). Так как масса иона значительно больше массы электрона (по меньшей мере в 2000 раз), то ионы можно считать покоящимися. Поэтому импульс плазмы практически не отличается от импульса электронов, но этот импульс не сохраняется из-за взаимодействия электронов с ионами. По этой причине в плазме в отсутствие внешнего поля тока электронов

нет. При включении поля уравнение $\dot{\mathbf{P}} = en\mathbf{E}$ уже не выполняется, так как справа должна стоять еще сила, действующая на электроны со стороны ионов. Не выполняется также и уравнение $d\mathbf{j}/dt = (e^2n/m_e)\mathbf{E}$.

Если, однако, исходить из упрощенной картины взаимодействия электронов с ионами и считать, что в течение некоторого промежутка времени τ электрон движется как свободная частица, а в конце промежутка сталкивается с ионом и полностью теряет свой импульс, то уравнением $d\mathbf{j}/dt = (e^2n/m_e)\mathbf{E}$ можно пользоваться в течение промежутка времени τ . Поэтому ток не будет расти безгранично и определится формулой

$$\mathbf{j} = \frac{e^2n}{m_e} \mathbf{E} \langle \tau \rangle,$$

где $\langle \tau \rangle$ — некоторое среднее значение τ . Эту величину называют *средним временем свободного пробега электрона*. (В дальнейшем мы будем обозначать ее τ .) Мы получаем, таким образом, закон Ома с электрической проводимостью

$$\sigma = \frac{e^2n}{m_e} \tau.$$

Это соотношение называют *формулой Друде*.

Умножив τ на среднюю скорость v_T теплового движения электронов, получим *среднюю длину свободного пробега электрона*: $l = v_T\tau$. Поэтому коэффициент электропроводности можно записать и так:

$$\sigma = \frac{e^2nl}{m_e v_T}.$$

Чем больше длина свободного пробега, тем больше σ . Конкретно для плазмы

$$l \sim \frac{1}{n} \left(\frac{\epsilon_0 kT}{e^2} \right)^2,$$

а так как $v_T \sim \sqrt{kT/m_e}$, то

$$\sigma \sim \frac{(kT)^{3/2} \epsilon_0}{\sqrt{m_e} e^4}.$$

Мы видим, что электрическая проводимость плазмы сильно, как $T^{3/2}$, растет с увеличением температуры.

9.7. Кинетическое уравнение для электронов

Вернемся теперь к кристаллическим проводникам. В какой мере применимы к ним результаты, полученные для газа свободных электронов? Казалось бы, металл можно рассматривать как плазму, состоящую из электронов проводимости и ионов, и поэтому для удельной электрической проводимости металлов естественно пользоваться формулой $\sigma = (e^2n/m_e)\tau$. Но при этом возникает

вопрос: что такое τ и как определить эту величину? Существенно то, что, во-первых, ионы в кристалле расположены в строго определенном порядке, а не хаотически, как в плазме, и, во-вторых, то, что электроны в металле подчиняются статистике Ферми, а не статистике Максвелла—Больцмана, как в случае плазмы. Что касается первого обстоятельства, то с ним связана модификация понятия импульса электрона, движущегося в кристалле. А именно: в квантовой механике показывается, что состояния электронов с импульсами \mathbf{p} и $\mathbf{p}' = \mathbf{p} + 2\pi\hbar\mathbf{b}$ физически эквивалентны. Здесь $\mathbf{b} = m_1\mathbf{b}_1 + m_2\mathbf{b}_2 + m_3\mathbf{b}_3$ — так называемый *вектор обратной решетки* (\mathbf{b}_i — основные периоды обратной решетки, связанные с основными периодами кристаллической решетки $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ соотношением $\mathbf{b}_i\mathbf{a}_k = \delta_{ik}$, $\delta_{ik} = 1$ при $i = k$ и $\delta_{ik} = 0$ при $i \neq k$ и m_i — целые числа). Таким образом,

вектор \mathbf{p} в периодической решетке может быть определен только с точностью до $2\pi\hbar\mathbf{b}$.

По этой причине вектор \mathbf{p} называют не импульсом, а *квазиимпульсом*.

Невозможность точной фиксации импульса приводит к несохранению квазиимпульса в процессах взаимодействия электронов в решетке. Например, если в решетке сталкиваются два электрона с квазиимпульсами \mathbf{p}_1 и \mathbf{p}_2 и в результате взаимодействия (столкновения) их квазиимпульсы становятся равными \mathbf{p}'_1 и \mathbf{p}'_2 , то справедливо соотношение

$$\mathbf{p}_1 + \mathbf{p}_2 = \mathbf{p}'_1 + \mathbf{p}'_2 + 2\pi\hbar\mathbf{b},$$

а не обычный закон сохранения импульса $\mathbf{p}_1 + \mathbf{p}_2 = \mathbf{p}'_1 + \mathbf{p}'_2$. Поэтому если даже отвлечься от взаимодействия электронов проводимости с ионами кристаллической решетки, то уже само взаимодействие электронов друг с другом приводит (в отличие от случая плазмы) к конечной проводимости металла.

Второе обстоятельство — иная статистика электронов в металле по сравнению с плазмой — также очень существенно, потому что для определения вектора плотности тока, как мы видели, нужно знать функцию распределения электронов f , а для того, чтобы найти ее (в присутствии электрического поля она отличается от фермиевской функции распределения), необходим учет принципа Паули.

Разъясним метод нахождения функции распределения электронов $f(t, \mathbf{p})$ в электрическом поле \mathbf{E} . Рассмотрим для этого баланс числа электронов в элементе d^3p пространства \mathbf{p} . В момент времени t это число равно $f(t; \mathbf{p})d^3p/(2\pi\hbar)^3$ (предполагается заданной проекция спина электрона, распределение электронов считается пространственно однородным). Но из-за столкновений в этот элемент объема одни частицы будут входить, а другие — выходить. Кроме того, электроны ускоряются в поле \mathbf{E} , поэтому меняется их импульс (точнее, квазиимпульс). Именно: если в момент времени t квазиимпульс был равен \mathbf{p} , то в момент времени $t + dt$ он составит $\mathbf{p} + e\mathbf{E}dt$.

Обозначим $(\partial f/\partial t)_+^{\text{coll}}$ и $(\partial f/\partial t)_-^{\text{coll}}$ числа электронов, входящих и уходящих за 1 с из-за различных процессов столкновений

в единичный объем пространства \mathbf{p} (возле точки \mathbf{p}). Тогда в момент времени $t + dt$ число электронов в элементе объема $d^3p = dp_x dp_y dp_z$ пространства \mathbf{p} составит

$$\left\{ f(t; \mathbf{p}) + \left[\left(\frac{\partial f}{\partial t} \right)_+^{\text{coll}} - \left(\frac{\partial f}{\partial t} \right)_-^{\text{coll}} \right] dt \right\} \frac{d^3p}{(2\pi\hbar)^3}.$$

Так как частицы не рождаются и не исчезают, то эта величина равна $f(t + dt; \mathbf{p} + e\mathbf{E}dt)$. Но

$$f(t + dt; \mathbf{p} + e\mathbf{E}dt) \approx f(t; \mathbf{p}) + \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial \mathbf{p}} e\mathbf{E}dt.$$

Поэтому, приравняв оба выражения, получим

$$\frac{\partial f}{\partial t} + e\mathbf{E} \frac{\partial f}{\partial \mathbf{p}} = \left(\frac{\partial f}{\partial t} \right)^{\text{coll}},$$

где $(\partial f/\partial t)^{\text{coll}} = (\partial f/\partial t)_+^{\text{coll}} - (\partial f/\partial t)_-^{\text{coll}}$. Эту величину называют *интегралом столкновений*, а полученное соотношение — к и н е т и ч е с к и м у р а в н е н и е м.

Мы получили кинетическое уравнение для однородного распределения электронов, когда функция f не зависит от координат. Предполагалось также, что сила, действующая на электроны, — это электростатическая сила $e\mathbf{E}$. В более общем случае неоднородного распределения $f(t; \mathbf{p}, \mathbf{r})$ и произвольной силы \mathbf{F} число электронов в момент $t + dt$ вблизи точки $\mathbf{r} + d\mathbf{r}$ в элементе объема d^3p в \mathbf{p} -пространстве равно

$$\begin{aligned} & f(t + dt; \mathbf{r} + d\mathbf{r}; \mathbf{p} + \mathbf{F}dt) d^3p \approx \\ & \approx \left\{ f(t; \mathbf{r}, \mathbf{p}) + \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial \mathbf{p}} \mathbf{F}dt + \frac{\partial f}{\partial \mathbf{r}} \frac{d\mathbf{r}}{dt} dt \right\} d^3p; \end{aligned}$$

именно это выражение мы должны приравнять величине $\{f(t; \mathbf{r}, \mathbf{p}) + (\partial f/\partial t)^{\text{coll}}\} d^3p$. Замечая, что $d\mathbf{r}/dt = \mathbf{v}$ есть скорость частицы, мы приходим к о б щ е м у к и н е т и ч е с к о м у у р а в н е н и ю

$$\frac{\partial f}{\partial t} + \mathbf{v} \frac{\partial f}{\partial \mathbf{r}} + \mathbf{F} \frac{\partial f}{\partial \mathbf{p}} = \left(\frac{\partial f}{\partial t} \right)^{\text{coll}}.$$

Интеграл столкновений определяется конкретными процессами взаимодействий, в которых участвуют электроны. Простейшую структуру имеет интеграл столкновений, соответствующий столкновениям электронов с «чужими» (по отношению к кристаллу) атомами, или, как говорят, *примесными атомами*. Так как масса атома значительно больше массы электрона, то столкновения электрона с примесными атомами являются в основном упругими (без изменения энергии электрона). Если $w_{\text{im}}(\mathbf{p}, \mathbf{p}')$ — вероятность перехода за 1 с электрона из состояния \mathbf{p} в состояние \mathbf{p}' (она пропорциональна концентрации примесей), то число электронных переходов $\mathbf{p} \rightarrow \mathbf{p}'$ за 1 с есть

$$\left(\frac{\partial f}{\partial t} \right)_-^{\text{coll}} = \sum_{\mathbf{p}'} w_{\text{im}}(\mathbf{p}, \mathbf{p}') f'_{\mathbf{p}} [1 - f(\mathbf{p}')],$$

где множитель $1 - f(\mathbf{p}')$ учитывает принцип Паули: переход возможен только в том случае, если состояние \mathbf{p}' не занято. Суммирование здесь производится по

всем состояниям p' . Аналогично, число переходов $p' \rightarrow p$ определяется формулой

$$\left(\frac{\partial f}{\partial t}\right)_+^{\text{coll}} = \sum_{p'} w_{\text{im}}(p', p) f(p') [1 - f(p)].$$

Вычтя из этого выражения $(\partial f / \partial t)_{\text{im}}^{\text{coll}}$, мы получим интеграл столкновений $(\partial f / \partial t)_{\text{im}}^{\text{coll}}$, соответствующий столкновениям электронов с примесными атомами. При этом следует учитывать, что $w(p, p') = w(p', p)$. Поэтому

$$\begin{aligned} \left(\frac{\partial f}{\partial t}\right)_{\text{im}}^{\text{coll}} &= \sum_{p'} w_{\text{im}}(p, p') \{f(p') [1 - f(p)] - f(p) [1 - f(p')]\} = \\ &= \sum_{p'} w_{\text{im}}(p, p') (f(p') - f(p)). \end{aligned}$$

Обратим внимание на важнейшее свойство интеграла столкновений — он обращается в нуль для равновесных распределений частиц, т. е. в данном случае для фермиевского распределения электронов. Поэтому для малых отклонений функции распределения f от фермиевской функции f_0 интеграл столкновений линеен по $f - f_0$:

$$(\partial f / \partial t)_{\text{im}}^{\text{coll}} = -w_{\text{im}}(f - f_0),$$

где w_{im} — некоторое усредненное значение вероятности $w_{\text{im}}(p, p')$ (в следующем параграфе эта величина будет оценена).

Разъясним теперь структуру интеграла столкновений, соответствующего рассеянию электронов электронами. Рассмотрим два электрона с квазиимпульсами p и p' , которые в результате столкновения переходят в состояние с квазиимпульсами p'' и p''' . Обозначим $w(p, p'; p'', p''')$ вероятность того, что этот процесс произойдет за l с. Тогда число переходов $p, p' \rightarrow p'', p'''$ за l с составит

$$w(p, p'; p'', p''') f(p) f(p') [1 - f(p'')] [1 - f(p''')],$$

где множители типа $1 - f$ учитывают принцип Паули: электроны смогут перейти в состояния p'' и p''' только в том случае, если состояния не заняты. Вероятности процессов $p, p' \rightarrow p'', p'''$ и $p'', p''' \rightarrow p, p'$ одинаковы. Поэтому интеграл столкновений, соответствующий взаимодействию электронов друг с другом, может быть записан в виде

$$\begin{aligned} \left(\frac{\partial f}{\partial t}\right)_{\text{el}}^{\text{coll}} &= \sum_{p', p'', p'''} w(p, p'; p'', p''') \{f(p'') f(p''') [1 - f(p)] [1 - f(p')] - \\ &- f(p) f(p') [1 - f(p'')] [1 - f(p''')]\}. \end{aligned}$$

В случае равновесия фигурная скобка здесь обращается в нуль; в этом легко убедиться, подставив вместо f фермиевскую функцию f_0 и используя закон сохранения энергии $\varepsilon + \varepsilon' = \varepsilon'' + \varepsilon'''$. Поэтому при малых отклонениях f от f_0 интеграл столкновений линеен по $f - f_0$:

$$(\partial f / \partial t)_{\text{el}}^{\text{coll}} = -w_{\text{el}}(f - f_0),$$

где w_{el} — некоторое усредненное значение $w(p, p'; p'', p''')$. Так как в случае металлов из-за вырождения электронного газа электронные переходы возможны только в зоне размытости фермиевского распределения, а в процессе участвуют два электрона, то величина w_{el} в металлах пропорциональна квадрату температуры.

Мы рассмотрели интегралы столкновений, соответствующие взаимодействию электронов с примесями и рассеянию электронов электронами. В принципе,

нужно учитывать одновременно как эти, так и другие процессы взаимодействий, в которых участвуют электроны. Поэтому интеграл столкновений представляется в виде суммы:

$$(\partial f / \partial t)^{\text{coll}} = \sum_i (\partial f / \partial t)_i^{\text{coll}},$$

где индекс i нумерует различные процессы взаимодействия. Важным свойством интеграла столкновений является обобщение его в нуль для фермиевского распределения $f_0 = \{e^{(\epsilon - \mu)/(kT)} + 1\}^{-1}$. Поэтому при $|f - f_0| \ll f_0$ интеграл столкновений линеен по $f - f_0$:

$$(\partial f / \partial t)_i^{\text{coll}} = -\omega_i (f - f_0),$$

где ω_i — некоторые усредненные значения вероятностей различных процессов.

Итак, если функция распределения близка к равновесной, то

$$(\partial f / \partial t)^{\text{coll}} = -(f - f_0) / \tau, \quad 1/\tau \equiv \omega = \sum_i \omega_i.$$

Величина ω представляет собой сумму усредненных вероятностей различных процессов рассеяния, в которых участвует электрон, а обратная величина τ — среднее время между столкновениями. Ее называют *средним временем свободного пробега* электрона. Вводя среднее время свободного пробега $\tau_i = 1/\omega_i$ по отношению к i -процессу взаимодействия электронов, можно записать

$$1/\tau = \sum_i (1/\tau_i).$$

Величины ω_i и τ_i могут зависеть как от энергии электрона, так и от температуры, плотности примесных атомов и т. д.

Обратимся теперь к кинетическому уравнению для электронов и подставим в него приведенное выражение для $(\partial f / \partial t)^{\text{coll}}$, справедливое при $|f - f_0| \ll f_0$:

$$\frac{\partial f}{\partial t} + e\mathbf{E} \frac{\partial f}{\partial \mathbf{p}} = -\frac{1}{\tau} (f - f_0).$$

Это уравнение определяет функцию f , по которой можно вычислить плотность тока \mathbf{j} . Переходя к этой задаче, предположим сперва, что $\mathbf{E} = 0$. Тогда $\frac{\partial}{\partial t} \delta f = -\frac{1}{\tau} \delta f$, где $\delta f = f - f_0$. Поэтому $\delta f = (\delta f_0) e^{-t/\tau}$, где (δf_0) — начальное значение отклонения функции распределения от равновесной функции распределения. Мы видим, что в отсутствие поля любое отклонение от равновесной функции распределения затухает и по истечении времени $t \sim \tau$ становится очень малым, так что τ играет роль времени релаксации к равновесному распределению.

Рассмотрим теперь стационарное распределение электронов в металле, считая, что $\dot{\mathbf{E}} \neq 0$. Из кинетического уравнения следует тогда, что

$$f = f_0 - \tau e\mathbf{E} \frac{\partial f}{\partial \mathbf{p}}.$$

Так как предполагается, что $|f - f_0| \ll f_0$, то в члене, содержащем \mathbf{E} , можно заменить f на f_0 . Так как

$$\frac{\partial f_0}{\partial \mathbf{p}} = \frac{df_0}{d\varepsilon} \mathbf{v},$$

где $\mathbf{v} = \partial\varepsilon/\partial\mathbf{p}$ — скорость электрона, то

$$f = f_0 - \tau e v \mathbf{E} \frac{df_0}{d\varepsilon}.$$

Мы видим, что отклонение от равновесного распределения, вызываемое полем, пропорционально $v\mathbf{E}$. Но величина $df_0/d\varepsilon$ существенно отлична от нуля только в зоне размытости $\Delta\varepsilon \sim kT$ фермиевского распределения вблизи граничной энергии $\varepsilon = \mu$. Поэтому только в этой области поле существенно меняет функцию распределения электронов, вне же ее распределение электронов практически не отличается от фермиевского. Так как $(-df_0/d\varepsilon)$ имеет резкий максимум при $\varepsilon = \mu$, то при вычислении интегралов, содержащих произведение $df_0/d\varepsilon$ на произвольную гладкую функцию $F(\varepsilon)$, можно пользоваться правилом

$$-\int (df_0/d\varepsilon) F(\varepsilon) d\varepsilon = F(\mu).$$

Погрешность, которая при этом возникает, пропорциональна $(kT/\mu)^2$.

Вычислим теперь плотность тока. Подставив выражение для f в общую формулу для \mathbf{j} и учитывая, что слагаемое f_0 не вносит вклада в плотность тока, получим

$$\mathbf{j} = -2e^2 \int \mathbf{v}(\mathbf{v}\mathbf{E}) \tau \frac{df_0}{d\varepsilon} v(\varepsilon) d\varepsilon \frac{d\Omega}{4\pi}$$

($v(\varepsilon)$ — плотность уровней электрона). Используя резкий максимум функции $df_0/d\varepsilon$, можно выполнить здесь интегрирование по энергии:

$$\mathbf{j} = 2e^2 v_0^2 \tau_0 v(\mu) \int \mathbf{n}(\mathbf{n}\mathbf{E}) \frac{d\Omega}{4\pi},$$

где $\mathbf{n} = \mathbf{v}/v$ и индекс 0 обозначает, что соответствующая величина относится к значению $\varepsilon = \mu$. Интеграл по телесному углу Ω равен $1/3 \mathbf{E}$, так что

$$\mathbf{j} = \sigma \mathbf{E}, \quad \sigma = \frac{2}{3} e^2 v_0^2 \tau_0 v(\mu).$$

Этой общей формуле для коэффициента электропроводности металла можно придать вид формулы Друде, если считать, что $\varepsilon = p^2/(2m)$ (вообще говоря, связь между энергией электрона ε и его квазиимпульсом \mathbf{p} носит более сложный характер). В этом случае величину $d^3p = p^2 dp d\Omega$ можно представить в виде $d^3p = \sqrt{2\varepsilon m}^{3/2} d\varepsilon d\Omega$. Используя далее соотношение

$$\frac{d^3p}{(2\pi\hbar)^3} = v(\varepsilon) d\varepsilon \frac{d\Omega}{4\pi},$$

получим отсюда

$$v(\varepsilon) = \sqrt{2\varepsilon} m_e^{3/2} / (2\pi^2 \hbar^3).$$

Учтем, наконец, что $\varepsilon_F \approx \mu$ и

$$2 \frac{4\pi}{3} \frac{\rho_F^3}{(2\pi\hbar)^3} = n; \quad \frac{\rho_0^2}{2m_e} = \mu$$

(n — плотность электронов). Мы получим тогда формулу Друде

$$\sigma = e^2 n \tau_0 / m_e,$$

в которую входит значение времени релаксации τ при $\varepsilon = \mu$. Ее можно переписать в виде

$$\sigma = e^2 n l / (m_e v_0),$$

где $l = \tau_0 v_0$ — длина свободного пробега электрона, обладающего граничной скоростью v_0 . Так как обратные времена свободного пробега складываются, то

$$1/l = \sum_i (1/l_i),$$

где $l_i = \tau_i v_0$ — длина свободного пробега электрона по отношению к i -процессу рассеяния.

Приведем еще формулу для удельного сопротивления $\rho = 1/\sigma$:

$$\rho = \sum_i \rho_i; \quad \rho_i = m_e v_0 / (e^2 n l_i).$$

Здесь ρ_i представляет собой удельное сопротивление, обусловленное i -процессом рассеяния электронов.

9.8. Электропроводность металлов

Входящие в формулу для σ плотность электронов и их граничная скорость являются характерными для каждого металла константами и не зависят от внешних факторов, таких, как температура. Только длина свободного пробега может зависеть от этих факторов, и мы перейдем теперь к изучению этой зависимости. Рассмотрим сперва гипотетический идеализированный кристаллический металл, который обладает абсолютно правильной периодической решеткой, с неподвижными ионами (не совершающими никаких колебаний!), не содержит никаких примесей и в котором электроны не взаимодействуют между собой. Такой идеализированный кристалл можно рассматривать как гигантскую молекулу, и подобно тому, как в обычной молекуле движение электронов происходит без столкновений с ионами, так и в нашем идеализированном кристалле благодаря его абсолютной периодичности движение электронов происходит без столкновений с ионами (в отличие от случая плазмы, где ионы распо-

ложены хаотически). Кроме того, по предположению, электроны не сталкиваются друг с другом. Поэтому длина свободного пробега электрона в таком кристалле бесконечна, а следовательно, бесконечна и проводимость и равно нулю сопротивление.

Конечность длины свободного пробега электрона обусловлена неидеальностью решетки и взаимодействием электронов между собой. Неидеальность же решетки, в свою очередь, связана с наличием примесей и искажений, а также с тепловыми колебаниями решетки. Таким образом, следует различать

три основных механизма рассеяния электронов — рассеяние электронов примесями атомами, рассеяние электронов колебаниями решетки и рассеяние электронов электронами.

Простейшим является рассеяние электронов примесными атомами. Если Q — эффективное сечение рассеяния электрона атомом примеси ($Q \sim 10^{-18} \text{ м}^2$), то длина свободного пробега электрона $l_{\text{им}}$ по отношению к этому рассеянию может быть определена по газокинетической формуле $l_{\text{им}} = 1/(n'Q)$, где n' — плотность примесных атомов.

Рассмотрим теперь рассеяние электронов, обусловленное тепловыми колебаниями решетки. Колебания эти могут происходить с разными частотами ω , которые зависят от длины волны колебаний или волнового вектора \mathbf{f} ; кроме того, частота зависит от поляризации волны j , а поляризаций существует три — в простейшем случае изотропного тела одна поляризация соответствует продольным колебаниям, а две другие — поперечным. Поэтому частоту следует записывать в виде $\omega_j(\mathbf{f})$. В случае длинных волн, когда $fa \ll 1$ (a — постоянная решетки), частота $\omega_j(\mathbf{f}) = s_j f$ пропорциональна волновому числу f , здесь s_j — скорость звука, соответствующая поляризации j .

С колебаниями решетки можно связать определенные частицы — фононы, энергия $\epsilon_{\mathbf{f}j}$ и квазиимпульс \mathbf{q} которых определяется соотношениями

$$\epsilon_{\mathbf{f}j} = \hbar\omega_j(\mathbf{f}); \quad \mathbf{q} = \hbar\mathbf{f}.$$

Рассеяние электронов, обусловленное колебаниями решетки, сводится к двум основным процессам — испусканию и поглощению электроном фонона. В этих процессах выполняются законы сохранения:

$$\epsilon = \epsilon' \pm \epsilon_{\mathbf{f}j}, \quad \mathbf{p} = \mathbf{p}' \pm \mathbf{q} + 2\pi\hbar\mathbf{b},$$

где ϵ — энергия электрона с квазиимпульсом \mathbf{p} до рассеяния, ϵ' — энергия электрона с квазиимпульсом \mathbf{p}' после рассеяния и \mathbf{b} — вектор обратной решетки (напомним, что он должен входить в уравнение потому, что в решетке может быть определен только квазиимпульс, а не импульс, и это относится как к электрону, так и к фонону).

Интеграл столкновений $(\partial f/\partial t)_{\text{ph}}^{\text{coll}}$ электронов с решеткой может быть записан как интеграл столкновений электронов с фононами. При этом следует учитывать, что вероятность поглощения фонона пропорциональна числу имеющихся фононов $N_{\mathbf{f}j}$, а вероятность испускания — числу имеющихся фононов плюс один, т. е. $N_{\mathbf{f}j} + 1$. Поэтому интеграл столкновений имеет следующую структуру:

$$\left(\frac{\partial f}{\partial t}\right)_{\text{ph}}^{\text{coll}} = \sum_{\mathbf{p}', \mathbf{f}} \omega_+(\mathbf{p}'; \mathbf{p}, \mathbf{f}) \{f(\mathbf{p}') [1 - f(\mathbf{p})] (N_{\mathbf{f}} + 1) - f(\mathbf{p}) [1 - f(\mathbf{p}')] N_{\mathbf{f}}\} + \\ + \sum_{\mathbf{p}'', \mathbf{f}} \omega_-(\mathbf{p}'', \mathbf{f}; \mathbf{p}) \{f(\mathbf{p}') [1 - f(\mathbf{p})] N_{\mathbf{f}} - f(\mathbf{p}) [1 - f(\mathbf{p}')] (N_{\mathbf{f}} + 1)\}.$$

Здесь первая сумма определяет число электронных переходов $\mathbf{p}' \rightarrow \mathbf{p}$ за 1 с испусканием фотона \mathbf{f} (индекс j опущен) и число обратных переходов $\mathbf{p} \rightarrow \mathbf{p}'$ с поглощением фотона \mathbf{f} (отнесенная к 1 с вероятностью обоих процессов одинакова и обозначена $\omega_+(\mathbf{p}', \mathbf{p}, \mathbf{f})$). Аналогичным образом вторая сумма определяет число переходов $\mathbf{p}'' \rightarrow \mathbf{p}$ за 1 с с поглощением фонона \mathbf{f} и число обратных переходов $\mathbf{p} \rightarrow \mathbf{p}''$ с испусканием фонона \mathbf{f} (отнесенная к 1 с вероятностью обоих процессов обозначена $\omega_-(\mathbf{p}'', \mathbf{f}, \mathbf{p})$). Суммирования производятся по всем возможным $\mathbf{p}', \mathbf{p}'', \mathbf{f}$ с учетом закона сохранения энергии и модифицированного закона сохранения квазиимпульса.

В состоянии статистического равновесия $(\partial f/\partial t)_{\text{ph}}^{\text{coll}} = 0$. Используя закон сохранения энергии, легко убедиться, что функция распределения

$$f_0(\mathbf{p}) = \{e^{(\epsilon - \mu)/(kT)} + 1\}^{-1}$$

для электронов и функция распределения

$$N_0(\mathbf{f}) = \{e^{\hbar\omega/(kT)} - 1\}^{-1}$$

для фононов обращают интеграл столкновений в нуль.

Первую функцию распределения называют фермиевской, а вторую — планковской; напомним, что в первом случае говорят, что частицы подчиняются статистике Ферми — Дирака, а во втором — статистике Бозе — Эйнштейна.

Чтобы найти длину свободного пробега электрона l_{ph} , обусловленную его взаимодействием с решеткой, следует учесть, что частота фонона ω ограничена сверху: $\omega \leq \omega_0$, где ω_0 по порядку величины равна s/a (s — скорость звука, a — постоянная решетки). Если поэтому считать, что решетка находится в состоянии статистического равновесия, то возникает характерный параметр $\hbar\omega_0/(kT)$, или T_D/T , где

$$T_D = \hbar\omega_0/k = \hbar s/(ka)$$

(эту величину называют температурой Дебая, она составляет несколько сотен кельвин), и длина свободного пробега оказывается функцией T/T_D .

Особенно прост тот случай, когда $T \gg T_D$. При этом планковская функция сводится к $kT/(\hbar\omega)$ и поэтому $(\partial f/\partial t)_{\text{ph}}^{\text{coll}}$ оказывается пропорциональным, а длина пробега — обратно пропорциональной температуре. Более точно, если $T \gg T_D$, то

$$l_{\text{ph}} \sim \hbar v_0/(kT) \quad (T \gg T_D).$$

Длина пробега l_{ph} может быть оценена и в области низких температур: $T \ll T_D$. В этом случае она растет с убыванием температуры быстрее, чем T^{-1} , а именно как T^{-5} :

$$l_{\text{ph}} \sim \frac{\hbar v_0}{kT} \left(\frac{T_D}{T} \right) \quad (T \ll T_D).$$

Это связано как с уменьшением интенсивности колебаний решетки (число фононов при $T \ll T_D$ пропорционально $(T/T_D)^3$), так и с уменьшением эффективного угла рассеяния электрона (он пропорционален T/T_D). В промежуточной области температур l_{ph} является сложной функцией T/T_D , и мы не будем здесь ее приводить.

Рассмотрим, наконец, третий механизм рассеяния электронов, обусловленный их взаимодействием между собой. Числа электронных переходов $\mathbf{p}, \mathbf{p}' \rightarrow \mathbf{p}'', \mathbf{p}'''$ и $\mathbf{p}'', \mathbf{p}''' \rightarrow \mathbf{p}, \mathbf{p}'$ пропорциональны, очевидно, произведению ширин двух зон размытости для электронов в состояниях \mathbf{p}, \mathbf{p}' (или $\mathbf{p}'', \mathbf{p}'''$), т. е. квадрату температуры. Этой же величине пропорционален и интеграл столкновений, соответствующий рассеянию электронов электронами. Поэтому длина свободного пробега электрона l_{el} , обусловленного взаимодействием электронов друг с другом, обратно пропорциональна T^2 . Если считать взаимодействие между электронами чисто кулоновским, то

$$l_{\text{el}} \sim \frac{\hbar v_0 \varepsilon_F}{(kT)^2} \left(\frac{\varepsilon_0 \hbar v_0}{e^2} \right)^2 \sim \frac{\hbar v_0 \varepsilon_F}{(kT)^2}.$$

В действительности, однако, длина свободного пробега электронов, обусловленная их рассеянием друг на друге, может значительно превышать это значение. Дело в том, что взаимодействие между электронами может быть гораздо меньшим, чем чисто кулоновское взаимодействие, так как кулоновское отталкивание может существенно компенсироваться притяжением, обусловленным обменом фононами между электронами (об этом эффекте см. гл. 13).

Оценив длины пробегов электрона, обусловленные различными процессами, мы можем теперь выяснить, как ведет себя электропроводность или сопротивление металла в зависимости от температуры в концентрации примесей. Удельное сопротивление ρ определяется формулой

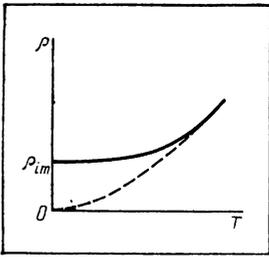
$$\rho = \rho_{\text{im}} + \rho_{\text{ph}} + \rho_{\text{el}},$$

где $\rho_{\text{im}}, \rho_{\text{ph}}, \rho_{\text{el}}$ — вклады в сопротивление, обусловленные рассеянием электронов примесными атомами, рассеянием электронов колебаниями решетки и взаимодействием электронов между собой. Они определяются формулами

$$\rho_{\text{im}} = \frac{m_e v_0}{e^2 n} \frac{1}{l_{\text{im}}}, \quad \rho_{\text{ph}} = \frac{m_e v_0}{e^2 n} \frac{1}{l_{\text{ph}}}, \quad \rho_{\text{el}} = \frac{m_e v_0}{e^2 n} \frac{1}{l_{\text{el}}}.$$

Что касается ρ_{im} , то эта величина не зависит от температуры и определяется концентрацией $c = n'/n$ примесных атомов:

$$\rho_{\text{im}} = \frac{\rho_F}{e^2 Q} c.$$



9.12.

Примерный ход зависимости удельного сопротивления металла от температуры

При низких температурах ($T \ll T_D$) и ρ_{ph} и ρ_{el} зависят от T (уменьшаясь с убыванием T), сопротивление же ρ_{im} от T не зависит. Поэтому при $T \ll T_D$ главную роль играет механизм рассеяния электронов примесными атомами. Обусловленное им сопротивление носит название *остаточного*. Величины ρ_{ph} и ρ_{el} при $T \ll T_D$ могут рассматриваться как небольшие температурные добавки к остаточному сопротивлению. Они определяются формулами

$$\rho_{ph} \approx \frac{m_e}{\hbar e^2} \frac{kT}{n} \left(\frac{T}{T_D} \right)^4 \quad (T \ll T_D),$$

$$\rho_{el} = \frac{m_e}{\hbar e^2} \frac{(kT)^2}{n \varepsilon_F} \xi \left(\frac{T}{T_D} \right),$$

где $\xi(T/T_D)$ — фактор, учитывающий уменьшение взаимодействия между электронами, обусловленное обменом фононами. Если бы этот фактор был порядка единицы, то главной поправкой к ρ_{im} при $T \ll \ll T_D$ было бы сопротивление $\rho_{el} \sim T^2$. Но для очень многих металлов $\xi \ll 1$ и поэтому $\rho_{ph} \gg \rho_{el}$. Примерный ход зависимости ρ от T показан на рис. 9.12. Обратим внимание на то, что остаточное сопротивление, пропорциональное концентрации примесных атомов, может служить мерой чистоты металла — на рисунке пунктиром показан ход кривой в случае чистого металла.

Заметим, в заключение этого параграфа, что электроны проводимости обуславливают не только электропроводность, но и теплопроводность металлов. Однако мы не будем изучать здесь явление теплопроводности, поскольку оно не относится непосредственно к электрическим и магнитным свойствам вещества.

9.9. Вольт-амперные характеристики полупроводников

Принципиально механизм проводимости одинаков и для металлов, и для полупроводников (заряд в тех и других переносится электронами), одинакова и причина электрического сопротивления: оно обуславливается теми же процессами взаимодей-

Так как $Q \sim 10^{-18} \text{ м}^2$ и $e^2/(\hbar \varepsilon_0 v_0) \sim 1$, то для оценки ρ_{im} можно пользоваться формулой $\rho_{im} \sim 10^{-14} \text{ с}$.

Величины ρ_{ph} и ρ_{el} существенно зависят от температуры, причем эта зависимость различна в областях высоких и низких температур. В области высоких температур ($T \gg T_D$) $\rho_{ph} \gg \rho_{el}$, причем

$$\rho_{ph} \sim \frac{m_e}{\hbar e^2} \frac{kT}{n} \quad (T \gg T_D),$$

т. е. ρ_{ph} растет линейно с T . Так как ρ_{im} не зависит от T , то при $T \gg T_D$ главную роль играет механизм рассеяния электронов, вызванного колебаниями решетки, и $\rho \sim T$.

ствия электронов с решеткой, с примесными атомами и друг с другом. Тем не менее зависимость тока от напряжения — *вольт-амперная характеристика* — для полупроводников может существенно отличаться от аналогичной зависимости для металлов (которая, как мы знаем, линейна и соответствует закону Ома $I = U/R$).

Такое отличие связано с основным различием между полупроводниками и металлами, заключающемся в том, что плотность носителей тока (электронов и дырок) в полупроводниках составляет обычно $10^{18} - 10^{25} \text{ м}^{-3}$ и, следовательно, на много порядков меньше плотности электронов в металлах (характерное значение которой 10^{29} м^{-3}). Поэтому газ носителей в полупроводниках, как правило, невырожден в отличие от вырожденного электронного газа в металлах.

Малость плотности электронов в полупроводниках позволяет упростить для них общее фермиевское распределение $f_0(\mathbf{p}) = (e^{(\epsilon_p - \mu)/(kT)} + 1)^{-1}$. Действительно, функция распределения f_0 в случае полупроводников мала ($f_0 \ll 1$), а это может быть только если слагаемое $e^{(\epsilon_p - \mu)/(kT)}$ велико по сравнению с единицей. Таким образом, равновесная функция распределения носителей в полупроводниках имеет вид

$$f_0(\mathbf{p}) \equiv f_M(\mathbf{p}) = e^{(\mu - \epsilon_p)/(kT)}.$$

Эту формулу называют *распределением Максвелла*. Ее можно переписать в виде $f_0(\mathbf{p}) = C e^{-\epsilon_p/(kT)}$, где $C \equiv e^{\mu/(kT)}$ — нормировочная постоянная, определяющаяся (как и химический потенциал μ) из условия

$$2 \int f_0(\mathbf{p}) d^3p / (2\pi\hbar)^3 = n_e$$

(n_e — плотность электронов). Если энергия электрона пропорциональна квадрату импульса: $\epsilon_p = p^2/(2m_e) = m_e v^2/2$, где v — скорость электрона (в этом случае говорят о *квадратичном законе дисперсии* и называют постоянной m_e *эффективной массой*), то формула для распределения Максвелла имеет такой же вид, как и в газокINETической теории:

$$f_0(\mathbf{p}) \propto e^{-m_e v^2/(2kT)}.$$

Если включить внешнее электрическое поле, то функция распределения электронов изменится. Чтобы найти ее, нужно воспользоваться кинетическим уравнением (см. § 9.7), в котором для случая полупроводников возникнет упрощение, а именно: так как $f \ll 1$, то множители типа $(1-f)$, учитывающие принцип Паули, могут быть заменены единицей.

Исследуя кинетическое уравнение для электронов в металле при наличии внешнего постоянного поля \mathbf{E} , мы считали, что изменение функции распределения, вызываемое полем, невелико, и искали функцию распределения в виде $f = f_0 + \delta f$, где δf — малая добавка к равновесной функции f_0 , линейно зависящая от \mathbf{E} . Такой подход

в случае металлов пригоден практически для любых сколь угодно сильных электрических полей. В случае же полупроводников такой подход годится только в очень слабых полях, так как даже умеренное поле существенно искажает функцию распределения электронов. Происходит это в первую очередь по той причине, что электрон, двигаясь в электрическом поле, может приобрести большую энергию. Эта энергия по порядку величины равна eEl , где l — средняя длина свободного пробега электрона. Такое выражение справедливо как для металлов, так и для полупроводников: различие лишь в том, с чем следует сравнивать эту величину.

В случае металлов ее нужно сравнивать с граничной фермиевской энергией ϵ_F , так как для проводимости существенны только те электроны, энергия которых близка к ϵ_F . Поэтому критерий слабости поля в случае металлов имеет вид $eEl \ll \epsilon_F$, или $eEl \ll kT_F$, где T_F — температура вырождения. Так как она очень велика ($T_F \sim 10^5$ К), то поля, меньшие $kT_F/(el) \sim 10^9$ В/м (мы положили для определенности $l \sim 10$ нм), должны в случае металлов считаться слабыми. Существенное изменение функции распределения электронов в металлах в электрическом поле должно наступать только при $E \geq kT_F/(el)$.

Ситуация существенно изменяется в случае полупроводников. Электронный газ в них невырожден, и значение приобретаемой в поле энергии eEl следует сравнивать со средней тепловой энергией электрона $\frac{3}{2} kT$. Поле, следует считать слабым, если $eEl \ll kT$, и сильным, если $eEl > kT$. Слабое поле мало искажает функцию распределения, в случае же сильного поля искажение будет большим.

Функция распределения электронов при наличии постоянного поля E имеет следующую структуру:

$$f(\mathbf{p}) = f_0(v) + \frac{v}{v} f_1(v),$$

где f_0 и f_1 — некоторые функции модуля скорости. В случае слабых полей $f_0(v)$ совпадает с распределением Максвелла f_M , для сильных же полей функция $f_0(v)$ существенно отличается от f_M . Это отличие определяется не только значением напряженности поля, но и характером зависимости длины пробега от скорости. При постоянной длине пробега, как это имеет место для рассеяния на атомах примеси, функция f_0 определяется формулой

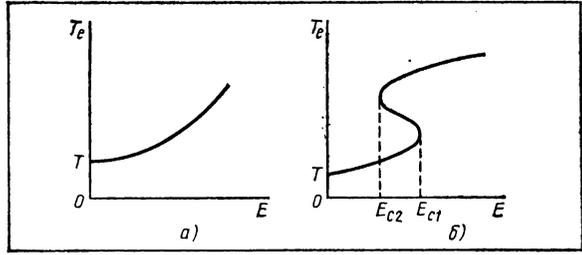
$$f_0(v) = C e^{-[m_0 v^2 / (2kT_e)]^2},$$

где C — нормировочная постоянная и T_e — так называемая электронная температура, равная $[m_i / (3m_0)]^{1/2} eEl$ (здесь m_i — масса атома примеси). Эта величина определяет среднюю тепловую энергию электрона при наличии поля, которая в рассматриваемом случае значительно превосходит $\frac{3}{2} kT$, где T — температура решетки.

Таким образом, в сильном поле электроны можно рассматривать как некоторый тепловой резервуар с температурой, много большей температуры решетки (рис. 9.13, а). Аналогичная ситуация может иметь место для газа, находящегося в электрическом поле; в этом случае возникают полностью аналогичные формулы. Вопрос о темпера-

9.13.

Зависимость электронной температуры полупроводника от напряженности приложенного поля



туре заряженных частиц газа при наличии электрического поля мы рассмотрим в § 11.2.

Векторная функция f_1 при $l = \text{const}$ имеет вид

$$f_1 = -\frac{E}{E} \sqrt{\frac{3m_0}{m_i} \frac{m_0 v^2}{kT_e}} f_0.$$

Она определяет плотность тока $j = 2e \int v f_1 d^3p / (2\pi\hbar)^3$. Вычисления показывают, что

$$j = C e n \frac{e E l}{\sqrt{m_0 T_e}},$$

где C — числовая константа.

Эта формула очень показательна. Если записать плотность тока в привычном для нас виде $j = \sigma E$, то для коэффициента электропроводности σ мы получим значение $\sigma = C e^2 n l / \sqrt{m_0 T_e}$. Знаменатель здесь представляет собой среднюю скорость электрона при наличии поля. Она пропорциональна \sqrt{E} . Поэтому σ зависит от напряженности поля, как $E^{-1/2}$. Мы видим, что в сильном поле в полупроводниках нарушается закон Ома. При $E \rightarrow \infty$ коэффициент электропроводности стремится к нулю. Аналогичная картина имеет место и при газовом разряде (см. гл. 11). Но приведенные нами формулы не являются универсальными: они справедливы в предположении, что $l = \text{const}$. Между тем постоянство длины пробега, как правило, не имеет места. Поэтому для функций f_0 и f_1 нет универсальных формул — эти функции существенным образом зависят от того, как меняется со скоростью длина свободного пробега электрона.

Дело осложняется еще и тем, что одной средней длины свободного пробега электрона вообще недостаточно для описания электропроводности полупроводников и определения их вольт-амперной характеристики. Это, собственно, касается и электропроводности металлов, поскольку в общее кинетическое уравнение для электронов входят интегралы столкновений, определяемые вероятностями различных процессов взаимодействия электронов. Когда мы писали формулу Друде, в которую входит среднее время свободного пробега электрона, то в основном имели в виду получение общей качественной картины электропроводности. В действительности, если уточнять картину, в формулу Друде входит средняя вероятность изменения за l с направ-

ления импульса электрона, которую следует, вообще говоря, отличать от средней вероятности изменения энергии электрона.

Если рассматривать слабые поля, то температура электронного газа в полупроводнике не отличается от температуры решетки и имеет место формула Друде для электропроводности $\sigma = e^2 n \tau / m_0$, в которой под τ нужно понимать обратную среднюю вероятность процесса изменения за 1 с направления импульса электрона. Эта величина является функцией температуры электронного газа T_e , которая в случае слабых полей совпадает с температурой решетки. Поэтому мы получаем закон Ома $j = \sigma E$ с не зависящим от E коэффициентом электропроводности σ .

Что произойдет, если поле станет сильным? Температура электронов T_e станет функцией напряженности поля E , поэтому от поля начнет зависеть величина $\tau \equiv \tau(T_e)$ — время релаксации электронов по импульсу — и σ станет функцией поля, т. е. перестанет быть справедливым закон Ома. Что же касается формулы

$$\sigma \equiv \sigma(T_e) = \frac{e^2 n}{m_0} \tau(T_e),$$

то она останется справедливой.

Вопрос теперь заключается в том, чем определяется температура электронов. Она определяется балансом энергии электрона, получаемой от поля и отдаваемой при рассеянии на атомах примеси или колебаниях решетки.

От поля электрон получает за 1 с энергию

$$jE/n = \sigma E^2/n = e^2 E^2 \tau(T_e)/m_0.$$

Средняя энергия электрона $\langle \epsilon \rangle$ равна $^{3/2} k T_e$. Поэтому энергия, теряемая за 1 с, равна $^{3/2} k (T_e - T)/(\tau_e(T_e))$, где $1/\tau_e$ — вероятность передачи энергии за 1 с. Уравнение баланса энергии имеет вид

$$\frac{d}{dt} \langle \epsilon \rangle = \frac{e^2 E^2 \tau(T_e)}{m_0} - \frac{3k(T_e - T)}{2\tau_e(T_e)}.$$

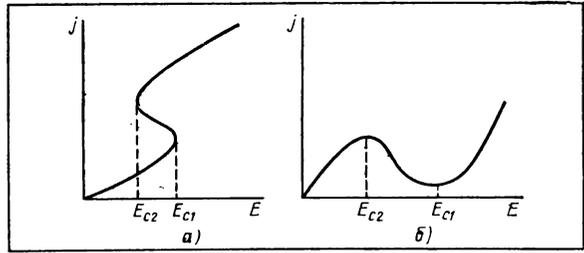
В стационарном случае $d\langle \epsilon \rangle/dt = 0$, и мы получаем следующее уравнение для определения электронной температуры как функции напряженности поля E :

$$\frac{e^2 E^2 \tau(T_e)}{m_0} = \frac{3k(T_e - T)}{2\tau_e(T_e)}.$$

Чтобы найти из него T_e , нужно знать функции $\tau(T_e)$ и $\tau_e(T_e)$. Они определяются конкретными механизмами рассеяния электронов и аппроксимируются обычно степенными функциями $\tau_e \sim T_e^s$, $\tau \sim T_e^q$. Показатели степеней имеют следующие значения: для взаимодействия со звуковыми колебаниями в обычных кристаллах $s = q = -1/2$ и в пьезоэлектрических кристаллах $s = q = 1/2$; для рассеяния заряженной примесью $q = 3/2$ и нейтральной примесью $q = 0$ (передача энергии при рассеянии примесью пропорциональна m_0/m_i и потому пренебрежимо мала).

9.14.

Зависимость плотности тока в полупроводнике от напряженности приложенного поля



Существенно, что написанное уравнение для определения T_e (при заданном E) может иметь несколько решений; в этом случае плотность тока является многозначной функцией напряженности поля. Например, возможна ситуация, когда в некотором интервале полей $E_{c2} < E < E_{c1}$ температура T_e — трехзначная функция E . Тогда и плотность тока — трехзначная функция напряженности поля, и мы получаем так называемую S-образную вольт-амперную характеристику (рис. 9.13, б, 9.14, а). При малых E она имеет прямолинейный участок, соответствующий закону Ома. Но на ней же есть участок, на котором плотность тока падает с ростом напряженности поля: участок, на котором дифференциальная проводимость $\sigma_d = dj/dE$ отрицательна.

Легко вывести формулу для дифференциальной проводимости:

$$\sigma_d = \frac{dj}{dE} = \sigma \frac{1 - \left(1 - \frac{T}{T_e}\right) \frac{d \ln(\tau_\epsilon / \tau)}{d \ln T_e}}{1 - \left(1 - \frac{T}{T_e}\right) \frac{d \ln(\tau_\epsilon / \tau)}{d \ln T_e}}$$

(σ — коэффициент электропроводности в слабом поле). Эта функция при подходящем виде функций $\tau(T_e)$ и $\tau_\epsilon(T_e)$ может в достаточно сильных полях E_{c1} и E_{c2} обратиться в бесконечность и иметь в области $E_{c2} < E < E_{c1}$ ветвь $\sigma_d < 0$. В этом случае мы и приходим к S-образной вольт-амперной характеристике. Условие ее возникновения имеет вид

$$\frac{d(\ln \tau \tau_\epsilon)}{d \ln T_e} > \frac{T_e}{T_e - T} > 1.$$

Такой механизм возникновения S-образной характеристики наблюдался, например, в полупроводнике n -InSb. Импульс носителей тока передавался в основном заряженной примеси ($q_s = 3/2$), а их энергия — звуковым колебаниям ($s = 1/2$, так как рассматриваемый кристалл является пьезоэлектриком).

Вольт-амперная характеристика полупроводника может быть и N-образной (рис. 9.14, б). В этом случае дифференциальная проводимость σ_d меняет знак при $E = E_{c1}$ и $E = E_{c2}$, обращаясь не в бесконечность, а в нуль. Условие возникновения N-образной характеристики имеет вид

$$\frac{d \ln (\tau_e / \tau)}{d \ln T_e} > \frac{T_e}{T_e - T} > 1.$$

Отсюда следует, что должно, в частности, выполняться неравенство $s - q > 1$.

Мы рассмотрели один из механизмов, приводящих к невыполнению в полупроводниках закона Ома; его называют *перегревным* механизмом. Известны еще и другие механизмы, например *концентрационный* механизм, заключающийся в том, что приложенное электрическое поле меняет концентрацию носителей тока. Концентрационным механизмом объясняется, в частности, N-образная вольт-амперная характеристика, наблюдающаяся в германии с примесью меди.

Полупроводники с N-образной вольт-амперной характеристикой нашли широкое применение в качестве генераторов сверхвысоких частот (десятки гигагерц). Энергия высокочастотного поля в таких генераторах черпается при этом из энергии постоянного электрического поля, напряженность которого должна лежать в области отрицательной дифференциальной проводимости, $E_{c2} < E < E_{c1}$ (обычно это $10^2 - 10^5$ В/м).

Полупроводники с S-образной вольт-амперной характеристикой используются в качестве переключающих устройств, например в элементах памяти ЭВМ. Их достоинство состоит в том, что время переключения с одной ветви вольт-амперной характеристики на другую очень мало ($\sim 10^{-10}$ с).

ОСНОВНЫЕ ФОРМУЛЫ

Распределение Ферми—Дирака

$$f = \frac{1}{e^{(s-\mu)/(kT)} + 1}$$

Распределение Бозе—Эйнштейна (для фононов или фотонов)

$$N = \frac{1}{e^{\hbar\omega/(kT)} - 1}$$

Ток термоэлектронной эмиссии (формула Ричардсона—Дешмана)

$$j_s = CT^2 e^{-\Phi/(kT)}$$

Температура вырождения электронного газа

$$T_F = (3\pi^2)^{2/3} \frac{\hbar^2 n^{2/3}}{2mk}$$

Теплоемкость электронного газа (отнесенная к одному электрону)

$$c \sim kT/T_F$$

Кинетическое уравнение для электронов

$$\frac{\partial f}{\partial t} + \mathbf{v} \frac{\partial f}{\partial \mathbf{r}} + \mathbf{F} \frac{\partial f}{\partial \mathbf{p}} = \left(\frac{\partial f}{\partial t} \right)^{\text{coll}}$$

Формула Друде

$$\sigma = \frac{e^2 n}{m} \tau, \quad \tau = \frac{l}{v_0}, \quad \frac{1}{l} = \sum_i \frac{1}{l_i}, \quad l_i = v_0 \tau_i$$

Длина пробега электрона по отношению к взаимодействию с решеткой

$$l_{ph} = \begin{cases} \hbar v_0 / (kT) & (T \gg T_D) \\ \frac{\hbar v_0}{kT} \left(\frac{T_D}{T} \right)^4 & (T \ll T_D) \end{cases}$$

Длина пробега электрона по отношению к взаимодействию с электронами

$$l_{ee} = \frac{\hbar v_0 \epsilon_F}{(kT)^2}$$

Длина свободного пробега электрона по отношению к расстоянию примесными атомами

$$l_{im} = \frac{1}{Qn'}$$

Электронная температура в электрическом поле

$$T_e = \left(\frac{m_i}{3m_e} \right)^{1/2} eEl$$

Электропроводность в сильном поле

$$\sigma \equiv \sigma(T_e) = \frac{e^2 n}{m} \tau(T_e)$$

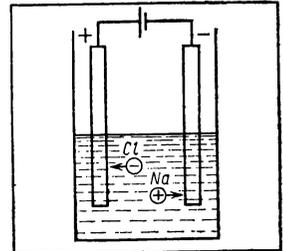
■

Глава 10. ЭЛЕКТРОЛИТЫ

10.1. Сильные и слабые электролиты

Если растворить в воде поваренную соль, то ее молекулы распадутся на положительные Na^+ и отрицательные Cl^- ионы. Поэтому, опустив в такой раствор два металлических электрода и присоединив их к источнику э. д. с. (рис. 10.1), мы получим электрическую цепь, через которую пойдет электрический ток. Ясно, что отрицательно заряженные ионы Cl^- (анионы) будут двигаться к положительному электроду — аноду, а положительно заряженные ионы Na^+ (катионы) — к отрицательному электроду — катоду. Свойством проводимости обладает не только раствор NaCl , но и растворы других солей, а также кислот и оснований. Такие вещества называют *электролитами* и говорят об *электролитической проводимости*.

В отличие от металлов, проводимость которых обусловлена электронами,



10.1. Электролитическая проводимость

в растворах электролитов ток представляет собой направленное движение ионов, возникающих вследствие распада, или диссоциации, молекул электролита в растворе.

Процесс диссоциации молекул вполне естествен, так как при протекании любой химической реакции $A + B \rightarrow C$ всегда идет и обратная реакция $C \rightarrow A + B$. Поэтому если молекула создается в результате соединения ионов разных знаков заряда, то всегда вместе с нейтральными молекулами должны находиться и составляющие их ионы.

Замечательным является то, что в ряде случаев возникает почти стопроцентная диссоциация молекул. Именно так обстоит дело при растворении в воде многих солей, сильных кислот и сильных оснований (если только раствор не слишком концентрирован). Такие вещества, очень сильно диссоциирующие в растворе, называют *сильными электролитами*. В отличие от них слабые кислоты и слабые основания, например амины, фенолы, большинство органических кислот, некоторые неорганические кислоты и основания (например, синильная кислота, гидрат окиси аммония), а также некоторые соли, например хлорная и цианистая ртуть, диссоциируют при обычных концентрациях в очень небольшой степени. Эти вещества называют *слабыми электролитами*. Соли слабых кислот и сильных оснований или слабых оснований и сильных кислот, а также соли слабой кислоты и слабого основания представляют собой, как правило, сильные электролиты.

Ясно, что растворение электролита в воде должно способствовать диссоциации молекул электролита. Дело в том, что вода обладает большой диэлектрической проницаемостью ($\epsilon \approx 80$); поэтому сила взаимодействия между ионами в молекуле электролита, находящейся в воде, ослабевает примерно в 80 раз.

Заметим, однако, что в ряде случаев диэлектрическая проницаемость растворителя играет второстепенную роль. Например, хлористый водород растворяется в этиловом спирте и образует сильный электролит, между тем раствор хлористого водорода в нитробензоле, диэлектрическая проницаемость которого почти такая же, как и этилового спирта, является слабым электролитом. Существенным здесь оказывается то обстоятельство, что молекула этилового спирта легко соединяется с ионом водорода, образуя комплексный ион $C_2H_5OH_2^+$, который и является формой существования иона водорода в спиртовом растворе. Нитробензол не образует такого соединения, поэтому степень диссоциации оказывается малой и раствор хлористого водорода в нитробензоле является слабым электролитом. Так как молекулы сильных электролитов в растворах почти полностью диссоциированы, то в реакциях между сильными электролитами в растворах участвуют фактически не молекулы, а отдельные ионы.

В отличие от сильных электролитов, почти все молекулы которых диссоциированы вплоть до очень больших концентраций раствора, степень диссоциации слабых электролитов существенно зависит от концентрации раствора, возрастая с ее уменьшением. Чтобы установить эту зависимость, воспользуемся законом действующих масс. Сформулируем этот закон. Пусть происходит какая-либо химическая реакция $\nu_1 A_1 + \nu_2 A_2 + \dots = \nu'_1 A'_1 + \nu'_2 A'_2 + \dots$, где A_1, A_2, \dots — символы химических веществ, вступающих в реакцию, A'_1, A'_2, \dots — символы веществ, возникающих в результате реакции, и $\nu_1, \nu_2, \dots, \nu'_1, \nu'_2, \dots$ — целые числа, показывающие количества молекул, участвующих в реакции. (Например, для реакции $2H_2 + O_2 \rightarrow 2H_2O$ имеем $\nu_{H_2} = 2, \nu_{O_2} = 1, \nu_{H_2O} = 2$.) Если реакция происходит в смеси идеальных газов, то после

достижения равновесия справедливо соотношение

$$\frac{[A_1]^{v_1} [A_2]^{v_2} \dots}{[A'_1]^{v'_1} [A'_2]^{v'_2} \dots} = K(p, T),$$

где $[A_i]$, $[A'_i]$ — концентрации соответствующих веществ,

$$K(p, T) = p^{-(v_1+v_2+\dots - v'_1 - v'_2 - \dots)} K(T),$$

p — давление и $K(T)$ — величина, зависящая только от температуры T и называемая *константой химического равновесия*. В частности, для газа с постоянной теплоемкостью $K(T) = K_0 T^n e^{-\omega/(kT)}$, где K_0 , n и ω — постоянные. Закон действующих масс справедлив и для реакций между растворенными веществами, если раствор является слабым, только в этом случае $K(p, T)$ зависит от давления более сложным образом.

Применим закон действующих масс к реакции диссоциации $M \rightarrow A^+ + B^-$ молекулы M на ионы A^+ и B^- ; получим $[A^+][B^-]/[M] = K$. Пусть всего растворено c молей электролита в литре воды; тогда $[A^+] = [B^-] = \alpha c$, $[M] = (1-\alpha)c$, где α называют *степенью диссоциации*. Тогда $c\alpha^2/(1-\alpha) = K$, откуда

$$\alpha = 2K(K + \sqrt{K^2 + 4Kc})^{-1}.$$

Мы видим, что степень диссоциации растет с уменьшением концентрации и стремится к единице при $c \rightarrow 0$.

10.2. Электролиз

Рассмотрим теперь подробнее, как проходит ток через раствор электролита. Так как всегда при этом имеются металлические электроды, погруженные в раствор электролита, то электрическая цепь состоит не только из раствора электролита, но и из металлических проводников, переносчики же заряда в них разные: в растворе — это ионы, а в металле — электроны. Отсюда вытекают важные следствия. Действительно, свободные ионы в растворе электролита не могут проходить через металл, а электроны металла не могут двигаться в электролите. Но ток представляет собой непрерывное течение электрического заряда, поэтому на границе электронного и ионного проводников должно происходить существенное превращение, нечто вроде «переодевания» ионов: на аноде анионы должны отдавать свои лишние электроны аноду и превращаться в нейтральные атомы (или группы атомов), а на катоде катионы должны получать добавочные электроны и также превращаться в нейтральные атомы. Только при таком превращении суммарный ионный ток в электролите будет равен электронному току в металлической части цепи. Таким образом,

прохождение тока через цепь, содержащую электронные и ионные проводники, должно сопровождаться химическим разложением электролита.

Это замечательное явление носит название *электролиза* (оно было открыто Фарадеем в 1833 г.).

Легко определить количество веществ, выделяющихся при электролизе на электродах. Для этого достаточно использовать закон сохра-

нения заряда. Действительно, рассмотрим ионы с зарядом ze (z — валентность, e — элементарный заряд). Ясно, что если через цепь пройдет заряд Q , то на каждом из электродов образуется $Q/(ze)$ ионов. Возьмем теперь один моль ионов, в нем содержится $N_0 = 6,02 \cdot 10^{23}$ ионов. Поэтому на каждом из электродов выделится по $Q/(zeN_0)$ молей ионов, т. е. по $Q/(zeN_0)$ молей атомов нейтрального вещества. Если A — масса его атома, то $M_A = AQ/(zeN_0)$ — количество вещества в граммах, выделяющееся на электроде при прохождении заряда Q . Входящая сюда величина eN_0 представляет собой универсальную константу — это заряд одного моля электронов (или моля протонов). Ее обозначают F и называют *постоянной Фарадея*. Заряд этот равен $F = eN_0 = 96484$ Кл/моль. Замечая, что $Q = It$, где I — сила тока, проходящего в течение времени t через цепь, можно переписать выражение для M_A в виде

$$M_A = \frac{A}{z} \frac{It}{F}.$$

Это соотношение известно под названием *закона электролиза Фарадея*.

10.3. Электропроводность растворов

Вясним теперь, чем определяется электропроводность растворов электролитов.

Если к раствору приложено постоянное электрическое поле \mathbf{E} , то на ион сорта a с зарядом ez_a действует сила $ez_a\mathbf{E}$. Однако ясно, что скорость иона не может все время возрастать, так как движение иона происходит с трением. Сила трения пропорциональна скорости иона \mathbf{v}_a , т. е. ее можно представить в виде $-\gamma_a\mathbf{v}_a$, где γ_a — коэффициент трения. В результате совместного действия электрической силы и силы трения устанавливается стационарное движение иона, при котором суммарная сила равна нулю: $ez_a\mathbf{E} - \gamma_a\mathbf{v}_a = 0$. Отсюда $\mathbf{v}_a = b_a ez_a\mathbf{E}$, где $b_a = 1/\gamma_a$. Таким образом, ионы движутся с постоянной скоростью, пропорциональной действующей на них электрической силе. Коэффициент пропорциональности b_a носит название *подвижности* иона.

Скорости ионов, как правило, невелики. Например, при поле $0,01 \text{ В} \cdot \text{м}^{-1}$ и температуре 25°С большинство ионов в воде движутся со скоростью порядка $5 \cdot 10^{-6}$ м/с. Исключение составляет водород, ион которого движется со скоростью $36,2 \cdot 10^{-6}$ м/с.

Умножив скорость \mathbf{v}_a на $ez_a n_a$, где n_a — концентрация ионов сорта a , найдем часть плотности тока, обусловленную этими ионами. Поэтому суммарная плотность тока

$$\mathbf{j} = \sum_a ez_a n_a \mathbf{v}_a = \sum_a b_a (ez_a)^2 n_a \mathbf{E}.$$

Мы видим, что *плотность тока пропорциональна напряженности поля*, т. е. растворы электролитов подчиняются закону Ома, а их удельная электрическая проводимость определяется формулой

$$\sigma = \sum_a b_a (ez_a)^2 n_a,$$

где суммирование распространяется на все сорта ионов. В простейшем случае электролита, молекулы которого диссоциируют на два одновалентных иона, эта формула дает $\sigma = e^2(b_+n_+ + b_-n_-)$, где $n_+ = n_-$ — концентрации ионов каждого знака заряда, а b_+ и b_- — их подвижности.

В таком общем виде эта формула справедлива и для слабых, и для сильных электролитов, но она приводит к разным зависимостям электропроводности от концентрации раствора для двух типов электролитов.

Рассмотрим сперва растворы слабых электролитов. Для них, очевидно, $n_+ = n_- = \alpha c N_0$, где α — степень диссоциации и c — молярная концентрация, т. е. отношение числа молей электролита к объему растворителя. Поэтому коэффициент электропроводности слабого электролита определяется формулой

$$\sigma = e^2 \alpha c N_0 (b_+ + b_-),$$

где согласно результатам § 10.1 степень диссоциации является функцией концентрации

$$\alpha = 2K (K + \sqrt{K^2 + 4Kc})^{-1}.$$

Что касается подвижностей b_+ и b_- , то они в случае слабого раствора могут считаться не зависящими от концентрации c . В электрохимии часто пользуются понятием *эквивалентной электрической проводимости* $\sigma' = \sigma/c$. Эта величина зависит от концентрации так же, как и степень диссоциации. При бесконечном разведении ($c \rightarrow 0$) степень диссоциации стремится к единице, поэтому $\sigma'_{c \rightarrow 0} \rightarrow e^2 N_0 (b_+ + b_-)$.

Рассмотрим теперь растворы сильных электролитов. Для них степень диссоциации равна единице (вплоть до очень больших концентраций), подвижности же нельзя считать не зависящими от концентрации. Действительно, так было бы, если бы подвижность определялась только силой трения, действующей на ионы со стороны молекул растворителя. Но в случае растворов сильных электролитов, когда имеется много ионов, становится существенным их взаимодействие друг с другом, приводящее к добавочному торможению ионов. Механизм этого торможения связан с тем, что каждый ион в электролите окружен атмосферой ионов противоположного знака заряда, которую он «тянет» с собой при движении. Но поле действует на нее с силой, направленной в противоположную сторону по отношению к силе, действующей на центральный ион; в результате возникает добавочное торможение иона, называемое *электрофоретической силой*.

Для того чтобы вычислить эту силу, нужно определить размеры ионной атмосферы, окружающей данный ион. Полагая, что ионы в электролите находятся в тепловом равновесии, мы должны считать, что в 1 м^3 объема содержится согласно распределению Больцмана

$$n_i(\mathbf{r}) = n_{i0} e^{-q_i \varphi(\mathbf{r}) / (kT)}$$

ионов сорта i (здесь $\varphi(\mathbf{r})$ — электростатический потенциал, q_i — заряд иона, n_{i0} — постоянная). Обычно электрическая энергия иона мала по сравнению с тепловой; $q\varphi \sim q^2 / (4\pi\epsilon_0 \langle r \rangle) \ll kT$ ($\langle r \rangle \sim n^{-1/3}$ — среднее расстояние между ионами). Это позволяет разложить больцмановскую экспоненту в ряд

$$n_i(\mathbf{r}) = n_{i0} [1 - q_i \varphi(\mathbf{r}) / (kT)].$$

Найдем теперь объемную плотность заряда: $\rho = \sum_i q_i n_i$. Учитывая, что в целом раствор электрически нейтрален, т. е. $\sum_i q_i n_{i0} = 0$, получим $\rho = -\varphi / (kT) \sum_i q_i^2 n_{i0}$. Подставляя это выражение в уравнение Пуассона $\epsilon_0 \Delta \varphi = -\rho$, имеем

$$\Delta \varphi = \frac{1}{r_D^2} \varphi; \quad r_D^2 = \frac{\epsilon_0 kT}{\sum_i q_i^2 n_{i0}}.$$

Величину r_D называют *радиусом Дебая*. В терминах r_D использованное нами условие малости электрической энергии имеет вид $\langle r \rangle \ll r_D$; в сфере дебаевского радиуса должно находиться много ионов.

Решим полученное уравнение относительно φ . Для этого учтем, что распределение ионов в атмосфере сферически симметрично и поэтому

$$\Delta \varphi = \frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d\varphi}{dr} \right).$$

В результате получим $\varphi = \text{const} \cdot \frac{1}{r} e^{-r/r_D}$. Но вблизи иона ($r \rightarrow 0$) потенциал не может отличаться от кулоновского, поэтому $\text{const} = q_i$, где q_i — заряд рассматриваемого иона. Таким образом, ион создает потенциал

$$\varphi = \frac{q_i}{4\pi\epsilon_0 r} e^{-r/r_D}.$$

Его называют *экранированным потенциалом*, поскольку он существенно отличен от нуля только на расстояниях, не превышающих r_D (в этой связи дебаевский радиус называют также *радиусом экранировки*).

Мы видим, что размеры ионной атмосферы имеют порядок r_D . Рассматривая эту атмосферу как шарик радиуса r_D , движущийся в среде с некоторой вязкостью η , можно, согласно формуле Стокса, найти скорость \mathbf{u} шарика: $\mathbf{u} = q_i \mathbf{E} / (6\pi\eta r_D)$, где $q_i \mathbf{E}$ — сила, действующая на ион. Умножив, эту скорость на коэффициент трения γ , обусловленного чистым растворителем, получим электрофоретическую силу:

$$\mathbf{f} = \frac{-\gamma}{6\pi\eta r_D} q_i \mathbf{E}.$$

В результате действия электрофоретической силы подвижность иона уменьшается и это уменьшение оказывается пропорциональным r_D^{-1} , т. е. корню квадратному из концентрации. В свою очередь, уменьшение подвижности приводит к уменьшению эффективной электропроводности раствора сильного электролита, которую можно представить в виде

$$\sigma' = \sigma'_0 \left(1 - \sqrt{\frac{c}{c_0}} \right),$$

где σ'_0 — эффективная электропроводность при бесконечном разведении ($c \rightarrow 0$) и c_0 — некоторая константа, которую мы здесь не выписываем.

10.4. Гальванические элементы

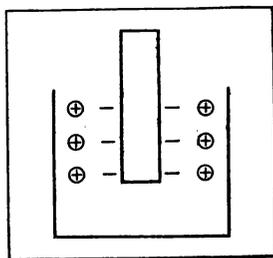
В гл. 9 мы показали, что между двумя соприкасающимися металлами или полупроводниками всегда возникает определенная разность потенциалов (контактная разность потенциалов). Легко видеть, что контактная разность потенциалов должна возникать и на границе металла с раствором электролита, а также на границе между двумя растворами электролитов. Действительно, если металл находится в воде (или в растворе электролита), то решетка поверхностной части металла, граничащей с водой, диссоциирует и ее положительно заряженные ионы уходят в растворитель, сам же металл заряжается отрицательно, так как в нем образуется избыток электронов. Процесс диссоциации не приводит, однако, к полному растворению металла, так как положительно заряженные ионы металла в растворе вместе с отрицательно заряженным металлом создают электрическое поле, направленное внутрь металла и поэтому препятствующее дальнейшему переходу ионов в раствор. Можно сказать, что на поверхности металла создается электрический диполь, а точнее говоря, двойной электрический слой (рис. 10.2), поле которого препятствует уходу ионов из металла. В конечном счете возникает равновесное состояние, характеризующееся некоторой равновесной плотностью ионов в растворе, которой соответствует вполне определенная разность потенциалов между металлом и водным раствором его ионов (или металлом и раствором электролита, в который погружен металл); эта разность потенциалов (ее называют *электродным потенциалом*) аналогична контактной разности потенциалов, возникающей на границе соприкосновения двух металлов. Таким же образом объясняется возникновение разности потенциалов на границе двух соприкасающихся растворов электролитов. Разные металлы обладают, вообще говоря, отличающимися электродными потенциалами. Поэтому если поместить в какой-либо раствор электролита два разных металла, то их потенциалы будут отличаться и между ними возникнет электрическое поле.

В гл. 9 мы видели, что в замкнутой цепи из разных металлов, имеющих одну и ту же температуру, сумма всех контактных разностей потенциалов равна нулю. Если же образовать замкнутую цепь из металлов и растворов электролитов, то сумма всех контактных разностей потенциалов не равна нулю даже при одной и той же температуре всех элементов цепи. В самом деле, сумма всех контактных разностей потенциалов, т. е. э. д. с. в цепи, представляет собой работу сил электрического поля (которое возникает между соприкасающимися проводниками) при перемещении единичного заряда вдоль замкнутой цепи. Если электрон перемещается вдоль замкнутой цепи, состоящей из одних только металлов, то при их одинаковой температуре в цепи не происходит никаких внутренних изменений, т. е. энергия ее не изменяется и поэтому общая работа сил электрического поля равна нулю. В отличие от такой цепи в цепи, содержащей и металлы, и растворы электролитов, при перемещении заряда происходит, как мы видели, при изучении электролиза, химическая реакция. При этом изменяется энергия системы и, следовательно, работа сил поля может быть отлична от нуля.

Существование отличной от нуля э. д. с. означает, что

в замкнутых цепях, содержащих металлические и электролитические проводники, могут самопроизвольно, т. е. без добавочных внешних устройств, протекать электрические токи.

Такие цепи носят общее название *гальванических элементов*. Простейшим примером гальванического элемента является элемент



10.2.
Двойной электрический слой на поверхности металла, погруженного в раствор электролита

Вольты, состоящий из цинкового и медного электродов, погруженных в разбавленную серную кислоту (что записывается в виде $Zn|H_2SO_4|Cu+$). При замыкании электродов в цепи течет электрический ток от медного электрода (анод) к цинковому электроду (катод).

ОСНОВНЫЕ ФОРМУЛЫ

Закон электролиза $M_A = \frac{A}{z} \frac{It}{F}$,

$$F = 96484 \text{ Кл/моль}$$

Удельная электрическая проводимость раствора $\sigma = \sum_a b_a (ez_a)^2 n_a$

Удельная электрическая проводимость слабого электролита $\sigma = e^2 \alpha c N_0 (b_+ + b_-)$,

$$\alpha = \frac{2K}{K + \sqrt{K^2 + 4Kc}}$$

Потенциал, создаваемый ионом в растворе (или плазме) $\varphi = \frac{ez_a}{4\pi\epsilon_0 r} e^{-r/r_D}$

Радиус экранировки (дебаевский радиус) $r_D = \sqrt{\frac{\epsilon_0 kT}{\sum_a e^2 z_a^2 n_{a0}}}$



Глава 11. ГАЗОВЫЙ РАЗРЯД

11.1. Анодный ток

Чистые газы являются идеальными изоляторами. Если приложить к ним постоянное электрическое поле, то в них не возникнет электрического тока, так как нейтральные молекулы, из которых состоит газ, будут лишь несколько деформироваться или поворачиваться, но не будут двигаться поступательно.

Ситуация существенно изменяется при введении в газ посторонних заряженных частиц: они начинают двигаться под действием приложенного к газу электрического поля, и в газе возникает электрический ток. Если напряжение поля достаточно велико, то внешние заряженные частицы, введенные в газ, могут приобрести в поле столь большую энергию, что ее оказывается достаточно для ионизации нейтральных газовых молекул, т. е. для отрыва от них одного или нескольких электронов. В результате наряду с посторонними носителями заряда в газе возникают собственные носители заряда — электроны и положительные ионы. Они, в свою очередь, при движении в поле мо-

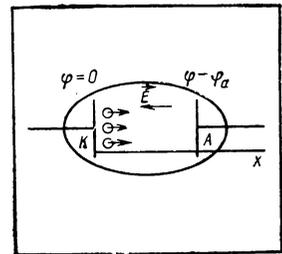
гут ионизовать молекулы, т. е. создавать вторичные электроны и ионы; вторичные частицы могут создавать третичные частицы и т. д. Таким образом, в принципе количество собственных носителей заряда может нарастать лавинообразно, а вместе с этим лавинообразно нарастает и ток в газе.

Ток в газе носит общее название *газового разряда*. Если он связан только с внешними носителями заряда, то разряд называют *несамостоятельным*. Если же вследствие лавинообразного нарастания числа носителей заряда разряд может протекать при сколь угодно малом затравочном внешнем заряде, то его называют *самостоятельным*.

Начнем с определения силы тока при движении отдельного заряда через газовый или вакуумный промежуток электрической цепи. На первый взгляд кажется, что мгновенный ток в цепи отсутствует во все время движения частицы между электродами и только в тот момент, когда частица достигнет электрода и отдаст ему свой заряд, возникает импульс тока. В действительности это не так: вследствие электростатической индукции на электродах наводятся заряды, которые при движении частицы движутся во внешней части цепи.

Найдем силу тока I . Для этого приравняем работу $\mathcal{E}I$ внешнего источника с э. д. с., равной \mathcal{E} , совершаемую за 1 с, изменению за то же время энергии движущегося заряда $e\mathbf{E}\mathbf{v}$; получим $I = e\mathbf{E}\mathbf{v}/\mathcal{E}$ (e — заряд и \mathbf{v} — скорость частицы). В частности, для плоскопараллельных электродов $E = \mathcal{E}/d$ (d — расстояние между ними), поэтому если заряд движется перпендикулярно электродам, то $I = ev/d$. В самом общем случае для отдельного заряда, движущегося в вакууме между электродами, возникает задача о движении частицы в заданном внешнем поле $\mathbf{E} = -\text{grad}\varphi$, которое определяется лишь геометрией электродов и приложенным к ним напряжением. При этом в силу потенциальности поля сохраняется полная энергия $W = \frac{mv^2}{2} + e\varphi = \text{const}$, поэтому в точке с координатами \mathbf{r} скорость частицы $v(\mathbf{r}) = \sqrt{2(W - \varphi(\mathbf{r}))/m}$ (m — масса частицы).

Если в вакууме движется не одна, а много заряженных частиц, то поле уже нельзя считать заданным: на него существенное влияние оказывает создаваемый ими объемный заряд. Рассмотрим, например, вакуумный диод (рис. 11.1), т. е. два плоских параллельных электрода, между которыми поддерживается определенная разность потенциалов φ_a , и один из электродов — катод K — непрерывно испускает электроны (для этого катод нагревают до высокой температуры). Под действием электрического поля, направленного от анода A к катоду K , электроны движутся к аноду. Возникает вопрос: какое влияние оказывает объемный заряд электронов на поле между электродами, а следовательно, и на само движение электронов? Ясно, что присутствие пространственного отрицательного заряда между электродами приводит к тому, что



11.1. Вакуумный диод

часть силовых линий, исходящих из анода A , заканчивается на отрицательных зарядах в промежутке между K и A и не доходит до катода K . Поэтому поле у катода K меньше, чем у анода A . Чем больше электронов вылетает из катода, тем больше их находится в пути к аноду, тем больше силовых линий, исходящих из анода, не доходит до катода и, следовательно, тем меньше поле у катода. В конечном счете поле у катода может упасть до нуля, и в этом случае ток электронов на анод достигнет максимально возможного значения (при заданном потенциале φ_a анода). Если число электронов, эмиттируемых из катода за 1 с, станет еще большим, то поле у катода вместо ускоряющего станет тормозящим и ток будет меньшим, чем максимально возможный предельный ток, соответствующий нулевому полю у катода.

Все величины, описывающие поле, зависят, очевидно, только от одной координаты x (ось x направлена по нормали к электродам). Поэтому уравнение Пуассона примет вид

$$\frac{d^2\varphi}{dx^2} = en/\epsilon_0$$

(e — заряд электрона, n — плотность электронов). Катод эмиттирует электроны равномерно во времени, поэтому плотность электронов не зависит от времени; рассматривая установившееся движение, будем считать и все другие величины не зависящими от времени. Тогда в силу закона сохранения заряда плотность тока j является вообще константой, т. е. не зависит от x :

$$j = env = \text{const.}$$

Воспользуемся теперь законом сохранения энергии $mv^2/2 = e\varphi$ (мы полагаем, что на катоде $v = 0$, и выбираем за $\varphi = 0$ потенциал катода). Получим $v = \sqrt{2e\varphi/m}$ и, следовательно,

$$en = \frac{j}{v} = j \sqrt{\frac{m}{2e}} \varphi^{-1/2}.$$

Наконец, подставляя это выражение в уравнение Пуассона, имеем

$$\epsilon_0 \frac{d^2\varphi}{dx^2} = j \sqrt{\frac{m}{2e}} \varphi^{-1/2}.$$

Это уравнение позволяет определить потенциал как функцию координаты x . Умножая его для этого на $d\varphi/dx$, приведем полученное уравнение к виду

$$\epsilon_0 \left(\frac{d\varphi}{dx} \right)^2 = 4j \sqrt{\frac{m}{2e}} \varphi^{1/2} + \epsilon_0 E_0^2,$$

где E_0^2 — значение $(d\varphi/dx)^2$ в точке, где $\varphi = 0$, т. е. квадрат напряженности поля у катода.

Предположим сперва, что $E_0 = 0$. Тогда $d\varphi/dx = \alpha j^{1/2} \varphi^{1/4}$, где $\alpha = (8m/e)^{1/4} \epsilon_0^{-1/2}$, откуда

$$\varphi(x) = \left(\frac{3\alpha}{4}\right)^{4/3} j^{2/3} x^{4/3}.$$

Мы видим, что из-за наличия объемного заряда при установившемся движении электронов потенциал между катодом и анодом изменяется с расстоянием пропорционально $x^{4/3}$, а не x , как было бы в случае вакуума. Соответственно напряженность поля изменяется как $x^{1/3}$, а не остается постоянным, как в случае вакуума:

$$E = \left| \frac{d\varphi}{dx} \right| = \frac{4}{3} \left(\frac{3\alpha}{4}\right)^{4/3} j^{2/3} x^{1/3}.$$

Таким образом, поле неоднородно и растет по направлению от катода к аноду.

Приведем выражения для скорости электронов и их объемного заряда как функции расстояния до катода:

$$v(x) = \sqrt{\frac{2e}{m_e} \varphi^{1/2}} = \sqrt{\frac{2e}{m_e} \left(\frac{3\alpha}{4}\right)^{2/3} j^{1/3} x^{2/3}};$$

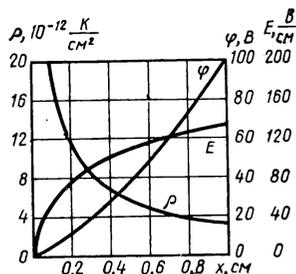
$$en(x) = \frac{j}{v} = \sqrt{\frac{m_e}{2e} \left(\frac{4}{3\alpha}\right)^{2/3} j^{2/3} x^{-2/3}}.$$

Обратим внимание на то, что объемная плотность заряда обращается в бесконечность на катоде. В действительности, конечно, это не так — плотность заряда конечна во всех точках, и полученный результат связан просто с грубостью сделанного нами предположения о равенстве нулю скорости вылетающих из раскаленного катода электронов. Если учесть, что эта скорость в действительности отлична от нуля, то мы придем к конечному значению для плотности электронов. Ход изменения потенциала φ , напряженности поля E и плотности заряда ρ для плоского диода с анодным потенциалом $\varphi_a = 100$ В и расстоянием между электродами $d = 1$ см показан на рис. 11.2.

Возвратимся к формуле, определяющей φ как функцию x , и положим в ней $x = d$. Мы найдем тогда потенциал анода φ_a как функцию плотности тока j и, следовательно, полного анодного тока $I_a = jS$ (S — площадь анода). Решая полученное соотношение относительно j , придем к так называемому «закону трех вторых»:

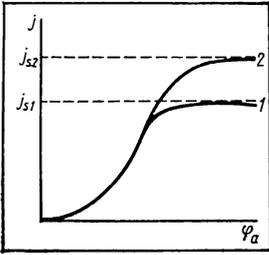
$$\frac{I_a}{S} = j = \frac{4\epsilon_0}{9} \sqrt{\frac{2e}{m_e} \frac{\varphi_a^{3/2}}{d^2}}.$$

Таким образом, сила тока на аноде пропорциональна анодному потенциалу в степени $3/2$, а не в первой степени, как это имеет место в обычных проводниках согласно закону Ома. Неподчинение закону Ома связано в этом случае с присутствием объемного заряда (который



11.2.

Распределение потенциала напряженности поля и плотности заряда в плоском диоде с расстоянием 1 см между электродами при анодном напряжении 100 В



11.3. Зависимость тока от анодного напряжения с учетом насыщения

быть больше плотности тока термоэлектронной эмиссии j_s . Этим определяется верхняя граница анодного потенциала, для которого еще справедлив «закон трех вторых». На рис. 11.3 изображены зависимости $j = j(\varphi_a)$ с учетом ограничения $j \leq j_s$ (ток j_s в этой связи называют *током насыщения*). Кривая 2 с большим током насыщения, чем в случае кривой 1, соответствует более высокой температуре катода (такого рода кривые называются вольт-амперными характеристиками).

полностью компенсируется зарядами противоположного знака как в металлах и полупроводниках, так и в электролитах, см. § 9.9).

Полученная формула для плотности тока соответствует тому случаю, когда напряженность поля на катоде $E_0 = 0$, но можно вычислить ток и в том случае, когда $E_0 \neq 0$. Оказывается, что сила тока в этом случае меньше силы тока при нулевом поле на катоде. Поэтому найденный ток является предельным при заданном анодном потенциале. Ясно, что, каков бы ни был анодный потенциал, плотность анодного тока не может

11.2. Температура заряженных частиц

В предыдущем параграфе мы полагали, что заряженные частицы движутся в вакууме. Рассмотрим теперь, как движутся частицы под действием электрического поля через находящийся в тепловом равновесии газ нейтральных молекул. Если бы поля не было, то заряженные частицы, сталкиваясь с нейтральными молекулами, обменивались бы с ними энергией и в результате установилось бы состояние равновесия, при котором стали бы одинаковыми средние значения кинетических энергий поступательного движения нейтральных молекул и заряженных частиц. Именно: независимо от массы и сорта молекул и заряженных частиц средняя кинетическая энергия стала бы равной $\frac{1}{2}m \langle v^2 \rangle = \frac{3}{2}kT$, где m и v — масса и скорость частицы или молекулы и T — общая температура газа нейтральных молекул и заряженных частиц. Иными словами, в отсутствие внешнего поля заряд частицы не сказывался бы.

Иная картина возникает, если имеется внешнее электрическое поле. В этом случае заряженные частицы непрерывно получают от поля энергию. Тем не менее энергия частицы возрастает не бесконечно. Действительно, частицы сталкиваются с молекулами и передают им часть своей энергии. В результате устанавливается некоторое стационарное состояние, но средние кинетические энергии частиц и молекул уже не одинаковы. Иными словами, различаются температуры газа нейтральных молекул и газа заряженных частиц, и это различие тем больше, чем сильнее электрическое поле.

Покажем, как можно оценить температуру T_e заряженных частиц (для определенности будем рассматривать электроны). Средняя тепловая скорость электрона всегда больше средней тепловой скорости молекул газа (при заданной температуре), поэтому молекулы можно считать покоящимися. В результате столкновения с электроном молекула приобретает некоторую энергию ϵ_m , пропорциональную кинетической энергии электрона ($\epsilon_m = \xi \epsilon$, где ξ — доля переданной электроном энергии). Эта величина существенно зависит от характера столкновения. Если столкновение электрона с молекулой является упругим, то, как непосредственно следует из законов сохранения энергии и импульса, максимальное значение $\xi_{\max} = 4(m_e/M)$, где m_e и M — масса электрона и молекулы ($m_e \ll M$). Минимальное значение ξ равно, очевидно, нулю, так что среднее значение доли переданной энергии ξ при упругом столкновении электрона с молекулой $\langle \xi \rangle = 2(m_e/M)$. Величина эта очень мала и составляет $2,8 \cdot 10^{-4}$ для He и $5,5 \cdot 10^{-6}$ для Hg. Однако малая потеря энергии характерна только для сравнительно медленных электронов, движущихся в одноатомных газах. С ростом энергии электронов и особенно при движении электронов в многоатомных газах доля теряемой электроном энергии значительно возрастает, так как столкновения становятся существенно неупругими, т. е. сопровождаются перестройкой молекулы и изменением ее внутренней энергии.

Разделив $\langle \xi \rangle \langle \epsilon \rangle$, где $\langle \epsilon \rangle$ — среднее значение кинетической энергии электрона, на среднее время τ свободного пробега электрона, найдем среднюю энергию $\delta \epsilon_-$, теряемую электроном за 1 с: $\delta \epsilon_- = \langle \xi \rangle \langle \epsilon \rangle / \tau$. В состоянии равновесия она должна быть равна средней энергии, получаемой электроном от электрического поля. Средняя энергия, которую электрон получает от поля за 1 с, $\delta \epsilon_+ = eEu$, где $u = beE$ — направленная скорость электрона в поле и подвижность электрона $b = \tau/m$. Написав $\delta \epsilon_+ = \delta \epsilon_-$ и вводя длину свободного пробега $l = \tau \langle v \rangle \sim \tau \sqrt{\langle \epsilon \rangle / m_e}$, получим

$$\langle \epsilon \rangle \equiv \frac{3}{2} kT_e \sim eEl \langle \xi \rangle^{-1/2}.$$

Длина свободного пробега не зависит от энергии электрона и равна $l \sim 1/(nQ)$, где n — плотность молекул и Q — сечение рассеяния электрона молекулой; поэтому температура электронов T_e пропорциональна напряженности электрического поля. (Этот вывод справедлив при не слишком слабых полях, когда $T_e \gg T$.) Для подвижности электронов получим $b \sim [l/(m_e eE)]^{1/2} \langle \xi \rangle^{1/4}$.

11.3. Самостоятельный разряд

Быстрый электрон, столкнувшись с атомом или молекулой, может выбить из них один или несколько электронов. В результате наряду с исходным электроном возникает несколько новых заряженных частиц — электронов и положительно заряженных ионов. Явление это носит общее название *ионизации*. Не следует думать, однако, что каждое столкновение электрона с атомом должно

обязательно приводить к ионизации атома. Необходимое (но не достаточное!) условие заключается в том, чтобы энергия электрона превосходила некоторую минимальную энергию — энергию связи электрона в атоме (или молекуле). Эту минимальную энергию называют *энергией ионизации*. Так, для получения иона аргона Ag^+ энергия налетающего электрона должна превосходить 15 эВ, а для получения Ag^{++} она должна быть более 45,0 эВ. Но даже и в том случае, когда энергия электрона превосходит энергию ионизации атома, столкновение электрона с атомом не обязательно приводит к ионизации атома. Дело в том, что наиболее часты упругие столкновения, при которых не меняется внутреннее состояние атома, и только в сравнительно небольшом числе случаев при столкновении происходит либо внутренняя перестройка атома, либо выбивание из атома электронов — одного или нескольких. Перестройка структуры атома обычно связана с возрастанием внутренней энергии атома, поэтому ее называют *возбуждением атома*. Перейдя при столкновении с электроном в возбужденное состояние, атом, однако, долго в нем не находится и по прошествии очень небольшого времени (порядка 10^{-8} — 10^{-10} с) переходит в основное состояние. Этот переход сопровождается излучением света («высвечивание» атома).

С ростом напряженности электрического поля возрастает энергия, приобретаемая заряженными частицами, движущимися в газе. В конечном счете она может достичь энергии ионизации и превзойти ее, и в этом случае сами заряженные частицы, а не только внешний ионизатор, будут ионизовать молекулы газа. Мы перейдем теперь к изучению этой собственной ионизации газа.

Интенсивность ионизации газа в электрическом поле можно характеризовать числом пар ионов противоположного знака заряда, образуемых заряженными частицами (электронами и положительными ионами) на единице пути их вдоль поля. Эти числа называют *коэффициентами ионизации* и обозначают обычно α в случае электронов и β в случае ионов. Покажем, как связать эти коэффициенты с напряженностью поля. Для определенности рассмотрим электроны. Пусть l — средняя длина свободного пробега электрона в газе. Тогда вероятность того, что электрон пройдет без столкновений расстояние, лежащее в интервале $(x, x + dx)$, есть

$$\omega(x) dx = e^{-kx/l} \frac{dx}{l}.$$

Пройдя расстояние x вдоль поля, электрон приобретает энергию eEx , и если она будет равна или больше энергии ионизации ε_i , т. е. $eEx \geq \varepsilon_i$, то в принципе электрон сможет ионизовать молекулу газа. Таким образом, расстояние, проходимое электроном вдоль поля без столкновений, должно быть не меньшим, чем $\varepsilon_i/(eE)$. Соответствующая вероятность составит

$$\omega dx = e^{-\varepsilon_i/(eEl)} \frac{dx}{l}.$$

Умножив эту величину на вероятность w_i ионизации газа электроном (данной энергии), можно найти число пар ионов, рождаемых электроном на участке пути dx . Таким образом, мы получим

$$\alpha = \frac{w_i}{l} e^{-\varepsilon_i/(eEl)}.$$

Длина пробега l обратно пропорциональна давлению p . Поэтому формулу для α можно переписать в виде

$$\alpha/p = Ae^{-pB/E},$$

где A и B — некоторые характерные для газа константы.

Пройдя путь dx , электрон, по определению рождает αdx пар заряженных частиц, т. е. αdx положительных ионов и столько же электронов. Эти электроны, двигаясь в поле, в свою очередь рожают новые электроны, так что число электронов в газе лавинообразно возрастает. Если число электронов на расстоянии x от электрода (катода) равно $N_e(x)$, то увеличение их числа на участке dx составляет $dN_e(x) = N_e(x)\alpha dx$, откуда

$$N_e(x) = N_0 e^{\alpha x},$$

где N_0 — число электронов на катоде (при $x = 0$). Эта формула и определяет закон лавинообразного нарастания электронов в газе, обусловленного последовательными процессами ионизации.

Если на расстоянии d от катода расположен анод, то $N_0 e^{\alpha d}$ — число электронов на аноде. Отсюда следует, что если из катода выходит электронный ток I_0 (возникающий, например, при освещении или нагревании катода), то электронный ток на аноде

$$I_a = I_0 e^{\alpha d}.$$

Итак, каждый освобожденный электрон на катоде рождает всего в газовом промежутке толщины d между катодом и анодом $e^{\alpha d}$ электронов; число же рожденных положительных ионов равно, очевидно, $e^{\alpha d} - 1$. Эти ионы движутся в обратном направлении по сравнению с электронами, т. е. к катоду. По пути они ионизируют молекулы газа и, кроме того, придя на катод, выбивают из него электроны.

Выясним, как влияют на лавину электронов электроны, освобождающиеся из катода при ударах ионов. Пусть ион, попавший на катод, выбивает из катода в среднем η электронов. Каждый электрон, дойдя до анода, создает $e^{\alpha d}$ электронов и $e^{\alpha d} - 1$ ионов. Последние, достигая катода, создают $\eta(e^{\alpha d} - 1)$ электронов. Эти электроны, в свою очередь, создают $\eta(e^{\alpha d} - 1)e^{\alpha d}$ электронов и т. д. Таким образом, из одного начального электрона на катоде всего возникнет на аноде

$$z = e^{\alpha d} + (e^{\alpha d} - 1)\eta e^{\alpha d} + (e^{\alpha d} - 1)^2 \eta^2 e^{\alpha d} + \dots$$

электронов. Если $\eta(e^{\alpha d} - 1) < 1$, то этот ряд сходится и дает $z = e^{\alpha d} [1 - \eta(e^{\alpha d} - 1)]^{-1}$. Таким образом, если первичный ток из катода равен I_0 , то анодный ток

$$I_a = I_0 \frac{e^{\alpha d}}{1 - \eta(e^{\alpha d} - 1)}.$$

Эта формула позволяет понять, как несамостоятельный разряд переходит в самостоятельный, т. е. в разряд, существующий в отсутствие внешних источников ионизации. Ясно, что для того, чтобы получить конечное значение тока I_a при бесконечно малом первичном токе I_0 , нужно приравнять нулю знаменатель выражения для I_a , т. е. положить

$$1 - \eta(e^{\alpha d} - 1) = 0.$$

Это соотношение и представляет собой *условие возникновения самостоятельного разряда*, когда сам разряд производит необходимые для его поддержания заряженные частицы. Частицы эти возникают в результате ионизации электронами молекул газа и выбивания ионами электронов из катода. Переписав написанное условие в виде $\alpha d = \ln(1 + 1/\eta)$ и подставляя сюда найденное выше выражения для α , получим

$$A p d e^{-pB/E} = \ln(1 + 1/\eta).$$

Отсюда может быть найдено критическое значение поля, при котором начинается самостоятельный разряд. Замечая, что $E = \varphi_a/d$, где φ_a — потенциал анода по отношению к катоду, получим для критического анодного потенциала, так называемого *потенциала зажигания*, выражение

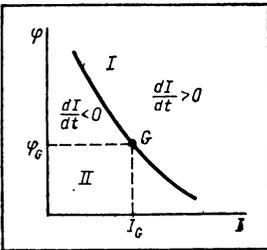
$$\varphi_a^{(c)} = B \frac{pd}{\ln(pd) + \delta}, \quad \delta = \ln \frac{A}{\ln(1 + 1/\eta)}.$$

Замечательно, что потенциал зажигания имеет минимум при некотором значении произведения давления на расстояние между электродами (он достигается при $pd \approx 2,7 \ln(1 + 1/\eta)/A$).

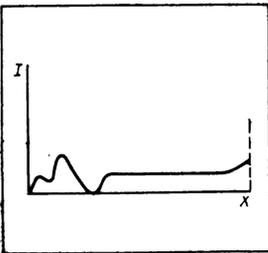
Разумеется, при самостоятельном разряде ток не становится бесконечно большим: если учитывать различные потери заряженных частиц, он всегда конечен. Более того, самостоятельный разряд может быть стационарным, для которого характерна определенная связь между силой тока и приложенной к газовому промежутку разностью потенциалов. Эта связь носит название *вольт-амперной характеристики разряда*. Типичная вольт-амперная характеристика разряда изображена на рис. 11.4. Если разряду соответствует некоторая точка G характеристики, то это значит, что процессы возникновения новых частиц и уничтожения частиц компенсируют друг друга, так что ток I_a , соответствующий напряжению φ_a , не меняется с течением времени.

Иными словами, плоскость (φ, I) , которую можно назвать плоскостью режимов разряда, разбивается вольт-амперной характеристикой на две части I и II , которым соответствуют неравновесные режимы, причем области I соответствует нарастание разрядного тока, а области II — спадание тока. Вольт-амперная характеристика, разграничивающая обе эти области, соответствует равновесным режимам разряда.

Характер газового разряда существенно зависит от силы тока. При очень малых токах ($I \leq 10^{-4}$ А) разряд происходит практически без свечения. Поэтому он называется *темным*. При больших токах постепенно возникают световые явления и такой разряд называют *тлеющим*. Основные физические процессы здесь те же, что и в темном разряде, т. е. объемная ионизация электронами и выбивание электронов ионами из катода. Типичным для разряда является наличие светящейся области на некотором расстоянии от катода (*катодная пленка*, или *катодный слой*). Перед катодной пленкой расположена темная область (*астоновое темное пространство*) и темное же пространство (*катодное темное пространство*) расположено сразу за пленкой.



11.4. Вольт-амперная характеристика газового разряда



11.5. Интенсивность свечения как функция расстояния от катода

Такая картина может быть объяснена просто. Действительно, испускаемые катодом (вследствие бомбардировки катода ионами) электроны имеют при выходе из катода очень малую скорость, так что они не могут сразу ни ионизовать, ни возбуждать атомы газа. Поэтому в аэрированном темном пространстве нет свечения. Но в этой области электроны, двигаясь в поле, набирают достаточно энергии, так что в катодной пленке становится возможным интенсивное возбуждение атомов, сопровождающееся излучением света. На катодную пленку приходится максимум вероятности возбуждения, поэтому за катодной пленкой, где энергия электронов еще больше, свечение становится значительно меньше, чем в катодной пленке. (Картина изменения интенсивности свечения вдоль тлеющего разряда изображена на рис. 11.5.)

Газовый разряд находит колоссальное количество чисто физических и технических применений. Сюда относятся ионизационная камера, газонаполненные фотоэлементы, счетчики с острием, служащие для измерения излучений; различного рода выпрямители (ртутные и железные) для преобразования тока; многочисленные источники света, использующие газовый разряд; различные сварочные дуги; наконец, разряды используют в качестве предохранителей и выключателей.

ОСНОВНЫЕ ФОРМУЛЫ

Плотность анодного тока («закон трех втормых»)

$$j = \frac{4\epsilon_0}{9} \sqrt{\frac{2e}{m_e}} \frac{\varphi_a^{3/4}}{d^2}$$

Температура электронов в газовом разряде

$$T_e \sim \frac{2eEl}{3k \sqrt{\langle \xi \rangle}}$$

Анодный ток, вырраженный через ионизационные коэффициенты

$$I_a = I_0 \frac{e^{\alpha d}}{1 - \eta (e^{\alpha d} - 1)}$$



Глава 12. МАГНЕТИКИ

12.1. Магнетик во внешнем поле

Намагниченность \mathbf{J} (см. гл. 6) зависит от магнитного поля в веществе, т. е. от магнитной индукции \mathbf{B} . Поэтому магнитная индукция и напряженность магнитного поля $\mathbf{H} = \mu_0^{-1} \mathbf{B} - \mathbf{J}$ не являются независимыми величинами и связаны между собой вполне определенным соотношением $\mathbf{B} = \mathbf{B}(\mathbf{H})$, характерным для каждого вещества. Для большинства веществ магнитная индукция является линейной функцией напряженности поля вплоть до ее очень больших значений, и только для так называемых ферромагнетиков

линейная зависимость наблюдается при малой напряженности поля и очень быстро сменяется сложной нелинейной зависимостью. Если рассматривать газы, жидкости и поликристаллы различных веществ, исключая ферромагнетики, то для них магнитная индукция \mathbf{B} пропорциональна напряженности магнитного поля \mathbf{H} , т. е.

$$\mathbf{B} = \mu_0 \mu \mathbf{H},$$

где μ — некоторая константа, характерная для данного вещества. Ее называют *относительной магнитной проницаемостью* или *магнитной проницаемостью*. Для монокристаллов векторы \mathbf{B} и \mathbf{H} не параллельны, хотя связь между ними по-прежнему является линейной. В этом случае декартовы составляющие B_i, H_i ($i = x; y; z$) векторов \mathbf{B} и \mathbf{H} связаны между собой линейными соотношениями

$$B_i = \mu_0 \sum_j \mu_{ij} H_j,$$

в которые входят девять коэффициентов пропорциональности μ_{ij} . Их называют *тензором магнитной проницаемости*. Тензор этот симметричен, т. е. $\mu_{ij} = \mu_{ji}$. Поэтому число независимых компонент этого тензора равно шести, но в зависимости от симметрии кристалла число различных компонент тензора магнитной проницаемости может быть и меньше шести.

Полагая $\mathbf{B} = \mu_0 \mu \mathbf{H}$ в соотношении $\mu_0^{-1} \mathbf{B} = \mathbf{H} + \mathbf{J}$, можно выразить намагниченность \mathbf{J} через напряженность поля \mathbf{H} :

$$\mathbf{J} = \chi \mathbf{H}, \quad \chi = (\mu - 1).$$

Величину χ называют *магнитной восприимчивостью* вещества.

Если проводить аналогию между магнитными величинами \mathbf{B} и \mathbf{H} , с одной стороны, и электрическими величинами \mathbf{D} и \mathbf{E} — с другой:

$$\mathbf{B} \rightarrow \mathbf{D}, \quad \mathbf{H} \rightarrow \mathbf{E},$$

то магнитная проницаемость μ аналогична электрической постоянной ϵ_0 , или диэлектрической проницаемости ϵ , а магнитная восприимчивость χ — диэлектрической восприимчивости α . Аналогия эта проявляется в граничных условиях на поверхности раздела двух сред:

$$B_{1n} = B_{2n}, \quad D_{1n} = D_{2n}; \quad H_{1t} = H_{2t}, \quad E_{1t} = E_{2t}.$$

Несмотря на аналогию между магнитной проницаемостью и электрической постоянной, между ними есть принципиальное различие. Дело в том, что диэлектрическая восприимчивость всегда положительна, поэтому $\epsilon > 1$, в то время как магнитная восприимчивость χ может быть как положительной, так и отрицательной, а следовательно, магнитная проницаемость может быть и больше и меньше единицы. В первом случае вещество называется *парамагнитным*, а во втором — *диамагнитным*.

Чтобы экспериментально убедиться, в существовании двух различных типов веществ — парамагнетиков и диамагнетиков, — выясним, как происходит взаимодействие между двумя постоянными

токама, находящимися не в вакууме, а в среде с определенной магнитной проницаемостью μ . Напряженность магнитного поля \mathbf{H} в веществе в статическом случае определяется, как мы знаем, уравнением $\text{rot} \mathbf{H} = \mathbf{j}$, в которое не входит магнитная проницаемость. Поэтому напряженность магнитного поля, порождаемого током в веществе, определяется тем же законом Био—Савара

$$\mathbf{H} = \frac{I}{4\pi} \oint \frac{[d\mathbf{r}]}{r^3},$$

что и в вакууме (здесь $d\mathbf{l}$ — элемент длины проводника, по которому течет ток I , и \mathbf{r} — радиус-вектор, проведенный из элемента $d\mathbf{l}$ в точку, где ищется поле \mathbf{H}). Что же касается вектора магнитной индукции \mathbf{B} , то он в $\mu_0\mu$ раз больше \mathbf{H} , т. е.

$$\mathbf{B} = \frac{\mu_0\mu I}{4\pi} \oint \frac{[d\mathbf{r}]}{r^3}.$$

Пусть теперь в рассматриваемой среде находится второй проводник, по которому течет ток I' . Какую силу испытывает этот проводник со стороны проводника с током I ? Если рассматривать отдельный заряд q , движущийся в среде со скоростью \mathbf{v} , то на заряд действует сила Лоренца $\mathbf{f} = q[\mathbf{v}\mathbf{B}]$. Эта формула ничем не отличается от формулы для силы, действующей на заряд со стороны магнитного поля в вакууме, — в обе формулы входит вектор \mathbf{B} . То, что в формулу для силы в случае движения частицы в среде должен входить именно вектор магнитной индукции \mathbf{B} , а не вектор напряженности поля \mathbf{H} , следует из самого смысла магнитной индукции — эта величина представляет собой усредненную напряженность микроскопического магнитного поля, действующего в среде (умноженную на μ_0).

Из написанного выражения для силы Лоренца, действующей на одиночный заряд, вытекает, что на элемент проводника $d\mathbf{l}'$ с током I' , находящимся в среде, действует сила

$$d\mathbf{F}' = I' [d\mathbf{l}'\mathbf{B}].$$

Подставляя сюда вместо \mathbf{B} магнитную индукцию, создаваемую током I , получим

$$d\mathbf{F}' = \mu_0\mu \frac{II'}{4\pi} \left[d\mathbf{l}' \oint \frac{[d\mathbf{r}]}{r^3} \right],$$

где \mathbf{r} — радиус-вектор, проведенный из $d\mathbf{l}$ в $d\mathbf{l}'$, и интегрирование совершается по контуру с током I . Эта формула показывает, что сила взаимодействия между токами в среде отличается от случая вакуума только множителем μ . Поэтому если среда парамагнитна, то сила между токами больше, чем в случае вакуума; если же среда диамагнитна, то сила меньше, чем в случае вакуума.

Как мы только что видели, магнитная индукция, порождаемая некоторым током в среде, в μ раз отличается от индукции для случая, когда проводник с этим же током находится в вакууме. В частности, если взять прямой длинный провод, по которому течет ток I , и помес-

тить провод в среду с магнитной проницаемостью μ , то напряженность поля H и индукция B вне проводника на расстоянии r от его оси определяются формулами

$$H = \frac{I}{2\pi r}, \quad B = \frac{\mu_0 \mu I}{2\pi r} \quad (r \geq a),$$

а внутри проводника — формулами

$$H = \frac{I}{2\pi a} \frac{r}{a}, \quad B = \frac{I \mu' \mu_0 r}{2\pi a a} \quad (r \leq a),$$

где a — радиус сечения и μ' — магнитная проницаемость проводника.

Для напряженности поля и индукции в длинном и тонком соленоиде имеем

$$H = In, \quad B = \mu_0 \mu In,$$

где μ — магнитная проницаемость вещества сердечника и n — число витков, приходящееся на единичную длину соленоида. Поэтому магнитный поток через площадь сечения S соленоида есть

$$\Phi = BS = \mu_0 \mu InS.$$

Заменяя здесь n на N/l , где N — полное число витков и l — длина соленоида, можно представить Φ в виде $\Phi = \mathcal{E}_m / R_m$, где $\mathcal{E}_m = \mu_0 IN$, $R_m = l/(\mu S)$.

Магнитный поток можно сравнивать с током, а соленоид рассматривать как магнитопровод, через который проходит магнитный поток. Аналогия эта связана с тем, что плотность токов \mathbf{j} в отсутствие зарядов удовлетворяет уравнению $\operatorname{div} \mathbf{j} = 0$ и точно такому же уравнению удовлетворяет магнитная индукция \mathbf{B} , представляющая собой плотность магнитного потока: $\operatorname{div} \mathbf{B} = 0$. Поэтому написанное уравнение для Φ можно интерпретировать как закон Ома для магнитной цепи — величина \mathcal{E}_m играет при этом роль магнитодвижущей силы, а величина R_m — магнитного сопротивления цепи. Магнитодвижущая сила пропорциональна числу ампервитков IN , а магнитное сопротивление связано с длиной и площадью сечением магнитопровода точно так же, как обычное сопротивление связано с длиной и площадью сечения проводника, только вместо коэффициента электропроводности в R_m входит магнитная проницаемость. Чем она больше, тем больше магнитный поток. Поэтому на практике для получения больших потоков (при заданном числе ампервитков) используют материалы с большой μ , например в трансформаторах и динамомашинках используются различные сорта железа и сплавы типа пермаллоя (20% Fe, 80% Ni), для которого магнитная проницаемость достигает значений порядка 10 000.

Закон Ома для магнитной цепи позволяет находить магнитный поток не только в простейшем случае замкнутого соленоида, заполненного однородным магнетиком, но и для более сложных магнитных цепей. Если, например, сердечник соленоида состоит из двух соединенных последовательно частей с различными магнитными проницаемостями μ_1 и μ_2 , то магнитный поток определится прежней формулой,

но только под магнитным сопротивлением R_m следует понимать сумму магнитных сопротивлений $R_m^{(1)}$ и $R_m^{(2)}$ обоих участков соленоида:

$$R_m = R_m^{(1)} + R_m^{(2)}; \quad R_m^{(1)} = l_1/(\mu_1 s_1); \quad R_m^{(2)} = l_2/(\mu_2 s_2)$$

(l_1 и l_2 — длины, s_1 и s_2 — площади сечения обоих участков соленоида). Магнитная цепь может быть и разветвленной; при параллельном соединении складываются обратные сопротивления:

$$\frac{1}{R_m} = \frac{1}{R_m^{(1)}} + \frac{1}{R_m^{(2)}}.$$

На магнетик, помещенный в магнитное поле, со стороны поля действует некоторая сила. Чтобы понять это, вспомним, что рамка с током стремится повернуться перпендикулярно полю и переместиться в те места, где поле сильнее. Но в магнетике имеются молекулярные токи, взаимодействие которых с полем и приводит к силам, действующим на магнетик. Исходя из такой картины, легко определить силы, действующие на магнетик в магнитном поле. Воспользуемся выражением $U = -\mu_0 \mathbf{m} \mathbf{H}_0$ для энергии взаимодействия магнитного момента \mathbf{m} молекулы с внешним полем \mathbf{H}_0 . Мы приходим к выводу, что на \mathbf{m} действует момент сил

$$\mathbf{N}_m = \mu_0 [\mathbf{m} \mathbf{H}_0].$$

Если же \mathbf{H}_0 зависит от координат, то действует еще и сила

$$= -\text{grad } U = \mu_0 \{(\mathbf{m} \text{ grad}) \mathbf{H}_0 + [\mathbf{m} \text{ rot } \mathbf{H}_0]\}.$$

Последнее слагаемое в этом выражении, очевидно, отсутствует, если в объеме, занимаемом магнетиком, нет внешних токов. В этом случае

$$\mathbf{f} = \mu_0 (\mathbf{m} \text{ grad}) \mathbf{H}_0.$$

Просуммировав выражения для \mathbf{f} и \mathbf{N}_m по всем магнитным моментам \mathbf{m} , найдем результирующие силу \mathbf{F} и момент сил \mathbf{N} , действующие на магнетик. Для единичного объема магнетика это

$$\mathbf{F} = \mu_0 (\mathbf{J} \text{ grad}) \mathbf{H}_0, \quad \mathbf{N} = \mu_0 [\mathbf{J} \mathbf{H}_0],$$

где \mathbf{J} — намагниченность магнетика. Здесь \mathbf{H}_0 обозначает внешнее магнитное поле, но если магнитная проницаемость магнетика мало отличается от единицы, то \mathbf{H}_0 мало отличается от поля \mathbf{H} внутри магнетика. Учитывая, что $\mathbf{J} = \chi \mathbf{H}$, получим

$$\mu_0^{-1} \mathbf{F} = \chi (\mathbf{H} \text{ grad}) \mathbf{H} = \frac{\chi}{2} \text{grad } \mathbf{H}^2.$$

Эта формула оказывается справедливой и для парамагнетиков ($\chi > 0$), и для диамагнетиков ($\chi < 0$). Таким образом,

парамагнетики стремятся перейти в места, где напряженность поля больше, а диамагнетики — уйти из этих мест.

Если внести магнетик во внешнее магнитное поле, то поле станет иным, чем до внесения магнетика. Действительно, пусть в вакууме

до внесения магнетика имеется статическое поле $\mu_0 \mathbf{H}^{(e)} = \mathbf{B}^{(e)}$. Оно определяется уравнениями

$$\text{rot } \mathbf{H}^{(e)} = \mathbf{j}; \quad \text{div } \mathbf{H}^{(e)} = 0,$$

где \mathbf{j} — плотность тока проводимости, создающего поля. После внесения магнетика первое из этих уравнений остается в силе и для нового поля, которое мы обозначим \mathbf{H} , т. е. $\text{rot } \mathbf{H} = \mathbf{j}$. Второе же уравнение заменится уравнением $\text{div } \mathbf{B} = 0$, из которого следует, что $\text{div } \mathbf{H} = -\text{div } \mathbf{J}$. Сравнивая это уравнение с уравнением электростатики $\text{div } \mathbf{E} = \rho/\epsilon_0$, можно сказать, что у поля \mathbf{H} возникают источники с объемной плотностью заряда («магнитного заряда») $\rho_m = -\text{div } \mathbf{J}$. Кроме того, на границе магнетика с вакуумом возникает поверхностный заряд с плотностью $\sigma_m = J_n$, где J_n — нормальная составляющая \mathbf{J} на границе магнетика (нормаль подразумевается внешняя). По этой причине поле \mathbf{H} складывается из двух слагаемых: поля $\mathbf{H}^{(e)}$, существовавшего до внесения магнетика, и добавочного поля \mathbf{H}' , создаваемого источниками ρ_m и σ_m . Это поле может быть определено по формуле электростатики $\mathbf{H}' = -\text{grad } \varphi_m$, где φ_m — статический («магнитостатический») потенциал, порождаемый «магнитными» зарядами:

$$\varphi_m = \frac{1}{4\pi} \left\{ \int_V \frac{\rho_m}{r} dV + \int_\Sigma \frac{\sigma_m}{r} ds \right\}$$

(V — объем и Σ — поверхность магнетика; r обозначает расстояние от dV или ds до точки, где ищется φ_m).

Итак, \mathbf{H} отличается от $\mathbf{H}^{(e)}$. В частности, если поле до внесения магнетика было однородным, то после внесения магнетика оно перестанет быть однородным. Если, однако, поле $\mathbf{H}^{(e)}$ однородно, а внесенный в него магнетик имеет вполне определенную форму, а именно форму произвольного эллипсоида, то эллипсоид намагнитится однородно, т. е. вектор \mathbf{J} внутри эллипсоида имеет повсюду одно и то же значение и одно и то же направление. При этом $\rho_m = 0$ и поле \mathbf{H}' внутри эллипсоида (но не вне его!) однородно, а следовательно, однородно и суммарное поле $\mathbf{H}^{(i)}$ внутри эллипсоида, но не вне его. Поле внутри эллипсоида $\mathbf{H}^{(i)}$ связано с однородным полем $\mathbf{H}^{(e)}$, существовавшим до внесения эллипсоида, соотношением

$$\mathbf{H}^{(i)} = \mathbf{H}^{(e)} - \hat{N}\mathbf{J},$$

где \hat{N} — некоторый тензор, не зависящий от поля и определяемый только формой эллипсоида и расположением его относительно $\mathbf{H}^{(e)}$. Он носит название *тензора размагничивающих коэффициентов*. Напомним, что такой же формулой определяется связь между внешним однородным электростатическим полем $\mathbf{E}^{(e)}$ и полем $\mathbf{E}^{(i)}$ внутри диэлектрического эллипсоида, внесенного в поле $\mathbf{E}^{(e)}$:

$$\mathbf{E}^{(i)} = \mathbf{E}^{(e)} - \epsilon_0^{-1} \hat{N} \mathbf{P},$$

где \mathbf{P} — вектор поляризации внутри эллипсоида. Одинаковая структура этих формул связана с тем, что граничные условия для векторов \mathbf{D} и \mathbf{E} такие же, как и для векторов \mathbf{B} и \mathbf{H} .

Вектор \mathbf{J} , входящий в формулу для $\mathbf{H}^{(i)}$, равен, очевидно, $\mathbf{J} = \mu_0^{-1} \mathbf{B}^{(i)} - \mathbf{H}^{(i)}$, где $\mathbf{B}^{(i)}$ — магнитная индукция в эллипсоиде. Если $\mathbf{B}^{(i)} = \mu_0 \mu \mathbf{H}^{(i)}$, то $\mathbf{J} = \chi \mathbf{H}^{(i)}$, где $\chi = \mu - 1$, и формула для $\mathbf{H}^{(i)}$ приобретает вид

$$\mathbf{H}^{(i)} = \mathbf{H}^{(e)} - \chi \hat{N} \mathbf{H}^{(i)}.$$

В системе координат, оси которой совпадают с главными осями эллипсоида, тензор \hat{N} имеет только диагональные элементы N_1, N_2, N_3 , причем $N_1 + N_2 + N_3 = 1$. Если $\mathbf{H}^{(e)}$ направлено параллельно одной из главных осей эллипсоида, например оси k , то уравнение, связывающее $\mathbf{H}^{(i)}$ и $\mathbf{H}^{(e)}$, приобретает вид $\mathbf{H}^{(i)} = \mathbf{H}^{(e)} - \chi N_3 \mathbf{H}^{(i)}$, откуда

$$\mathbf{H}^{(i)} = \frac{\mathbf{H}^{(e)}}{1 + \chi N_3}.$$

Приведем значения N_1, N_2, N_3 для некоторых случаев. В случае шара

$$N_1 = N_2 = N_3 = 1/3.$$

Если магнетик имеет форму цилиндра, ось которого выбрана в качестве оси x_1 , то

$$N_1 = 0, \quad N_2 = N_3 = 1/2.$$

Для сплюснутого ($a > c$) эллипсоида вращения $(x_1^2 + x_2^2)/a^2 + x_3^2/c^2 = 1$ имеем

$$N_1 = N_2 = \frac{1}{2} (1 - N_3), \quad N_3 = \frac{1 + e^2}{e^3} (e - \arctg e), \quad e = \sqrt{\left(\frac{a}{c}\right)^2 - 1}.$$

Для вытянутого ($a > b$) эллипсоида вращения $x_1^2/a^2 + (x_2^2 + x_3^2)/b^2 = 1$, имеем

$$N_1 = \frac{1 - e^2}{2e^3} \left(\ln \frac{1 + e}{1 - e} - 2e \right), \quad N_2 = N_3 = \frac{1}{2} (1 - N_1),$$

$$e = \sqrt{1 - \left(\frac{b}{a}\right)^2}.$$

12.2. Магнитная восприимчивость парамагнитного газа

Пока мы формально разделили вещества на две группы: парамагнетики ($\mu > 1$) и диамагнетики ($\mu < 1$). Теперь мы покажем, что различие между ними связано с возможностью существования у атомов *собственного*, или *спонтанного*, *магнитного момента*, т. е. отличного от нуля магнитного момента в отсутствие внешнего магнитного поля. Если такой момент есть, то вещество будет парамагнитным, если же его нет — диамагнитным.

Собственный магнитный момент атома может иметь двойное происхождение: во-первых, он возникает вследствие движения атомных электронов и, во-вторых, электрон обладает сам по себе собственным магнитным моментом (так называемым спиновым магнитным моментом).

Магнитный момент, связанный с движением электронов, называют *орбитальным*. Возникновение его легко понять. Действительно, рассмотрим электрон, равномерно вращающийся со скоростью v по окружности радиуса r . Очевидно, сила тока, связанного с этим движением, равна $i = e/T$, где T — период обращения электрона по окружности. Умножив i на s_m , где s_m — площадь ориентированной площадки, обтекаемой током (направление s_m связано с направлением тока правилом правого винта), найдем орбитальный магнитный момент электрона $m_L = is_m$. Учитывая, что $s_m = \frac{1}{2}[rv]T$, можно переписать m_L в виде

$$m_L = \frac{1}{2}e[rv].$$

С другой стороны, орбитальный момент импульса электрона $K = m_e[rv]$, где m_e — масса электрона; поэтому

$$m_L = \frac{e}{2m_e} K.$$

Мы получили эту формулу, основываясь на определении момента импульса, данном в классической механике. Но

такая же формула справедлива и в квантовой механике, только K при этом не обычный, а «квантовый» вектор. Это значит, что его проекция на избранную ось (направление внешнего магнитного поля) может принимать не любые, а только вполне определенные дискретные значения $m\hbar$, где $\hbar = 10^{-34}$ Дж·с — квантовая постоянная и m — целое положительное или отрицательное число, заключенное в некоторых пределах $-l \leq m \leq l$ (l — положительное целое число). При этом квадрат орбитального момента импульса электрона $K^2 = \hbar^2 l(l+1)$.

Мы приходим, таким образом, к выводу, что проекция орбитального магнитного момента электрона m_L на направление поля (ось z) может принимать только следующие значения:

$$(m_L)_z = -\mu_B m, \quad -l \leq m \leq l,$$

где $\mu_B = e\hbar/(2m_e) = 9,2732 \cdot 10^{-24}$ Дж/Тл — так называемый *магнетон Бора*.

Эта величина представляет собой минимальную «порцию» магнитного момента, или квант магнитного момента.

Такие же формулы справедливы и для атома в целом — следует лишь под K понимать суммарный орбитальный момент импульсов всех электронов атома. Так как он выражается в единицах \hbar , то его записывают в виде $\hbar L$, где L — безразмерный квантовый вектор, т. е. вектор с целочисленными проекциями $-l, (-l+1), \dots, 0, 1, 2, \dots, l$ на направление поля. Орбитальный магнитный момент атома можно, следовательно, записать в виде

$$m_L = \frac{e\hbar}{2m_e} L.$$

(Так как $e < 0$, то вектор \mathbf{m}_L направлен противоположно вектору \mathbf{L} .)

Но и покоящийся электрон также обладает магнитным моментом. Он пропорционален внутреннему моменту количества движения электрона, называемому *спином* электрона.

Спин электрона представляет собой квантовый вектор, проекция которого на избранное направление — направление магнитного поля — может принимать только два дискретных значения: $-\hbar/2$ и $+\hbar/2$. Этим моментом электрон обладает всегда, т. е. независимо от того, покоится ли он или движется. Со спином связан магнитный момент электрона, равный магнетону Бора.

Если обозначить спиновый магнитный момент \mathbf{m}_s , то для электрона

$$\mathbf{m}_s = \frac{e\hbar}{m_e} \mathbf{s},$$

где $\hbar\mathbf{s}$ — спиновый момент количества движения электрона и \mathbf{s} — безразмерный квантовый вектор с двумя возможными проекциями $\pm 1/2$, на направление магнитного поля. Вектор \mathbf{m}_s может ориентироваться либо вдоль магнитного поля, либо против него, и проекции его равны при этом $\pm \mu_B$. Другие ориентации спинового магнитного момента электрона невозможны.

Формулой такого же типа определяется спиновый магнитный момент всего атома в целом — нужно лишь заменить \mathbf{s} на сумму спиновых моментов импульсов всех электронов атома: $\mathbf{S} = \sum \mathbf{s}$. Таким образом, спиновый магнитный момент атома равен

$$\mathbf{m}_S = \frac{e\hbar}{m_e} \mathbf{S}.$$

Сложив \mathbf{m}_L и \mathbf{m}_s , мы получим *суммарный магнитный момент* атома \mathbf{m} :

$$\mathbf{m} = \mathbf{m}_L + \mathbf{m}_S = \frac{e\hbar}{2m_e} (\mathbf{L} + 2\mathbf{S}).$$

Сумма орбитального и спинового моментов импульсов дает полный момент импульса атома $\hbar\mathbf{J}_a$:

$$\mathbf{J}_a = \mathbf{L} + \mathbf{S}.$$

Мы видим, что суммарный магнитный момент атома \mathbf{m} не параллелен суммарному моменту импульса атома, хотя для обеих составляющих — орбитального и спинового моментов — параллельность и имеет место. Если $\mathbf{L} = \mathbf{S} = 0$, то собственный магнитный момент атома равен нулю.

Посмотрим теперь, как будет вести себя во внешнем магнитном поле газ, атомы которого обладают собственным магнитным моментом. Сложив магнитные моменты \mathbf{m} отдельных атомов в единице объема газа, мы найдем намагниченность газа: $\mathbf{J} = \sum \mathbf{m}$. В отсутствие поля такая сумма обращается в нуль, так как моменты отдельных атомов ориентированы самым различным образом. При $\mathbf{H} \neq 0$ возникает избранное направление — направление \mathbf{H} и проекция \mathbf{J} на это направление становится отличной от нуля. Вычислим эту проекцию, предполагая,

что газ находится в состоянии статистического равновесия. В этом случае вероятность w различных ориентаций \mathbf{m} относительно \mathbf{H} определяется формулой Больцмана $w \propto \exp[-U/(kT)]$, где T — температура газа и $U = -\mu_0 \mathbf{m} \mathbf{H}$ — потенциальная энергия магнитного момента \mathbf{m} в поле \mathbf{H} .

Будем сперва считать, что магнитный момент атома имеет орбитальное происхождение. Тогда $U = \mu_0 \mu_B m H$ ($-l \leq m \leq l$), и вероятность того, что проекция \mathbf{m}_L на \mathbf{H} равна $m \mu_B$, определится формулой

$$w_m = C \exp[-m \mu_0 \mu_B H / (kT)],$$

где $C = 1/Z$ — нормировочная константа, а

$$Z \equiv Z[\mu_0 \mu_B H / (kT)] = \sum_{m=-l}^l \exp[-m \mu_0 \mu_B H / (kT)].$$

Среднее значение m составляет

$$\langle m \rangle = \frac{\sum_{-l}^l m w_m}{\sum_{-l}^l w_m} = \frac{\sum_{-l}^l m e^{-m \xi}}{\sum_{-l}^l e^{-m \xi}},$$

где $\xi = \mu_0 \mu_B H / (kT)$. Эту величину можно также представить в виде

$$\langle m \rangle = -\frac{d}{d\xi} \ln Z(\xi).$$

Легко вычислить $\langle m \rangle$ в случае слабых полей или высоких температур, когда $\xi \ll 1$. При этом

$$Z(\xi) = \sum_{m=-l}^l \left(1 + \frac{1}{2} [m^2 \xi^2]\right) = (2l+1) \left[1 + \frac{1}{6} l(l+1) \xi^2\right].$$

Отсюда следует, что если $\xi \ll 1$, то

$$\langle m \rangle = \frac{Z'(\xi)}{Z(\xi)} = \frac{1}{3} l(l+1) \xi.$$

Умножив $\langle m \rangle$ на μ_B и на n (концентрация атомов), мы получим проекцию J_{\parallel} вектора намагниченности \mathbf{J} на направления поля:

$$J_{\parallel} = \frac{\mu_0 \mu_B^2}{3kT} \frac{1}{3} l(l+1) n H.$$

Проекция \mathbf{J} , перпендикулярная \mathbf{H} , равна нулю, так что $J_{\parallel} = J$. Вспоминая определение магнитной восприимчивости $\mathbf{J} = \chi \mathbf{H}$, мы видим, что восприимчивость рассматриваемого газа равна

$$\chi = \mu_0 \frac{\mu_B^2 n}{3kT} l(l+1).$$

Она положительна, т. е. газ, атомы которого обладают отличным от нуля собственным магнитным моментом, является парамагнетиком.

Полученная формула для χ справедлива при $\mu_0 \mu_B H \ll kT$. В этих условиях восприимчивость обратно пропорциональна температуре (это утверждение называют законом Кюри). Так как $\mathbf{m}_L = \mu_B \mathbf{L}$ и $L^2 = l(l+1)$, то формулу для χ можно записать в виде

$$\chi = \frac{\mu_0 m_L^2 n}{3kT}$$

Эта формула аналогична формуле $\alpha = d_0^2 n / (3\epsilon_0 kT)$ для электрической восприимчивости газа полярных молекул (см. гл. 8). Она может быть получена (так же, как и формула для α), если считать \mathbf{m}_L классическим вектором и пространственное распределение векторов \mathbf{m}_L определить по формуле

$$dn = C \exp[-U/(kT)] d\Omega,$$

где $U = -\mu_0 \mathbf{m}_L \mathbf{H} = -\mu_0 m_L H \cos \nu$ и $d\Omega = 2\pi \sin \nu d\nu$ (ν — угол между векторами \mathbf{m}_L и \mathbf{H}). При этом среднее значение $\cos \nu$ определяется формулой

$$\langle \cos \nu \rangle = \int \cos \nu \exp\left(\mu_0 \frac{m_L H}{kT} \cos \nu\right) d\Omega \left/ \int \exp\left(\mu_0 \frac{m_L H}{kT} \cos \nu\right) d\Omega \right.$$

и намагниченность — формулой $J = nm \langle \cos \nu \rangle$.

До тех пор, пока рассматриваются высокие температуры ($\mu_0 \mu_B H \ll kT$), различие в обоих подходах (\mathbf{m}_L — «квантовый» вектор и \mathbf{m}_L — «классический» вектор) не проявляется. Однако если перейти к низким температурам $\mu_0 \mu_B H \geq kT$, то различие будет очень существенным. Более того, формула, которая получается для J , если предполагать, что угол ν между векторами \mathbf{m}_L и \mathbf{H} может принимать произвольные значения, — формула Ланжевена

$$J = nm_L L\left(\frac{\mu_0 m_L H}{kT}\right), \quad L(\xi) = \coth \xi - \frac{1}{\xi}$$

($L(\xi)$ называют функцией Ланжевена) является, строго говоря, неправильной. Дело в том, что из нее вытекает, что энтропия газа при $T \rightarrow 0$ стремится к бесконечности, тогда как по теореме Нернста она должна стремиться к нулю. При квантовом подходе этого противоречия с теоремой Нернста не возникает.

Мы предполагали при расчете χ , что собственный магнитный момент атома имеет орбитальное происхождение. Предположим теперь, что он имеет спиновое происхождение, т. е. $\mathbf{m} = \mathbf{m}_S = (e\hbar/m_e) \mathbf{S}$. В этом случае возможные значения проекции \mathbf{m}_S на \mathbf{H} равны $2\mu_B m'$, где $m' = -S, -S+1, \dots, S$ и S — максимальное значение проекции спина на направление \mathbf{H} (для одного электрона $s = 1/2$). Потенциальная энергия момента в поле по-прежнему определяется формулой $U = -\mu_0 \mathbf{m}_S \mathbf{H} = 2\mu_0 \mu_B H m'$, а вероятность данного значения m' — формулой

$$\omega_{m'} = e^{-2\mu_0\mu_B H m' / (kT)} \left/ Z [2\mu_0\mu_B H / (kT)]; \quad Z(\xi) = \sum_{m'=-S}^S e^{-\xi m'} \right.$$

Поэтому среднее значение m'

$$\langle m' \rangle = \frac{\sum m' e^{-m' \xi}}{\sum e^{-m' \xi}} = - \frac{d}{d\xi} Z(\xi); \quad \xi = \frac{2\mu_0\mu_B H}{kT}$$

Отсюда может быть найдена намагниченность: $J = 2\mu_B n \langle m' \rangle$. Предположим, для простоты, что $s = 1/2$. Тогда

$$\langle m' \rangle = \frac{1}{2} \frac{e^{\xi/2} - e^{-\xi/2}}{e^{\xi/2} + e^{-\xi/2}} = \frac{1}{2} \operatorname{th} \left(\mu_0 \frac{\mu_B H}{kT} \right)$$

и, следовательно,

$$J = \mu_B n \operatorname{th} \left(\mu_0 \frac{\mu_B H}{kT} \right).$$

Эта формула в соответствии со сказанным выше не совпадает с формулой Ланжевена. Если $\xi \ll 1$, то $\operatorname{th} \xi \approx \xi$; поэтому при $\mu_0\mu_B H \ll kT$

$$J = \frac{\mu_0 n \mu_B^2}{kT} H.$$

Таким образом, газ атомов, обладающих спиновым магнитным моментом, парамагнитен и его парамагнитная восприимчивость подчиняется закону Кюри.

Мы считали, что $S = 1/2$. При произвольном S формула для χ имеет вид

$$\chi = \frac{n\mu_0\mu_B^2}{3kT} 4S(S+1).$$

Если отличны от нуля и m_L и m_S и поле \mathbf{H} является слабым, то магнитная восприимчивость газа определяется формулой

$$\chi = \frac{n\mu_0\mu_B^2 g_L^2 j(j+1)}{3kT},$$

где j — максимальное значение проекции суммарного момента импульса атома $\mathbf{J}_a = \mathbf{L} + \mathbf{S}$ на направление напряженности поля \mathbf{H} и

$$g_L = \{1 + [j(j+1) + s(s+1) - l(l+1)]\} / 2j(j+1)$$

(эту величину называют *множителем Ланде*). Итак,

во всех случаях, когда атом обладает отличным от нуля собственным магнитным моментом, газ атомов парамагнитен, причем его парамагнитная восприимчивость зависит от температуры, убывая с ее ростом.

Так как магнитная проницаемость $\mu = 1 + \chi$ зависит от температуры, то величину

$$f_M = \mathbf{BH} / 2 = \mu_0 \mu H^2 / 2$$

при $T = \text{const}$ следует рассматривать как свободную энергию магнитного поля единичного объема газа, но она включает как энергию собственного магнитного поля, так и свободную энергию газа, связанную с магнитным полем.

12.3. Парамагнетизм вырожденного электронного газа

В предыдущем параграфе мы определили парамагнитную восприимчивость классического, т. е. невырожденного газа. Теперь мы рассмотрим вырожденный электронный газ в магнитном поле и определим его магнитную восприимчивость. Таким газом является, например, газ электронов проводимости в металлах.

Так как электрон обладает собственным спиновым магнитным моментом μ_B , то в поле H он приобретает добавочную энергию $\pm \mu_0 \mu_B H$ (знаки «+» и «-» соответствуют ориентациям спина по полю и против него). Поэтому если в отсутствие поля функция распределения вырожденного газа электронов, который мы будем предполагать равновесным, имела вид $f(\epsilon) = \{\exp[(\epsilon - \epsilon_0)/(kT)] + 1\}^{-1}$, где ϵ — энергия электрона, ϵ_0 — химический потенциал, то в поле мы получим две функции распределения $f_{\uparrow}(\epsilon)$ и $f_{\downarrow}(\epsilon)$ для ориентации спина — по полю и против него:

$$f_{\uparrow}(\epsilon) = \{\exp[(\epsilon + \mu_0 \mu_B H - \epsilon_0)/(kT)] + 1\}^{-1};$$

$$f_{\downarrow}(\epsilon) = \{\exp[(\epsilon - \mu_0 \mu_B H - \epsilon_0)/(kT)] + 1\}^{-1}.$$

Теперь легко найти намагниченность J газа: так как проекция спинового магнитного момента электрона на H равна $-\mu_B$ при ориентации спина электрона вдоль поля и μ_B при противоположной ориентации, то

$$J = \mu_B \int [f_{\downarrow}(\epsilon) - f_{\uparrow}(\epsilon)] d^3p / (2\pi\hbar)^3 = \mu_B \int \{[\exp((\epsilon - \mu_0 \mu_B H - \epsilon_0)/(kT)) + 1]^{-1} - [\exp((\epsilon + \mu_0 \mu_B H - \epsilon_0)/(kT)) + 1]^{-1}\} \nu(\epsilon) d\epsilon \frac{d\Omega}{4\pi},$$

где $\nu(\epsilon) d\epsilon d\Omega / (4\pi)$ — число состояний электрона в интервале энергий $(\epsilon, \epsilon + d\epsilon)$ и интервале телесных углов $d\Omega$ импульса (точнее квазиимпульса) p . Разность фермиевских распределений, входящих в этот интеграл, может быть заменена в случае слабых полей на $-2\mu_0 \mu_B H df/d\epsilon$, а так как производная $df/d\epsilon$ имеет резкий максимум при $\epsilon = \epsilon_0$, то

$$-2\mu_0 \mu_B H \frac{df}{d\epsilon} \approx 2\mu_0 \mu_B H \delta(\epsilon - \epsilon_0).$$

Поэтому интегрирование может быть выполнено сразу: $J = 2\mu_0 \mu_B^2 H \nu(\epsilon_0)$. Таким образом,

вырожденный электронный газ парамагнитен и его парамагнитная восприимчивость равна

$$\chi_p = 2\mu_0 \mu_B^2 \nu(\epsilon_0).$$

В отличие от восприимчивости невырожденного газа восприимчивость вырожденного газа не зависит от температуры и определяется, по существу, только значением плотности уровней энергии электрона на границе фермиевского распределения. Парамагнетизм вырожденного газа носит название *паулевского парамагнетизма*.

Если считать, что $\epsilon = p^2/(2m_e)$ (в металле такая зависимость энергии электрона ϵ от его квазиимпульса p , вообще говоря, не имеет места), то $\nu(\epsilon) = \sqrt{2\epsilon}m_e^{3/2}/(2\pi^2\hbar^3)$. Учитывая, что

$$2(4\pi/3)p_F^3/(2\pi\hbar)^3 = n, \quad p_F^2/(2m_e) = \epsilon_0$$

(p_F — граничный импульс, n — плотность электронов), получим

$$\chi_p = \frac{\mu_0}{\pi} \left(\frac{3}{\pi} \right)^{1/3} \frac{\mu_B^2 n^{1/3} m_e}{\hbar^3}$$

По порядку величины $\chi_p \sim \mu_0 n \mu_B^2 / (kT_0)$, где T_0 — температура вырождения (эта формула отличается от формулы для парамагнитной восприимчивости обычного газа заменой T на T_0). Полагая $n \simeq 10^{29} \text{ м}^{-3}$, $T_0 \simeq 10^5 \text{ К}$, получим $\chi_p \simeq 10^{-6}$.

При вычислении магнитной восприимчивости электронного газа мы считали, что влияние магнитного поля сводится лишь к изменению энергии электрона. Иными словами, мы считали, что зависимость ϵ от p не изменяется при включении поля и роль его сводится только к тому, что к $\epsilon(p)$ прибавляется $\pm \mu_0 \mu_B H$. Между тем поле изменяет и характер движения электрона. Действительно, как мы знаем, магнитное поле искривляет траекторию частицы (в плоскости, перпендикулярной \mathbf{H} , движение электрона происходит по окружности). Учет этого эффекта приводит к диамагнетизму электронов, в результате чего парамагнитная восприимчивость свободного электронного газа уменьшается на $1/3$ своего значения (*диамагнетизм Ландау*). Явление это играет важную роль в магнетизме металлов.

12.4. Восприимчивость диамагнитного газа

Мы должны показать теперь, что если атомы вещества не обладают собственным магнитным моментом (моментом в отсутствие внешнего магнитного поля), то вещество является диамагнитным, т. е. его магнитная восприимчивость отрицательна. Выясним с этой целью, какое влияние оказывает внешнее магнитное поле на движение электронов в атоме.

Если атом находится в магнитном поле с индукцией \mathbf{B} , то каждый его электрон испытывает действие силы Лоренца $e[\mathbf{vB}]$ (\mathbf{v} — скорость электрона в инерциальной системе отсчета). Поэтому если на электрон со стороны ядра и других электронов атома действует сила \mathbf{f} , то суммарная сила, действующая на электрон, составляет $\mathbf{F} = \mathbf{f} + e[\mathbf{vB}]$. Предположим, что сила \mathbf{f} имеет центральный характер, т. е. зависит только от расстояния r между электроном и ядром, и перейдем к системе отсчета, вращающейся с постоянной угловой скоростью

ω относительно исходной (инерциальной) системы отсчета. В неинерциальной системе отсчета появятся две добавочные силы — сила Кориолиса $2m_e[\mathbf{v}'\omega]$ и центробежная сила $m_e[\omega[\mathbf{r}\omega]]$, где \mathbf{r} — радиус-вектор, определяющий положение частицы, и $\mathbf{v}' = \mathbf{v} - [\omega\mathbf{r}]$ — ее скорость в неинерциальной системе отсчета. Поэтому результирующая сила в неинерциальной системе

$$\mathbf{F}' = \mathbf{f}' + e[\mathbf{v}'\mathbf{B}] + 2m_e[\mathbf{v}'\omega] + m_e[\omega[\mathbf{r}\omega]],$$

где \mathbf{f}' — сила, соответствующая \mathbf{f} в неинерциальной системе. Так как мы предполагаем, что сила \mathbf{f} зависит только от расстояния r , а оно не меняется при переходе к неинерциальной системе, то, очевидно, $\mathbf{f}' = \mathbf{f}$. Предположим теперь, что ω выбрана равной

$$\omega = \omega_L \equiv -e\mathbf{B}/(2m_e).$$

Тогда сила Лоренца сократится с силой Кориолиса. Если теперь считать поле \mathbf{H} достаточно слабым, то в выражении для силы \mathbf{F}' в неинерциальной системе можно будет пренебречь центробежной силой, которая пропорциональна ω^2 , и мы получим $\mathbf{F}' = \mathbf{f}' = \mathbf{f}$.

Таким образом, если угловая скорость неинерциальной системы равна ω_L , то сила, действующая на электроны в этой системе, с точностью до квадратичных по ω_L членов совпадает с силой, действующей на электроны в инерциальной системе в отсутствие магнитного поля. Поэтому можно утверждать, что электроны в атоме, находящемся в слабом однородном и постоянном магнитном поле, движутся так же, как они двигались бы в отсутствие поля, но во вращающейся системе отсчета, угловая скорость которой равна ω_L . Это же утверждение можно сформулировать и иначе:

в слабом магнитном поле весь атом, как целое, прецессирует вокруг направления магнитного поля с угловой скоростью ω_L .

Это утверждение носит название *теоремы Лармора*, а величину ω_L называют *ларморовой частотой* (заметим, что ларморова частота в два раза меньше циклотронной частоты электрона, движущегося по круговой орбите в магнитном поле \mathbf{H}).

Подчеркнем, что теорема Лармора справедлива только в случае достаточно слабого магнитного поля, а именно необходимо, чтобы центробежная сила была значительно меньше силы Кориолиса, т. е. должно выполняться неравенство $v_a \gg \omega_L a$, где a — длина порядка размеров атома и v_a — скорость порядка скорости электронов в атоме. Ясно также, что справедливость теоремы Лармора тесно связана с тем, что отношение e/m одинаково для всех подвижных частиц атома (т. е. электронов). Если бы в состав атома входили подвижные частицы различных сортов, для которых отношение заряда к массе было бы неодинаковым, то теорема Лармора не имела бы места.

В справедливости теоремы Лармора можно убедиться еще следующим путем. Выше мы видели, что если атом обладает орбитальным моментом импульса \mathbf{K} , то он обладает и магнитным моментом $\mathbf{m}_L = [e/(2m_e)]\mathbf{K}$. С другой стороны, магнитный момент \mathbf{m}_L в магнитном поле имеет потенциальную энергию $U = -\mathbf{m}_L \mu_0 \mathbf{H} = -m_L B \cos\vartheta$

(ϑ — угол между \mathbf{m}_L и \mathbf{H}). Поэтому на магнитный момент действует пара сил с моментом $N = -dU/d\vartheta = m_L B \sin\vartheta$, или в векторной форме

$$\mathbf{N} = [m_L \mathbf{B}].$$

Производная по времени от момента импульса \mathbf{K} должна быть равна моменту сил: $\dot{\mathbf{K}} = [m_L \mathbf{B}]$. Вспоминая связь между \mathbf{m}_L и \mathbf{K} , получим отсюда

$$\dot{\mathbf{K}} = \frac{e}{2m_e} [\mathbf{K}\mathbf{B}].$$

Эта формула имеет простой физический смысл: вспоминая, что скорость точек вращающегося твердого тела определяется формулой $\mathbf{v} \equiv d\mathbf{r}/dt = [\boldsymbol{\omega}\mathbf{r}]$, где $\boldsymbol{\omega}$ — вектор угловой скорости, \mathbf{r} — радиус-вектор точки и \mathbf{v} — ее скорость, легко заключить, что вектор \mathbf{K} совершает прецессионное движение около направления магнитного поля с угловой скоростью $\boldsymbol{\omega}_L = -e\mathbf{B}/(2m_e)$. Это и есть ларморова частота. Вместе с вектором \mathbf{K} в прецессии Лармора участвует весь атом.

Но с другой стороны, мы приходим и к некоторому парадоксу. Действительно, при прецессии вектора \mathbf{K} около направления поля \mathbf{H} ни модуль этого вектора, ни его наклон по отношению к \mathbf{H} не изменяются. Магнитный момент атома \mathbf{m}_L (это собственный момент!) пропорционален моменту импульса атомов \mathbf{K} . Поэтому вектор \mathbf{m}_L также совершает равномерную прецессию около направления поля \mathbf{H} , при которой ни модуль \mathbf{m}_L , ни угол наклона \mathbf{m}_L по отношению к \mathbf{H} не изменяются. Если теперь рассмотреть идеальный газ атомов с различно ориентированными магнитными моментами \mathbf{m}_L , то мы увидим, что эти моменты независимо прецессируют около \mathbf{H} , не изменяя наклона к полю; следовательно, газ никогда не сможет намагнититься. Разрешение этого парадокса заключается в том, что газ не является строго идеальным — в нем всегда происходят столкновения между атомами, и именно эти столкновения (либо другие процессы взаимодействия атомов с объектами, не входящими в состав газа) и приводят к намагничению газа, т. е. к изменению ориентации магнитных моментов его атомов относительно поля. Ясно, что чем столкновения реже, тем дольше длится процесс намагничения, так что можно говорить о *времени релаксации намагничения*.

Теперь мы можем доказать, что если собственный магнитный момент у атомов отсутствует, то газ таких атомов — диамагнитный. Будем исходить из формулы для орбитального магнитного момента атома

$$\mathbf{m}_L = \sum_{a \in \text{ат}} \frac{e}{2} [\mathbf{r}_a \mathbf{v}_a],$$

где \mathbf{r}_a и \mathbf{v}_a — радиус-вектор и скорость a -го электрона в атоме и суммирование производится по всем электронам, входящим в состав атома. Вектор скорости электрона можно согласно теореме Лармора записать в виде $\mathbf{v}_a = \mathbf{v}_a^0 + [\boldsymbol{\omega}_L \mathbf{r}_a]$, где \mathbf{v}_a^0 — скорость в отсутствие магнитного поля и $\boldsymbol{\omega}_L$ — ларморова частота. Подставляя это выражение в формулу для \mathbf{m}_L , получим

$$\mathbf{m}_L = \sum_a \frac{e}{2} [\mathbf{r}_a \mathbf{v}_a^0] + \sum_a \frac{e}{2} [\mathbf{r}_a [\boldsymbol{\omega}_L \mathbf{r}_a]].$$

Первое слагаемое здесь представляет собой собственный магнитный момент атома, который мы обозначим \mathbf{m}_L^0 , а второе слагаемое — магнитный момент, обусловленный ларморовой прецессией. Это слагаемое, которое мы обозначим \mathbf{m}_d , можно переписать в виде

$$\mathbf{m}_d = \sum_a \frac{e}{2} [\mathbf{r}_a [\boldsymbol{\omega}_L \mathbf{r}_a]] = \sum_a \frac{e}{2} \{ \boldsymbol{\omega}_L r_a^2 - \mathbf{r}_a (\boldsymbol{\omega}_L \mathbf{r}_a) \}.$$

Будем полагать, что атом сферически симметричен. Тогда

$$\sum_a x_a y_a = \sum_a y_a z_a = \sum_a z_a x_a = 0;$$

$$\sum_a x_a^2 = \sum_a y_a^2 = \sum_a z_a^2 = \frac{1}{3} \sum_a r_a^2$$

и поэтому $\mathbf{m}_d = \frac{e}{3} \boldsymbol{\omega}_L \sum_a r_a^2$. Подставляя сюда $\boldsymbol{\omega}_L = -e\mathbf{B}/(2m_e)$ и вводя обозначение $\sum_a r_a^2 = Z \langle r^2 \rangle$, где Z — число электронов в атоме и $\langle r^2 \rangle$ — среднее значение квадрата расстояния электрона до ядра, получим

$$\mathbf{m}_d = -\frac{e^2 \mu_0}{6m_e} Z \langle r^2 \rangle \mathbf{H}.$$

Мы видим, что магнитный момент атома, обусловленный ларморовой прецессией, направлен против магнитного поля (этот вывод не зависит от знака заряда электрона!).

Полный магнитный момент атома

$$\mathbf{m}_L = \mathbf{m}_L^0 - \frac{e^2 \mu_0}{6m_e} Z \langle r^2 \rangle \mathbf{H}.$$

Поэтому намагниченность газа определяется формулой

$$J = n m_L^0 \langle \cos \vartheta \rangle - \frac{e^2 \mu_0}{6m_e} Z \langle r^2 \rangle n H,$$

где $\langle \cos \vartheta \rangle$ — среднее значение косинуса угла между векторами \mathbf{m}_L^0 и \mathbf{H} . Первое слагаемое приводит к парамагнитной, а второе — к диамагнитной восприимчивости. Обычно первое слагаемое значительно больше второго, поэтому при $m_L^0 \neq 0$ второе слагаемое можно не учитывать. Но если $m_L^0 = 0$, то остается только второе слагаемое; в этом случае газ является диамагнитным с диамагнитной восприимчивостью

$$\chi_d = -\frac{e^2 \mu_0}{6m_e} Z \langle r^2 \rangle n.$$

В отличие от парамагнитной восприимчивости χ_p обычных газов, зависящей от температуры, диамагнитная восприимчивость от температуры не зависит. Величина χ_d очень мала. Например, для гелия

$\chi_d = -0,2 \cdot 10^{-6}$, для висмута $\chi_d = -14,5 \cdot 10^{-6}$. Для газов χ_p гораздо больше $|\chi_d|$. По порядку величины отношение

$$|\chi_d| / \chi_p = \frac{e^2}{m_e} \frac{\langle r^2 \rangle}{\mu_B^2} kT,$$

а так как $\mu_B = e\hbar/(2m_e)$ и $\hbar^2/(m_e \langle r^2 \rangle)$ определяет по порядку величины энергию атома E_a , то

$$|\chi_d| / \chi_p \sim kT/E_a.$$

Учитывая, что $E_a \sim 1 \text{ эВ} \sim 10\,000 \text{ К}$, мы видим, что для невырожденных газов отношение $|\chi_d|/|\chi_p|$ очень мало.

Диамагнитная восприимчивость пропорциональна квадрату размеров атома (или молекулы). Поэтому $|\chi_d|$ относительно больше для веществ, состоящих из больших молекул (примером является бензол).

12.5. Гальваномагнитные явления в металлах

Влияние магнитного поля на движение электронов проводимости в металлах приводит к ряду явлений, объединенных под названием *гальваномагнитных*. Во-первых, под действием магнитного поля изменяется сопротивление проводника. Во-вторых, если в отсутствие магнитного поля напряженность электрического поля \mathbf{E} и плотность тока имеют одинаковые направления (металл считается изотропным), то при наличии поля возникает добавочное электрическое поле, перпендикулярное \mathbf{B} и \mathbf{j} . Это поле называют *холловским*, а само явление — *эффектом Холла*. Таким образом, если при $\mathbf{B} = 0$, согласно закону Ома, $\mathbf{E} = \rho_0 \mathbf{j}$, то при $\mathbf{B} \neq 0$ электрическое поле и плотность тока связаны соотношением

$$\mathbf{E} = \rho \mathbf{j} + R[\mathbf{Bj}],$$

где ρ_0 и ρ — удельные сопротивления проводника при $\mathbf{B} = 0$ и $\mathbf{B} \neq 0$ и R — некоторая постоянная (*постоянная Холла*).

Для того чтобы понять, как возникают холловские поля, представим себе, что сопротивление отсутствует, т. е. электроны не испытывают столкновений. Тогда на них действует только сила $\mathbf{F} = e(\mathbf{E} + [\mathbf{vB}])$. При стационарном движении электронов эта сила должна обращаться в нуль. Как легко убедиться, для этого скорость электронов должна быть равна так называемой *скорости дрейфа* $\mathbf{v}_d = [\mathbf{EB}]B^2$. Умножив \mathbf{v}_d на en (n — плотность электронов), найдем плотность тока:

$$\mathbf{j} = en\mathbf{v}_d = \frac{en}{B^2} [\mathbf{EB}].$$

Если $\mathbf{E} \perp \mathbf{B}$, то

$$\mathbf{E} = R[\mathbf{Bj}], \quad R = 1/(en)$$

Это и есть холловское поле. Мы видим, что постоянная Холла обратно пропорциональна плотности электронов.

К этому полю должно быть добавлено поле, связанное с столкновениями электронов. Чтобы найти его, нужно определить плотность тока при наличии полей \mathbf{E} и \mathbf{V} с учетом столкновений.

Задача эта должна решаться с помощью кинетического уравнения для функции распределения электронов $f(\mathbf{p})$ в присутствии полей \mathbf{E} и \mathbf{V} . Кинетическое уравнение при $\mathbf{V} = 0$ было сформулировано в гл. 9. Если $\mathbf{V} \neq 0$ и функция распределения f мало отличается от равновесного фермиевского распределения f_0 , то кинетическое уравнение имеет вид

$$e(\mathbf{E} + [\mathbf{vB}]) \frac{\partial f}{\partial \mathbf{p}} = -\frac{f - f_0}{\tau},$$

где $\mathbf{v} = d\mathbf{e}/d\mathbf{p}$ — скорость электрона и τ — среднее время между столкновениями. Это уравнение отличается от кинетического уравнения гл. 9 только тем, что в него входит сила Лоренца $e(\mathbf{E} + [\mathbf{vB}])$ вместо электрической силы $e\mathbf{E}$.

Вводя отклонение функции распределения от равновесной $f - f_0 = \varphi$ (причем $|\varphi| \ll f_0$) и учитывая, что

$$[\mathbf{vB}] \frac{df_0}{d\mathbf{p}} = [\mathbf{vB}] \frac{df_0}{d\mathbf{e}} \quad \mathbf{v} = 0,$$

получим из кинетического уравнения следующее уравнение для определения неизвестной функции $\varphi(\mathbf{p})$:

$$\frac{1}{\tau} \varphi(\mathbf{p}) + e[\mathbf{vB}] \frac{\partial \varphi(\mathbf{p})}{\partial \mathbf{p}} = -e\mathbf{E}\mathbf{v} \frac{df_0}{d\mathbf{e}}.$$

Электрическое поле считается слабым, и поэтому здесь пренебрегают членом $e\mathbf{E}\partial\varphi/\partial\mathbf{p}$.

Рассмотрим тот случай, когда $\mathbf{E} \perp \mathbf{V}$. Направив \mathbf{V} вдоль оси z , перепишем уравнение для φ в виде

$$\frac{1}{\tau} \varphi + eB \left(v_y \frac{\partial}{\partial p_x} - v_x \frac{\partial}{\partial p_y} \right) \varphi = -e(E_x v_x + E_y v_y) \frac{df_0}{d\mathbf{e}}.$$

Будем сперва предполагать, что энергия ϵ электрона проводимости связана с его импульсом (точнее, квазиимпульсом) \mathbf{p} таким же соотношением $\epsilon = p^2/(2m_e)$, как и в случае свободного электрона. Тогда функцию φ следует искать в виде

$$\varphi(\mathbf{p}) = \Phi_1(\epsilon) v_x + \Phi_2(\epsilon) v_y,$$

где Φ_1 и Φ_2 — некоторые функции ϵ . Подстановка этого выражения в уравнение для φ дает

$$\frac{1}{\tau} \Phi_1 - \omega_c \Phi_2 = -eE_x \frac{df_0}{d\epsilon}; \quad \frac{1}{\tau} \Phi_2 + \omega_c \Phi_1 = -eE_y \frac{df_0}{d\epsilon},$$

откуда

$$\Phi_1 = \frac{-e\tau(E_x + \omega_c \tau E_y)}{1 + \omega_c^2 \tau^2} \frac{df_0}{d\epsilon}; \quad \Phi_2 = \frac{-e\tau(E_y - \omega_c \tau E_x)}{1 + \omega_c^2 \tau^2} \frac{df_0}{d\epsilon},$$

где $\omega_c = eB/m_e$ — циклотронная частота. Найдем плотность тока:

$$\mathbf{j} = 2e \int \mathbf{v} f \mathbf{v}(\epsilon) d\epsilon \frac{d\Omega}{4\pi},$$

где $\mathbf{v}(\epsilon)$ — плотность уровней электрона. Учитывая, что $df_0/d\epsilon = -\delta(\epsilon - \epsilon_0)$, получим

$$j_x = \sigma_0 \frac{E_x + \omega_c \tau_0 E_y}{1 + \omega_c^2 \tau_0^2}; \quad j_y = \sigma_0 \frac{E_y - \omega_c \tau_0 E_x}{1 + \omega_c^2 \tau_0^2},$$

где $\tau_0 \equiv \tau(\varepsilon_0)$ и $\sigma_0 = e^2 n \tau_0 / m_e$ — электрическая проводимость при $H = 0$. Пусть ток течет вдоль оси x , тогда $j_y = 0$, и, следовательно, $E_y = \omega_c \tau_0 E_x$. Подставляя это выражение в формулу для j_x , получим $j_x = \sigma_0 E_x$. Таким образом, мы пришли к старой формуле для плотности тока с проводимостью σ_0 , не зависящей от B . Действия поля B сводится лишь к возникновению поперечного (холловского) поля

$$E_y = \omega_c \tau_0 E_x = \frac{\omega_c \tau_0}{\sigma_0} j_x = R B j_x,$$

где $R = 1/(en)$ — постоянная Холла.

Такой вывод связан с нашим предположением о характере зависимости энергии электрона от импульса (квазиимпульса); именно: мы предполагали, что $\varepsilon = p^2/(2m_e)$, а в этом случае значение скорости на границе фермиевского распределения $\varepsilon = \varepsilon_0$ одинаково для всех электронов и действие поперечного электрического поля E_y компенсирует поперечную силу, создаваемую магнитным полем. Электроны находятся поэтому только под действием продольного поля E_x и столкновений. Если значения скорости при $\varepsilon = \varepsilon_0$ различны для разных электронов, то ситуация существенно меняется. Разъясним это на примере двух групп электронов проводимости, находящихся в двух энергетических зонах. Будем считать, что в каждой из зон энергия электрона пропорциональна квадрату квазиимпульса, но массы электрона различны для обеих зон. Различны также времена свободного пробега и могут отличаться знаки носителей заряда в обеих зонах. Тогда плотности тока $j^{(1)}$ и $j^{(2)}$ для каждой из зон определяются прежними формулами

$$j_x^{(\alpha)} = \frac{\sigma_\alpha E_x + R_\alpha B \sigma_\alpha^2 E_y}{1 + R_\alpha^2 B^2 \sigma_\alpha^2}, \quad j_y^{(\alpha)} = \frac{\sigma_\alpha E_y - R_\alpha B \sigma_\alpha^2 E_x}{1 + R_\alpha^2 B^2 \sigma_\alpha^2},$$

где индекс $\alpha = 1, 2$ нумерует зоны. Проекция суммарного тока $j_x = j_x^{(1)} + j_x^{(2)}$; $j_y = j_y^{(1)} + j_y^{(2)}$.

Положив, как и выше, $j_y = 0$, найдем холловское поле:

$$E_y = \frac{R_1 \sigma_1^2 + R_2 \sigma_2^2 + R_1 R_2 \sigma_1^2 \sigma_2^2 B^2 (R_1 + R_2)}{\sigma_1 + \sigma_2 + \sigma_1 \sigma_2 (R_1^2 \sigma_1 + R_2^2 \sigma_2) B^2} B E_x.$$

Подставив далее это выражение в формулы для $j_x^{(\alpha)}$, можно связать j_x и E_x , а именно $j_x = \rho E_x$, где ρ — сопротивление при наличии магнитного поля. Записав ρ в виде $\rho = \rho_0 + \Delta\rho$, где ρ_0 — сопротивление в отсутствие поля, получим

$$\frac{\Delta\rho}{\rho_0} = \frac{\sigma_1 \sigma_2 (\sigma_1 R_1 - \sigma_2 R_2)^2 B^2}{(\sigma_1 + \sigma_2)^2 + B^2 \sigma_1^2 \sigma_2^2 (R_1 + R_2)^2}.$$

Эта величина отлична от нуля, если $\sigma_1 R_1 \neq \sigma_2 R_2$, т. е. если $\tau_1/(m_1 e_1) \neq \tau_2/(m_2 e_2)$. Мы видим, что стоит лишь предположить $m_1 \neq m_2$, как $\Delta\rho$ становится отличным от нуля, даже при $\tau_1 = \tau_2$ и $e_1 = e_2$. Далее мы видим, что в слабых полях $\Delta\rho \sim B^2$.

Подставив в формулу для E_y поле $E_x = (1/\rho)j_x$, найдем постоянную Холла $R = E_y/(B j_x)$. Она, так же как и ρ , зависит от B и времен свободного пробега $\tau_{1,2}$. Для слабых полей

$$R = \frac{R_1 \sigma_1^2 + R_2 \sigma_2^2}{(\sigma_1 + \sigma_2)^2}.$$

Обратим внимание на то, что эта величина зависит не только от плотностей электронов, но и от времен их свободных пробегов.

Посмотрим теперь, как ведут себя магнитосопротивление ρ и постоянная Холла R в области сильных магнитных полей. Легко убедиться, что в этом случае $1/R = 1/R_1 + 1/R_2$, откуда

$$R = \frac{1}{e_1 n_1 + e_2 n_2}.$$

Мы специально выписали здесь разные заряды для носителей зарядов в обеих зонах, так как электроны в почти заполненной зоне ведут себя, как частицы с противоположным зарядом — дырки. Поэтому если проводимость обеих зон электронная, то $R = e(n_1 + n_2)^{-1}$. Если же проводимость первой зоны электронная, а второй — дырочная, то $R = e(n_1 - n_2)^{-1}$.

Поведение магнитосопротивления в области больших B различно в зависимости от того, отлична от нуля или равна нулю сумма холловских постоянных для обеих зон. Если $R_1 + R_2 \neq 0$, то в области больших полей магнитосопротивление достигает насыщения:

$$\frac{\Delta\rho}{\rho_0} = \frac{1}{\sigma_1 \sigma_2} \left(\frac{\sigma_1 R_1 - \sigma_2 R_2}{R_1 + R_2} \right)^2.$$

Напротив, если число электронов равно числу дырок, т. е. $R_1 + R_2 = 0$, то $\Delta\rho = \rho - \rho_0$ всегда пропорционально B^2 :

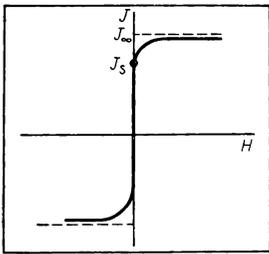
$$\frac{\Delta\rho}{\rho_0} = \frac{\sigma_1 \sigma_2 (\sigma_1 R_1 - \sigma_2 R_2)^2}{(\sigma_1 + \sigma_2)^2} B^2.$$

Мы выяснили, как ведет себя $\Delta\rho$ для модели двух энергетических зон, которые мы предполагали изотропными. Полученные результаты справедливы и для общего анизотропного случая, когда энергия электрона сложным образом зависит от квазимпульса. Зависимость сопротивления от B связана при этом с тем, что эффективная масса электрона зависит от направления его движения. В общем анизотропном случае в области слабых магнитных полей $\Delta\rho$ по-прежнему пропорционально B^2 . В сильных полях, как правило, $\Delta\rho$ не зависит от B . Исключением является случай равенства числа электронов и дырок, когда $\Delta\rho$ остается пропорциональным B^2 при сколь угодно сильных полях.

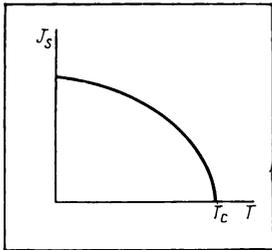
Рассмотренные гальваномагнитные явления могут быть названы *поперечными*, так как мы предполагали, что магнитное поле перпендикулярно электрическому полю и направлению электрического тока. На первый взгляд кажется, что при другой конфигурации полей никаких эффектов происходить не может, так как сила Лоренца отлична от нуля только в том случае, когда магнитное поле не параллельно электрическому току. Между тем существует еще *продольный гальваномагнитный эффект* — сопротивление металла меняется в магнитном поле, имеющем то же направление, что и плотность тока. Эффект этот имеет квантовое происхождение.

12.6. Характерные особенности ферромагнетиков

У большинства парамагнетиков магнитная восприимчивость в обычных условиях мала, поэтому их намагниченность J (даже в сильных полях) мала по сравнению с предельным значением $J_\infty = nm$, соответствующим бесконечному значению величины $\mu_0 m H / (kT)$ (m — магнитный момент атома, n — плотность атомов). Однако существует ряд твердых парамагнетиков, в первую очередь



12.1.
Идеализированная кривая намагничения ферромагнетика



12.2.
Зависимость спонтанной намагниченности от температуры

железо, у которых в обычных условиях уже при ничтожно малых полях намагничение достигает значений, сравнимых с J_∞ (у чистых монокристаллов железа это наблюдается в полях $H \sim 1$ А/м). Это так называемые *ферромагнетики*. К ним относятся помимо железа никель, кобальт, гадолиний, диспрозий и многочисленные соединения и сплавы, содержащие эти металлы (например, пермаллой, содержащий 20% Fe и 80% Ni, или сплав Гейслера, состоящий из 25% Ni, 60% Cu, 15% Al).

Схематически идеализированная кривая намагничения ферромагнетика изображена на рис. 12.1; здесь подчеркнуто то обстоятельство, что при $H \rightarrow 0$ имеется конечная намагниченность J_s . Ее называют *спонтанной намагниченностью*. Именно наличие спонтанной намагниченности отличает ферромагнетики от других кристаллов (парамагнетиков, диамагнетиков). При этом спонтанная намагниченность не обязательно сравнима с предельным значением намагниченности J_∞ ; существует целый ряд ферромагнетиков (так называемые *слабые ферромагнетики*, например соединение $ZrZn_2$) с малым по сравнению с J_∞ значением спонтанной намагниченности.

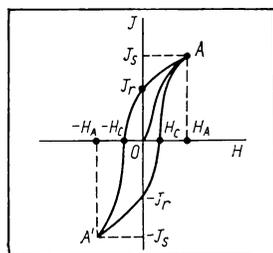
Спонтанная намагниченность J_s зависит от температуры и уменьшается с ее ростом. Характер этой зависимости иллюстрирует рис. 12.2. При некоторой температуре T_c , называемой *температурой Кюри*, J_s обращается в нуль. Так, температура Кюри равна 1043 К для железа, 631 К для никеля, 1388 К для кобальта, 289 К для гадолиния и 600 К для Cu_2AlMn . В области низких температур $T \ll T_c$ намагниченность J_s мало отличается от ее значения при $T = 0$; именно: *разность между этими величинами теоретически должна быть пропорциональна $T^{3/2}$* (закон Блоха). В частности, для ферромагнетиков, у которых намагниченность при нуле температур достигает своего предельного возможного значения J_∞ (к ним относятся, в частности, все ферромагнитные диэлектрики), имеем $1 - J_s/J_\infty \sim (T/T_c)^{3/2}$.

При температурах выше точки Кюри специфические ферромагнитные свойства исчезают и вещество становится обычным парамагнетиком, магнитная восприимчивость которого определяется формулой

$$\chi = C/(T - T_c),$$

где C — некоторая константа. Эту зависимость χ от температуры называют *законом Кюри — Вейсса*. В точке Кюри, при переходе от ферромагнитного в обычное парамагнитное состояние, наблюдается скачок в теплоемкости вещества.

На схематическом рис. 12.1 кривая намагничивания в точке $H = 0$ поднимается вертикально вверх, что соответствует идеализированному случаю $(dJ/dH)_{H=0} = \infty$, т. е. бесконечной магнитной восприимчивости и, следовательно, бесконечно большой магнитной проницаемости μ . В действительности величина dJ/dH конечна и поэтому конечна магнитная проницаемость, хотя она может быть и очень большой. Например, у железа она может достигать порядка нескольких тысяч, а у пермаллоя — сотен тысяч и даже миллиона.



12.3.
Петля гистерезиса

Но магнитная восприимчивость не может служить характеристикой магнитного состояния ферромагнетика. Дело в том, что в отличие от обычных парамагнетиков и диамагнетиков, которые можно назвать слабомагнитными веществами и которые характеризуются в области малых магнитных полей вполне определенной магнитной восприимчивостью, зависящей только от внутренних параметров, свойственных веществу (таких, как магнитный момент атома и плотность атомов), а также от температуры, для ферромагнетиков не существует однозначной и определенной магнитной восприимчивости, а в тех случаях, когда ее можно ввести, она оказывается зависящей от модуля и направления напряженности магнитного поля, процесса намагничивания, характера обработки ферромагнетика и присутствующих в нем примесей. Существующую здесь ситуацию разъясняет рис. 12.3.

Спонтанная намагниченность ферромагнетика J_s достигается в поле H_A . Для того чтобы перемагнитить образец в противоположном направлении, надо двигаться в плоскости (H, J) вдоль ветви $(A, -H_c, A')$, изображенной на рисунке замкнутой кривой (эту замкнутую кривую называют *петлей гистерезиса*, а само явление — *гистерезисом*). При возвращении в прежнее состояние мы движемся вдоль ветви (A', H_c, A) . Если образец был первоначально размагничен, то намагничению соответствует какая-либо кривая (например, OA), лежащая внутри петли гистерезиса. Намагниченность J_r называют *остаточной*, а поле H_c — *коэрцитивной силой*. Коэрцитивная сила существенно зависит от характера обработки ферромагнитного образца; она равна $\sim 0,3$ А/м в специальных сортах трансформаторного железа и $\sim 1,6 \cdot 10^6$ А/м в специальных постоянных магнитах большой стабильности. Что касается J_s , то она не зависит от характера обработки образца, а определяется типом ферромагнетика и зависит только от температуры.

Значение напряженности поля, при котором достигается намагниченность J_s , зависит, в частности, от ориентации \mathbf{H} относительно кристаллографических осей. Те направления, для которых это поле минимально, называют *направлениями легчайшего намагничивания*. Для железа, например, они совпадают с направлениями ребер кубической решетки этого металла.

Заметим, что когда говорят о большой магнитной восприимчивости

Ферромагнетиков, то имеют в виду большое значение производной dJ/dH на круто поднимающихся частях гистерезисной кривой, но необходимо подчеркнуть, что крутизна кривой намагничения зависит от H и поэтому μ не может служить характеристикой магнитного состояния ферромагнетика.

12.7. Природа ферромагнетизма

Перейдем к вопросу о природе ферромагнетизма. Как мы видели в § 12.2, вещество парамагнитно, если его атомы обладают собственным магнитным моментом, могущим иметь как орбитальное, так и спиновое происхождение. В случае ферромагнетиков основную роль играют спиновые магнитные моменты определенных групп атомных электронов. То, что ферромагнетизм имеет в основном спиновое происхождение, следует из опыта, известного под названием опыта Эйнштейна — де-Гааза. Идея опыта основана на том, что орбитальный магнитный момент атома m_L пропорционален орбитальному моменту импульса $\hbar L$, а спиновый момент m_S — спиновому моменту импульса $\hbar S$, но с удвоенным коэффициентом пропорциональности. Поэтому при намагничении тела (любого парамагнетика!) должен меняться момент импульса его электронов. Если проекция J вдоль какой-либо оси z изменится на ΔJ_z , то претерпит изменение и проекция момента импульса электронов на эту же ось. Для единичного объема эта величина, которую мы обозначим ΔK_z , пропорциональна ΔJ_z , $\Delta K_z/\Delta J_z = 1/g$, где $g = \mu_B/\hbar = e/(2m_e)$, если магнитный момент имеет орбитальное, и $g = 2\mu_B/\hbar = e/m_e$, если магнитный момент имеет спиновое происхождение (если в магнитный момент атома вносят вклад и орбитальный и спиновый моменты, то в это отношение должен входить еще множитель Ланде).

Учтем теперь, что при намагничении тела к нему не прикладываются внешние механические силы. Поэтому при намагничении сохраняется момент импульса тела, а так как изменение намагничения связано с изменением момента импульса электронов, то должен изменяться и момент импульса остова тела — его кристаллической решетки. Это значит, что изменение намагниченности подвешенного образца должно приводить к вращению образца относительно оси, вдоль которой изменяется намагничение, и, наоборот, вращение ферромагнитного стержня должно приводить к его намагничению. Именно этот последний эффект и был обнаружен в точных опытах Гарнета, который показал, что для железного стержня $g = e/m_e$. Этим была экспериментально доказана спиновая природа ферромагнетизма.

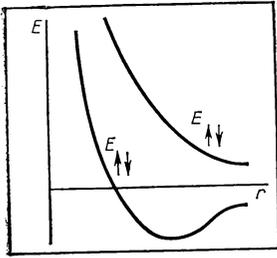
Мы приходим, таким образом, к выводу, что в ферромагнетике спины электронов (более точно, спины определенной группы электронов, так называемых *магнитных электронов*: d -электронов в металлах группы железа и их соединениях и f -электронов в редкоземельных металлах и их соединениях) в ничтожно слабом внешнем магнитном поле в подавляющем большинстве ориентированы по полю, и этой их ориентации не препятствует тепловое движение, если только температура ниже температуры Кюри; при более же высоких температу-

рах преимущественная ориентация исчезает и вещество становится обычным парамагнетиком. В этом случае отдельные магнитные моменты распределены в отсутствие внешнего поля хаотически — установлению порядка в системе магнитных моментов препятствует тепловое движение. Влияние его таково, что при $H \sim 1 \text{ А/м}$ только один момент на 10^9 — 10^{10} моментов ориентируется по полю, в то время как в ферромагнетике при $T < T_C$ такое поле приводит к намагниченности насыщения порядка 10^5 — 10^6 А/м. Можно сказать, что ферромагнетик характеризуется магнитным упорядочением, или сильной корреляцией, между направлениями магнитных моментов отдельных носителей магнетизма (в железе d -электронов). Такая корреляция должна быть следствием определенного взаимодействия между магнитными моментами, но с самого начала ясно, что это взаимодействие не может быть магнитным, так как магнитное взаимодействие двух магнитных моментов очень мало, да к тому же оно имеет место у всех парамагнетиков, а из них далеко не все являются ферромагнетиками.

Так как ферромагнетизм имеет спиновое происхождение, то спины определенной группы электронов имеют в ферромагнетиках при $T \ll T_C$ преимущественно одинаковую ориентацию. Этот эффект является не классическим, а чисто квантовым и связан с тем, что энергия системы электронов существенно зависит от свойств симметрии волновой функции этой системы.

Волновая функция может обладать разной пространственной симметрией, и различным симметриям отвечают различные значения энергии системы электронов. Например, если взять молекулу водорода, в состав которой входят два электрона, то электронная волновая функция может либо не меняться, либо изменять свой знак при перестановке местами электронов. В первом случае — пространственно симметричной волновой функции — энергия молекулы меньше, чем во втором — случае пространственно антисимметричной волновой функции. Но электроны подчиняются принципу Паули, согласно которому волновая функция электронов с учетом их спинов всегда меняет свой знак, если поменять местами электроны и при этом поменять еще их спины. Поэтому в случае молекулы водорода с пространственно симметричной волновой функцией спины электронов имеют разные ориентации, так что суммарный спин молекулы равен нулю. Напротив, в случае антисимметричной пространственной волновой функции спины обоих электронов имеют одинаковую ориентацию, так что суммарный их спин равен единице. Таким образом, вследствие неразличимости электронов и действия принципа Паули возникает ситуация, при которой энергия молекулы водорода оказывается зависящей от значения суммарного спина электронов. Этот эффект носит название *обменного*.

Обменный эффект лежит в основе гомеоплярной химической связи — в системе двух атомов водорода энергия атомов $E_{\uparrow\downarrow}$ при антипараллельной ориентации спинов обоих электронов меньше энергии $E_{\uparrow\uparrow}$, соответствующей параллельной ориентации спинов, причем зависимость $E_{\uparrow\downarrow}$ от расстояния между атомами (т. е. их ядрами) имеет



12.4.
К объяснению обменного эффекта

минимум, которому и соответствует устойчивое состояние молекулы H_2 (рис. 12.4).

Обменный эффект лежит также и в основе ферромагнетизма. Ферромагнитный кристалл можно рассматривать как гигантскую молекулу, причем минимуму ее энергии соответствует параллельная ориентация спинов всех ее электронов (точнее, магнитных электронов), а не антипараллельная, как в случае молекулы водорода. Состояние с минимумом энергии осуществляется при 0 К, а этот минимум соответствует максимальному суммарному спину электронов, т. е. максимально возможному намагничению кристалла. При

достаточно низких температурах $T < T_c$ намагничение еще очень велико. (Это объяснение природы ферромагнетизма было дано независимо Гейзенбергом и Френкелем и Дорфманом.)

Если бы можно было описывать спин классическим вектором, то часть энергии двух электронов, зависящая от их спинов, содержала бы скалярное произведение спинов s_1 и s_2 обоих электронов, т. е. имела бы вид $E_s = -Is_1s_2$, где I — некоторая функция от расстояния между электронами (ее называют *обменным интегралом*). Все дело в знаке этой функции. Если $I > 0$, то параллельной ориентации спинов соответствует меньшая энергия, чем антипараллельной; если же $I < 0$, то меньшей энергии соответствует антипараллельная ориентация спинов. Второй случай осуществляется в молекуле водорода, а первый — в кристалле ферромагнетика. Поэтому состояние, в котором спины всех электронов ферромагнетика расположены параллельно друг другу, обладает минимальной энергией. Это состояние и осуществляется при 0 К. Значение $I = \xi e^2 / (4\pi\epsilon_0 a)$ определяется электростатической энергией взаимодействия электронов двух соседних атомов, здесь a — расстояние между атомами и ξ — числовой коэффициент, который оказывается значительно меньше единицы. Итак,

спиновая корреляция — это квантовый эффект, связанный с зависимостью энергии системы электронов от их суммарного спина, а магнитное упорядочение в ферромагнетиках есть следствие корреляции спинов.

О зависимости энергии системы электронов от их суммарного спина говорят, что она обусловлена обменным взаимодействием, а часть энергии электронов $E_s = -Is_1s_2$ называют *обменной энергией*.

Не следует думать, что спиновая корреляция и связанное с ней магнитное упорядочение всегда носят такой характер, как в ферромагнетиках, т. е. что всегда в магнитоупорядоченных кристаллах минимуму энергии соответствует одинаковая ориентация спинов, а следовательно, и одинаковая ориентация магнитных моментов электронов. Могут быть и более сложные случаи магнитного упорядочения. Например, кристаллы многих солей металлов типа железа представляют собой совокупности двух вставленных друг в друга подрешеток, каждая из которых имеет отличную от нуля намагниченность, но сум-

ма намагниченностей равна нулю. Подобные кристаллы называют *антиферромагнетиками*. Магнитная упорядоченность антиферромагнетиков, так же как и ферромагнетиков, ограничена областью низких температур, а именно температура не должна превосходить некоторой определенной температуры, называемой *температурой Нееля*. Наконец, существуют магнитоупорядоченные кристаллы — *ферриты*, которые состоят из нескольких магнитных подрешеток, каждая со своим намагничением, но векторная сумма всех намагниченностей в отличие от антиферромагнетиков не обращается в нуль в отсутствие внешнего поля. К числу ферритов относятся, например, соли $\text{MnO} \cdot \text{Fe}_2\text{O}_3$, $3\text{Y}_2\text{O}_3 \cdot 5\text{Fe}_2\text{O}_3$.

Обменное взаимодействие электронов в ферромагнетике приводит к параллельной ориентации их спинов, а следовательно, и их магнитных моментов. Но так же действует и внешнее магнитное поле. Поэтому феноменологически обменное взаимодействие можно учесть, если добавить к внешнему магнитному полю $\mathbf{H}^{(e)}$ некоторое эффективное поле \mathbf{H}_W . Это эффективное поле (оно является фиктивным!) называют внутренним *молекулярным полем Вейсса*. Так как в нашем распоряжении имеется только один вектор — намагниченность \mathbf{J} , то естественно предположить, что $\mathbf{H}_W = \nu \mathbf{J}$, где ν — некоторая безразмерная константа. Таким образом, мы будем предполагать, что обменное взаимодействие вместе с внешним полем эквивалентны по своему действию результирующему полю $\mathbf{H}^* = \mathbf{H}^{(e)} + \nu \mathbf{J}$.

Вспомним теперь, что намагниченность парамагнитного газа, парамагнетизм которого имеет спиновую природу, определяется формулой

$$J = n\mu_B \operatorname{th} \frac{\mu_0 \mu_B H^{(e)}}{kT}$$

Так как ферромагнетизм также имеет спиновую природу, то естественно применить эту же формулу к ферромагнетику, заменив в ней $H^{(e)}$ на H^* :

$$J = n\mu_B \operatorname{th} \frac{\mu_0 \mu_B (H^{(e)} + \nu J)}{kT}$$

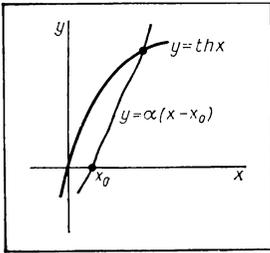
Эта формула представляет собой трансцендентное уравнение для намагниченности $J = J(H^{(e)}, T)$, определяющее ее как функцию $H^{(e)}$ и T . Вводя безразмерные переменные

$$y = \frac{J}{n\mu_B}, \quad x = \frac{\mu_0 \mu_B H^*}{kT}, \quad x_0 = \frac{\mu_0 \mu_B H^{(e)}}{kT},$$

получим

$$y = \operatorname{th} x, \quad y = \frac{kT}{\mu_0 n \mu_B^2 \nu} (x - x_0).$$

Нам нужно найти значения y и x , удовлетворяющие этим уравнениям, для чего удобно изобразить графически выписанную зависимость y от x (рис. 12.5, $\alpha = kT/(\mu_0 n \mu_B^2 \nu)$) и определить точки пересечения ги-



12.5.

Решение уравнения Вейсса

перболической тангенсоиды и прямой. Мы видим, что при $x_0 = 0$ (т. е. при $H^{(e)} = 0$) помимо тривиального решения $y = x = 0$ может быть и нетривиальное решение $y \neq 0$. Для этого необходимо, чтобы прямая $y = \alpha x$ образовывала с осью x угол, меньший угла наклона касательной к тангенсоиде при $x = 0$. Тангенс последнего равен 1, поэтому должно выполняться неравенство $kT/(n\mu_B^2 \nu \mu_0) < 1$, или $T < T_C$, где $T_C = n\mu_B^2 \nu \mu_0/k$.

Таким образом, мы приходим к заключению, что и в отсутствие внешнего магнитного поля ферромагнетик может обладать отличной от нуля намагниченностью — для этого необходимо лишь, чтобы температура была ниже некоторой определенной температуры T_C — температуры Кюри. Но это есть основное свойство ферромагнетика, и оно объясняется, как мы видим, предположением о внутреннем молекулярном поле. Более того, гипотеза молекулярного поля объясняет и другие свойства ферромагнетиков — поведение их при $T > T_C$ и скачок теплоемкости при $T = T_C$. Если $T > T_C$, то намагниченность в отсутствие внешнего поля отсутствует. Поэтому при малом $H^{(e)}$ мала и J и можно считать, что $y = th x \sim x$. Таким образом, мы приходим к уравнению $x = (x - x_0)T/T_C$, откуда $y = x = x_0 T / (T - T_C)$, т. е. $J = \chi H$, где

$$\chi = \frac{\mu_0 n \mu_B^2}{k(T - T_C)}.$$

Но это и есть закон Кюри—Вейсса.

Если $T \rightarrow 0$, то из уравнения Вейсса следует, что $J = n\mu_B [1 - 2\exp(-2T_C/T)]$. Это соотношение, однако, не подтверждается опытом: при $T \ll T_C$ имеет место другое соотношение: $J(0) - J(T) \sim (T/T_C)^{3/2}$. Таким образом, гипотеза молекулярного поля, объясняя основные черты ферромагнетизма, в первую очередь наличие спонтанного намагничивания, не может считаться полностью адекватной в области низких температур.

Оценив в заключение этого параграфа молекулярное поле $H_W = \nu J$. Полагая $n\mu_B \equiv J_\infty \sim 1,4 \cdot 10^5$ А/м и $T_C \sim 10^3$ К, найдем $\nu \sim 10^4$ и, следовательно, $H_W \sim 1,2 \cdot 10^9$ А/м. Мы получили колоссальное значение, которое никак нельзя объяснить магнитным взаимодействием между спиновыми магнитными моментами. Величина H_W соответствует обменному взаимодействию, энергия которого определяется электростатической энергией взаимодействия электронов.

12.8. Обменная энергия и энергия магнитной анизотропии

Взяв сумму обменных энергий всех пар электронов, мы получим общую обменную энергию W_e ферромагнетика:

$$W_e = -\frac{1}{2} \sum_{i,j} I_{ij} \mathbf{s}_i \mathbf{s}_j$$

где \mathbf{s}_i — спин i -го электрона и I_{ij} — обменный интеграл между i -м и j -м электронами (множитель $1/2$ возникает потому, что каждую пару электронов нужно учитывать один раз). Обменный интеграл I для двух электронов, принадлежащих соседним атомам, определяет молекулярное поле Вейсса H_W и температуру Кюри T_C ферромагнетика: $\mu_0 \mu_B H_W \sim kT_C \sim I$ (например, для железа $kT_C \sim 2I$).

Считая спины \mathbf{s}_i классическими векторами, можно переписать обменную энергию ферромагнетика в виде

$$W_e = -\frac{1}{2} s^2 \sum_{i,j} I_{ij} \cos \varphi_{ij},$$

где φ_{ij} — угол между \mathbf{s}_i и \mathbf{s}_j . Так как $I_{ij} > 0$, то минимуму энергии W_e , т. е. энергии основного состояния ферромагнетика, соответствует одинаковая ориентация всех спинов: $\varphi_{ij} = 0$. При малых углах φ_{ij} превышение энергии (обменной) над энергией основного состояния

$$\delta W_e = \frac{1}{4} s^2 \sum_{i,j} I_{ij} \varphi_{ij}^2.$$

Если ориентации спинов электронов неодинаковы, то вектор намагниченности \mathbf{J} также имеет различную ориентацию в разных местах ферромагнетика, хотя модуль его повсюду одинаков. Поэтому изменение вектора \mathbf{J} при $T \ll T_C$ сводится к его вращениям относительно намагниченности \mathbf{J}_0 , соответствующей основному состоянию. При малых поворотах φ , очевидно, $\delta J = |\mathbf{J} - \mathbf{J}_0| \approx J_0 \varphi$, т. е. угол поворота пропорционален отклонению \mathbf{J} относительно \mathbf{J}_0 . Отсюда следует, что φ_{ij} — линейная функция разности векторов $\delta \mathbf{J}(\mathbf{r}_i) - \delta \mathbf{J}(\mathbf{r}_j) = \mathbf{J}(\mathbf{r}_i) - \mathbf{J}(\mathbf{r}_j)$, где $\mathbf{J}(\mathbf{r}_i)$ — намагниченность в точке \mathbf{r}_i . Величина $I_{ij} \equiv I(|\mathbf{r}_i - \mathbf{r}_j|)$ быстро уменьшается с ростом $|\mathbf{r}_i - \mathbf{r}_j|$, поэтому основной вклад в сумму W_e вносят пары электронов соседних атомов. Но в этом случае

$$\mathbf{J}(\mathbf{r}_i) - \mathbf{J}(\mathbf{r}_j) \approx (|\mathbf{r}_i - \mathbf{r}_j| \frac{\partial}{\partial r}) \mathbf{J}(\mathbf{r}),$$

т. е. φ_{ij} — линейная функция пространственных производных вектора $\mathbf{J}(\mathbf{r})$. Поэтому δW_e — квадратичная функция пространственных производных $\mathbf{J}(\mathbf{r})$. Например, отнесенное к единичному объему изменение обменной энергии при возникновении неоднородного намагничивания (его называют *плотностью энергии магнитной неоднородности*) для ферромагнитного кристалла кубической симметрии, скажем Fe, имеет вид

$$\delta W_e = \frac{\alpha \mu_0}{2} \left[\left(\frac{\partial \mathbf{J}}{\partial x} \right)^2 + \left(\frac{\partial \mathbf{J}}{\partial y} \right)^2 + \left(\frac{\partial \mathbf{J}}{\partial z} \right)^2 \right],$$

где α — некоторая константа, пропорциональная I . Ее называют *константой неоднородного обменного взаимодействия* по порядку ве-

личины $\alpha = kT_C a^2 / (\mu_0 \mu_B^2 n)$ (n — число магнитных электронов в единичном объеме).

Как мы видим, параллельной ориентации спинов всех электронов соответствует минимум обменной энергии ферромагнетика. При этом, однако, эта ориентация может быть произвольной, так как с точки зрения обменной энергии все направления эквивалентны — в выражение для W_e входят только скалярные произведения спинов, зависящие от углов между спинами, никакое же избранное направление в W_e не входит. В соответствии с этим минимуму плотности обменной энергии W_e отвечает однородная намагниченность \mathbf{J} , однако направление \mathbf{J} энергией W_e не определяется.

Между тем однородная намагниченность ферромагнитного кристалла \mathbf{J} имеет вполне определенное направление даже в отсутствие внешнего поля (направление легчайшего намагничивания). Это значит, что различным направлениям \mathbf{J} (модуль \mathbf{J} неизменен) должны соответствовать различные значения энергии ферромагнитного кристалла — полной энергии, составной частью которой является обменная энергия, не зависящая от направления \mathbf{J} . Направлению же легчайшего намагничивания должен соответствовать минимум энергии. Чтобы понять это явление — явление магнитной анизотропии, — следует иметь в виду, что электроны; ответственные за ферромагнетизм, движутся во внутрикристаллическом электрическом поле \mathbf{E} . Поэтому на электрон действует магнитное поле $\mathbf{H} = \epsilon_0 [\mathbf{vE}]$, где \mathbf{v} — скорость электрона, и, следовательно, энергия электрона в этом поле составляет $-\mu_0 \mathbf{m}_s \mathbf{H} = 2\epsilon_0 \mu_B \mathbf{s} [\mathbf{vE}] / c^2$, где $\mathbf{m}_s = -2\mu_B \mathbf{s}$ — спиновый магнитный момент электрона. Эта энергия зависит от ориентации спина электрона \mathbf{s} по отношению к внутрикристаллическому электрическому полю \mathbf{E} . Мы должны теперь взять сумму этих энергий для всех электронов. В результате получим добавочную (по отношению к обменной энергии) энергию ферромагнитного кристалла, зависящую от ориентации вектора \mathbf{J} относительно поля \mathbf{E} , а следовательно, и относительно кристаллографических осей ферромагнетика. Эту энергию называют *энергией магнитной анизотропии*.

Оценим значение этой энергии. Заметим, для этого, что $E \sim \sim e/(4\pi\epsilon_0 a^2)$, где a — длина порядка постоянной атомной решетки. Так как $\mu_B = e\hbar/(2m_e)$, то:

$$\mu_0 m_s H \sim \frac{e\hbar}{m_e} \frac{v}{c^2} \frac{e}{4\pi\epsilon_0 a^2}.$$

Но $\hbar/a \sim m_e v$, где v — величина порядка скорости атомного электрона, а $e^2/(4\pi\epsilon_0 a)$ имеет порядок величины обменной энергии электронов. Поэтому энергия магнитной анизотропии составляет по порядку величины $(v/c)^2$ от обменной энергии. При $v \sim 10^8$ м/с отношение обеих энергий составляет 10^{-4} .

Энергия магнитной анизотропии зависит от направления вектора намагниченности \mathbf{J} относительно кристаллографических осей ферромагнетика. Для одноосных кристаллов, например кристаллов кобальта, принадлежащего к гексагональной симметрии, энергия анизотропии является функцией угла θ между \mathbf{J} и гексагональной осью.

Если ввести плотность энергии магнитной анизотропии ω_a , то для кобальта приближенно ее можно записать в виде $\omega_a = K \sin^2\theta$, где K — некоторая величина, не зависящая от угла θ , но зависящая от температуры и от упругих напряжений, действующих в кристалле. При комнатной температуре константа K для кобальта равна $0,4 \cdot 10^6$ Дж·м⁻³. Для кристаллов кубической симметрии, например железа,

$$\omega_a = K_1 (\alpha_1^2 \alpha_2^2 + \alpha_2^2 \alpha_3^2 + \alpha_3^2 \alpha_1^2) + K_2 \alpha_1^2 \alpha_2^2 \alpha_3^2,$$

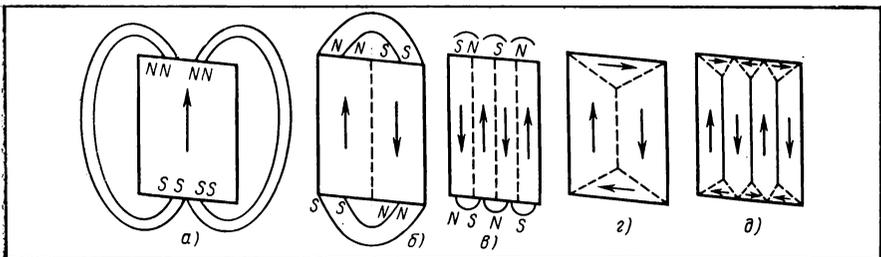
где $\alpha_1, \alpha_2, \alpha_3$ — направляющие косинусы вектора \mathbf{J} относительно ребер куба, а K_1 и K_2 — константы. При комнатной температуре для железа $K_1 = 4,2 \cdot 10^4$ Дж·м⁻³, $K_2 = 1,5 \cdot 10^4$ Дж·м⁻³. Мы видим, что для железа энергия анизотропии на порядок меньше, чем для кобальта. Несмотря на относительно малое значение энергии магнитной анизотропии, она играет существенную роль, так как именно из-за того, что она существует, в кристалле устанавливается определенное направление спонтанной намагниченности. Для железа — это направление любого из ребер куба кристаллической ячейки, для никеля — диагонали граней куба, для кобальта — гексагональная ось.

Того же порядка величины, что и энергия магнитной анизотропии, оказывается магнитная энергия — энергия магнитного поля $\mathbf{H}^{(m)}$, создаваемого самими спиновыми магнитными моментами ферромагнетика. Плотность этой энергии, очевидно, составляет $\omega_m = \mu_0 (H^{(m)})^2 / 2$. Наконец, если ферромагнетик находится во внешнем поле $\mathbf{H}^{(e)}$, то появляется еще энергия взаимодействия его магнитного момента с этим полем, плотность которой есть $\omega_H = -\mu_0 \mathbf{JH}^{(e)}$.

12.9. Доменная структура ферромагнетиков

Ферромагнетик состоит из отдельных соприкасающихся друг с другом макроскопических участков — доменов, каждый из которых обладает намагниченностью J_s , но эти намагниченности имеют разную ориентацию, так что намагниченность всего образца в отсутствие поля равна нулю.

Покажем, что в отсутствие внешнего магнитного поля ферромагнетик с доменной структурой (идея существования доменов принадлежит Вейссу) обладает меньшей энергией, чем ферромагнетик с однородным намагничиванием. Связано это с уменьшением магнитной энергии образца в том случае, если он имеет доменную структуру. Рис. 12.6



12.6. Объяснение доменной структуры ферромагнетика

разъясняет эту ситуацию. Рис. 12.6, *a* соответствует однодоменному образцу. Здесь изображены магнитные силовые линии вне кристалла — они охватывают объем, пропорциональный кубу размеров верхней и нижней поверхности образца. Энергия поля при этом велика, и ясно, что если мы рассмотрим двухдоменную структуру, когда образец состоит из двух половинок с противоположно ориентированными намагниченностями (рис. 12.6, *b*), то магнитная энергия уменьшена по сравнению с энергией однодоменного образца примерно в два раза. В случае четырехдоменной структуры (рис. 12.6, *в*) энергия уменьшится еще больше — примерно в четыре раза, а при структуре, содержащей N доменов, — примерно в N раз. Можно себе представить домены и другой формы и другого взаимного расположения, например домены, изображенные на рис. 12.6, *г*, *д*.

На первый взгляд кажется, что чем больше доменов, тем структура энергетически выгоднее. Однако на самом деле это не так, потому что кроме магнитной энергии нужно учитывать еще обменную энергию, которая минимальна в случае однородного намагничивания. Между тем при переходе от домена к домену происходит резкое изменение направления намагничивания, что должно приводить к большому увеличению обменной энергии. Это видно из выражения $-\frac{1}{2}S^2\sum_{ij}I_{ij}\cos\varphi_{ij}$ для обменной энергии, из которого следует, что резкий поворот спинов энергетически невыгоден и, напротив, энергетически выгоднее, чтобы поворот спинов происходил постепенно.

Отсюда можно заключить, что энергетически выгодно, чтобы между двумя соседними доменами существовал граничный слой (его называют *стенкой Блоха*), в котором угол между направлениями соседних спинов был бы мал и только угол между спинами по обе стороны граничного слоя был бы равен 180° . Действительно, повороту двух спинов друг относительно друга на угол φ соответствует дополнительная обменная энергия $\frac{1}{4}S^2I\varphi^2$ (угол φ считается малым). Если же между этими спинами поместить N спинов, так что поворот между каждыми двумя соседними спинами происходит на угол φ/N , то обменная энергия двух соседних спинов станет равной $\frac{1}{4}S^2I(\varphi/N)^2$, а суммарная обменная энергия N спинов будет пропорциональна $Is^2\varphi^2/N$, т. е. будет в N раз меньше, чем энергия, соответствующая резкому непосредственному повороту спинов на угол φ .

Возникает вопрос: чему равно число N ? Для того чтобы найти его, нужно учесть не только обменную энергию, но и энергию магнитной анизотропии. Вычислим суммарную энергию $\sigma_{\text{ш}}$ (обменную и энергию магнитной анизотропии) стенки Блоха, приходящуюся на ее единичную площадь. Для каждой цепочки из N атомов вдоль толщины стенки обменная энергия равна $\pi^2 Is^2/(4N)$. Таких цепочек приходится на единичную площадь a^{-2} (a — атомное расстояние), поэтому обменная энергия равна $\pi^2 Is^2/(4Na^2)$. Если, далее, энергия магнитной анизотропии, приходящаяся на единичный объем, равна K , то энергия, приходящаяся на единичную площадь слоя, равна KNa . Поэтому

$$\sigma = \frac{\pi^2 Is^2}{4Na^2} + KNa$$

Эта величина является функцией N и достигает, очевидно, минимума при $N = [\pi^2 I_s^2 / (4K a^3)]^{1/2}$. Соответствующая ширина слоя

$$\delta = Na = a \left(\frac{\pi^2 I_s^2}{4K a^3} \right)^{1/2}.$$

Вспоминая, что $I \sim kT_C$ и что $K \sim 10^4$ Дж·м⁻³ (10⁵ эрг/см³), найдем $\delta \sim 300$ а ~ 100 нм. Энергию слоя, приходящуюся на его единичную площадь, можно трактовать как поверхностную энергию ферромагнитного домена:

$$\sigma_W = 2\pi (IK_s^2/a)^{1/2}.$$

Для железа эта величина составляет $\sigma_W \sim 10^{-3}$ Дж·м⁻².

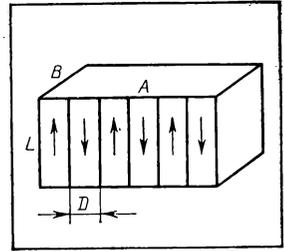
Мы можем теперь снова вернуться к вопросу о доменной структуре ферромагнетика. Чтобы определить, сколько должно возникнуть доменов предполагаемой нами конфигурации, следует найти суммарную энергию ферромагнетика с учетом поверхностной энергии доменов и минимизировать ее относительно числа доменов или параметров, характеризующих их вид и расположение. Рассмотрим, например, доменную структуру, изображенную на рис. 12.7. Число доменов мы найдем из условия минимума энергии ферромагнетика, которая складывается из магнитной энергии и поверхностной энергии доменов. Магнитная энергия ферромагнетика определяется общей формулой

$$W_M = - \frac{\mu_0}{2} \int \mathbf{H}^{(m)} dV,$$

где $\mathbf{H}^{(m)}$ — напряженность собственного магнитного поля, порождаемого намагниченностью \mathbf{J} , и интегрирование совершается по всему объему ферромагнетика. Как видно из формулы, достаточно знать поле $\mathbf{H}^{(m)}$ только внутри ферромагнетика. Здесь оно при большем числе доменов сосредоточено в основном вблизи верхней и нижней границ тела, причем внутри поле $\mathbf{H}^{(m)}$ простирается на глубину порядка D , т. е. порядка ширины домена. Поле в этой приграничной области, близко к $-\mathbf{J}$. Поэтому магнитная энергия имеет порядок $W_M \sim \sim 2\mu_0 J_0^2 DAB$, где A и B — размеры сечения ферромагнетика, перпендикулярного вектору намагничивания. (Коэффициент 2 обусловлен наличием двух границ ферромагнетика — точный расчет дает коэффициент, равный 1,7.)

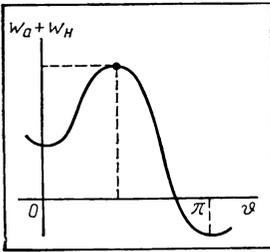
Поверхностная энергия равна энергии блоховской стенки $\sigma_W LB$ (здесь L — высота образца), умноженной на число поверхностей раздела, равное A/D . Таким образом, вся энергия ферромагнетика

$$W = \sigma_M \frac{L}{D} AB + 2\mu_0 J_0^2 DAB.$$



12.7.

К определению числа доменов в образце



12.8. Объяснение коэрцитивной силы

Это выражение нужно теперь минимизировать по D . Условие минимума, как легко видеть, имеет вид $\sigma_W L/D = 2\mu_0 J_0^2 D$, откуда

$$D = \left(\frac{\sigma_W L}{2\mu_0 J_0^2} \right)^{1/2}.$$

При $L = 1$ см мы получим, считая, что $\sigma_W \sim 2 \cdot 10^{-2}$ Дж·м⁻² и $J \sim 10^5$ А·м⁻¹, ширину домена порядка $D \sim 10^{-5}$ м. Напомним, что ширина блоховской стенки δ порядка 10^{-7} м, так что $\delta \ll D$, как и должно быть, чтобы понятие домена имело смысл.

При оптимальной ширине домена энергия доменной структуры $W_D = \omega S$, где $\omega = 2(\mu_0 2J_0^2 \sigma_W / L)^{1/2}$ и $S = AB$. При $L = 1$ см имеем $\omega \sim 7$ Дж·м⁻², в то время как в однодоменном кристалле энергия ферромагнетика была бы равна примерно $\mu_0 J^2 LS$, что при $J \sim 10^5$ А·м⁻¹ и $L = 1$ см в 10^3 раз больше энергии W_D . Мы видим, таким образом, что образование доменов приводит к значительному уменьшению энергии ферромагнетика.

Обратим внимание на то, что размеры домена D зависят от размеров образца L , причем $D \sim L^{1/2}$. При достаточно малых толщинах образца, в случае тонких пленок с $L \sim 10^{-8}$ м, D имеет порядок L . В этом случае образование доменов энергетически невыгодно и они не образуются.

Для экспериментального наблюдения доменов используется метод нанесения на образец тонкого слоя коллоидной суспензии очень мелких ферромагнитных частиц. Эти частицы располагаются вдоль границ между доменами; наблюдая их через микроскоп, можно видеть границы доменов.

Доменная структура ферромагнетика ответственна за возникновение гистерезиса. Дело в том, что в ненамагниченном образце магнитные моменты доменов с противоположными ориентациями спонтанной намагниченности взаимно компенсируются: при включении поля доменные стенки начинают перемещаться таким образом, что домены, намагниченные вдоль поля, растут за счет доменов, намагниченных против поля. При перемещении доменная стенка «цепляется» за примеси и другие дефекты кристаллической структуры, именно поэтому кривая намагничения сильно меняется от образца к образцу, завися от его механической обработки и наличия примесей.

В сильных полях главную роль играет не перемещение доменной стенки параллельно самой себе, а вращение вектора спонтанной намагниченности. Остановимся на этом процессе несколько подробнее. От ориентации вектора \mathbf{J} зависят два слагаемых в выражении для плотности энергии: энергия магнитной анизотропии $w_a = K \sin^2 \theta$ и магнитная энергия во внешнем поле $w_H = \mu_0 J_s H^{(e)} \cos \theta$ (θ — угол между первоначальным направлением намагниченности и полем $\mathbf{H}^{(e)}$, которое направлено вдоль легкой оси). Функция $w(\theta) \equiv w_a +$

$+ \omega_H$ имеет (если $H^{(e)} < 2K/(J_s \mu_0)$) вид, представленный на рис. 12.8. Поэтому хотя при $\theta = \pi$ энергия меньше, чем при $\theta = 0$, однако при перемагничении необходимо преодолеть потенциальный барьер, соответствующий $\cos\theta = \mu_0 J_s H^{(e)}/(2K)$.

ОСНОВНЫЕ ФОРМУЛЫ

Закон Ома для магнитной цепи

$$\Phi = \frac{\mathcal{E}_M}{R_M}$$

Сила, действующая на единичный объем магнетика

$$F = \frac{\chi}{2} \text{grad } H^2$$

Орбитальный магнитный момент атома

$$m_L = \frac{e\hbar}{2m_e} \mathbf{L}$$

Спиновый магнитный момент атома

$$m_S = \frac{e\hbar}{m_e} \mathbf{S}$$

Формула Ланжевена

$$M = nm_L L\left(\frac{\mu_0 m_L H}{kT}\right),$$

$$L(\xi) = \text{cth } \xi - \frac{1}{\xi}$$

Закон Кюри для магнитной восприимчивости парамагнетика

$$\chi = \frac{\text{const}}{T}$$

Парамагнитная восприимчивость вырожденного электронного газа

$$\chi_p = 2\mu_0 \mu_B^2 \nu(\epsilon_0)$$

Диамагнитная восприимчивость газа

$$\chi_d = \frac{-\mu_0 e^2}{6m_e} Z \langle r^2 \rangle n$$

Частота Лармора

$$\omega_L = -\frac{eB}{2m_e}$$

Эффект Холла

$$\mathbf{E} = \rho \mathbf{j} + R [\mathbf{Bj}]$$

Закон Кюри — Вейсса для магнитной восприимчивости ферромагнетика

$$\chi = \frac{C}{T_c - T}$$

Закон Блоха для спонтанной намагниченности ферромагнетика при низких температурах

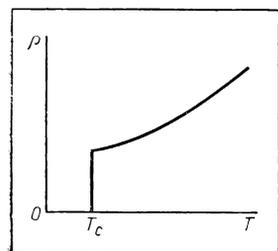
$$1 - \frac{J_s}{J_\infty} \sim \left(\frac{T}{T_c}\right)^{3/2}$$

Глава 13. СВЕРХПРОВОДНИКИ

13.1. Явление сверхпроводимости

Как мы видели в гл. 9, удельное сопротивление металлов непрерывно уменьшается с понижением температуры, и если бы кристаллическая решетка была идеальной (т. е. не содержала бы примесей и дефектов), то сопротивление обращалось бы в нуль при 0 К. Однако такой ход изменения удельного сопротивления характерен далеко не для всех металлов. У ряда металлов, например у свинца, ртути, олова, удельное сопротивление, уменьшаясь с понижением температуры, при некоторой, характерной для каждого металла, температуре внезапно, скачком обращается в нуль и остается нулем вплоть до 0 К (рис. 13.1). Такие металлы называют *сверхпроводниками*, а само явление — *сверхпроводимостью* (оно было открыто Камерлинг-Оннесом в 1911 г.). Сверхпроводимость наблюдается у 27 металлов и более чем у 1 тыс. металлических соединений и сплавов. Температура T_c , при которой она наступает (ее называют *критической*), изменяется у известных сверхпроводников от 0,01 до 20 К. Наиболее низкой критической температурой обладает вольфрам, для которого $T_c = 0,1$ К. Приведем еще несколько значений T_c : 0,85 К для цинка; 1,2 К для алюминия; 7,2 К для свинца, 4,15 К для ртути, 9,4 К для ниобия, 5,8 К для соединения Ti_3Sn и 18,0 К для Nb_3Sn .

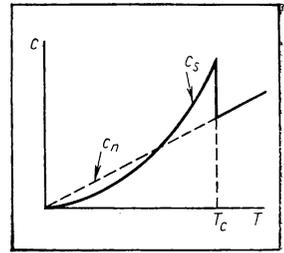
При $T > T_c$ сверхпроводящие металлы обладают электрическим сопротивлением и оно изменяется с температурой так же, как и у обычных, или, как их называют, *нормальных*, металлов, при критической же температуре сопротивление скачком обращается в нуль. Поэтому естественно ожидать, что отсутствие сопротивления есть проявление некоторого нового металлического состояния — сверхпроводящего, которое отличается от нормального состояния. И действительно, нормальное и сверхпроводящее состояния представляют собой две различные фазы металла, при понижении же температуры в точке $T = T_c$ происходит фазовый переход из нормальной в сверхпроводящую фазу.



Этот переход не связан, однако, с изменением кристаллической структуры и не сопровождается ни выделением, ни поглощением теплоты. Это значит, что переход из нормальной в сверхпроводящую фазу не есть переход первого рода. Он представляет собой фазовый переход второго рода, при котором не происходит скачка внутренней энергии, а скачок испытывает теплоемкость, причем теплоемкость электронная. Для нормальной металла, как мы видели, электронная теплоемкость (обозначим ее c_n) изменяется с температурой линейно: $c_n \sim kT/T_F$; в сверх-

13.1.
Зависимость удельного сопротивления сверхпроводника от температуры

проводящей же фазе электронная теплоемкость (обозначим ее c_s) изменяется с температурой экспоненциально: $c_s \sim ke^{-\Delta_s/(kT)}$, где величина $\Delta_s = 1,75 kT_c$ пропорциональна критической температуре. При критической температуре (ее называют также *температурой сверхпроводящего перехода*) возникает скачок теплоемкости, как это показано на рис. 13.2.



13.2. Зависимость теплоемкости сверхпроводника от температуры

Чем же отличаются друг от друга обе металлических фазы — нормальная и сверхпроводящая? Отличие заключается в том, что в сверхпроводящей фазе некоторая доля электронов проводимости как бы сконденсирована в «сверхтекучую жидкость», которая как единое целое может двигаться через кристалл. Движение это происходит беспрепятственно, т. е. без трения, и именно по этой причине в сверхпроводящей фазе отсутствует электрическое сопротивление. Это значит, что электрический ток в сверхпроводнике может течь, не затухая, сколь угодно долго без какого-либо внешнего источника э. д. с. При 0 К все электроны проводимости входят в состав сверхтекучей жидкости. С ростом температуры доля электронов, входящих в состав сверхтекучей жидкости, уменьшается, и наконец, при критической температуре обращается в нуль.

Разъясним теперь причину, по которой возникает сверхтекучая электронная жидкость. Двигаясь в периодическом поле кристаллической решетки, электроны, согласно закону Кулона, отталкиваются друг от друга. Но помимо отталкивания между электронами действуют еще и специфические силы притяжения. Дело в том, что электрон, сталкиваясь с ионами, может либо возбуждать, либо поглощать колебания решетки — фононы. Представим себе теперь два электрона, из которых один испускает фонон, а другой поглощает его. Такой обмен фононами связывает электроны, в результате чего возникает некоторое взаимодействие, приводящее к притяжению между электронами. Таким образом,

наряду с кулоновским отталкиванием между электронами проводимости существует еще и притяжение, обусловленное обменом фононами.

Все дело теперь в том, какие силы больше — отталкивания или притяжения. Если силы притяжения превосходят силы отталкивания, причем как угодно мало, то возникает замечательное явление — образуется связанное состояние двух электронов — нечто, подобное двухэлектронной молекуле. Такое образование является, однако, следствием не только превалирования сил притяжения над силами отталкивания, но еще и того в высшей степени важного обстоятельства, что электроны подчиняются статистике Ферми — Дирака. (Объяснение сверхпроводимости было дано Бардином, Купером и Шриффером в 1956 г., т. е. почти через полвека после открытия этого явления.)

Говорить о связанной двухэлектронной молекуле можно, впрочем, лишь условно. Дело в том, что размеры такой связанной пары электронов очень велики, в 10^2 — 10^4 раз больше постоянной решетки. Поэтому размеры пары оказываются во много раз больше среднего расстояния между парами. Тем не менее существует сильная корреляция между состояниями электронов внутри пары, и мы можем исходить из представления о связанных электронных парах. Эти пары и образуют сверхтекучую электронную жидкость. Все они находятся в одинаковом состоянии и движутся через кристалл как единое целое. Изменить это состояние, т. е. возбудить это единое целое, не так легко, потому что для этого нужно разорвать по крайней мере одну пару, т. е. нужно затратить энергию, равную энергии связи пары. Эта энергия в точности равна $2\Delta_s$. Поэтому до тех пор, пока возмущающие воздействия не приносят энергии, равной или большей $2\Delta_s$, состояние сверхпроводящей жидкости неизменно и устойчиво и она ведет себя, как сверхтекучая жидкость, двигающаяся без трения. Ее электрическое сопротивление, как мы уже говорили, равно нулю.

При 0 К все электроны проводимости спарены, т. е. весь электронный газ входит в состав сверхтекучей жидкости. При конечной температуре всегда есть конечная вероятность (пропорциональная $e^{-2\Delta_s/(kT)}$) того, что пара будет разрушена на отдельные электроны. Такие не связанные попарно электроны образуют нормальный электронный газ, точнее нормальную электронную жидкость, движущуюся по кристаллу обычным образом, т. е. с трением. С повышением температуры уменьшается количество сверхтекучей и увеличивается количество нормальной электронной жидкости. При $T = T_c$ количество сверхтекучей жидкости обращается в нуль и все электроны становятся нормальными.

13.2. Диамагнетизм сверхпроводников

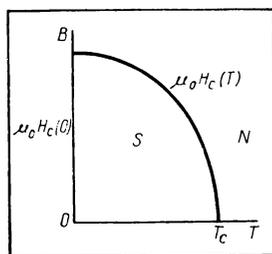
Сверхпроводники характеризуются не только отсутствием электрического сопротивления. Они обладают еще другим замечательным свойством: магнитное поле внутри сверхпроводников равно нулю ($\mathbf{B} = 0$). Поэтому сверхпроводники можно назвать абсолютными диамагнетиками. Равенство нулю магнитного поля в сверхпроводниках тесно связано с отсутствием электрического сопротивления. Действительно, представим себе, что сверхпроводник вносится во внешнее магнитное поле. Тогда если бы магнитное поле проникло в сверхпроводник, то по крайней мере в начальный момент времени производная от поля по времени была бы положительной, т. е. \mathbf{B} отличалась бы от нуля. Но в таком случае, согласно закону электромагнитной индукции, в сверхпроводнике возникло бы электрическое поле \mathbf{E} . Так как электроны в сверхпроводнике движутся, не испытывая сопротивления, то сколь угодно малое поле \mathbf{E} вызвало бы сколь угодно большой ток. Индуцируемый ток, как мы знаем, направлен всегда так, чтобы ослабить причину, вызывавшую его, т. е. так, чтобы собственное магнитное поле тока было направлено против внешнего магнитного поля. При этом ток нарастает до тех пор, пока не

произойдет полной компенсации полей, т. е. пока магнитное поле внутри проводника не обратится в нуль. Ток же в сверхпроводнике обязательно сохранится — без источника внешней э. д. с., так как его протекание необходимо для выполнения закона $\mathbf{B} = 0$.

Таким образом, отсутствие сопротивления у сверхпроводников (т. е. равенство нулю электрического поля, так как $\mathbf{E} = \mathbf{j}/\sigma$) является, по существу, следствием их магнитных свойств. Здесь необходимо подчеркнуть, что, в то время как равенство $\mathbf{E} = 0$ вытекает из равенства $\mathbf{B} = 0$, равенство $\mathbf{B} = 0$ не вытекает из равенства $\mathbf{E} = 0$: в идеальном проводнике с $\sigma = \infty$ и $\mathbf{E} = 0$ магнитное поле может существовать.

Итак, внешнее магнитное поле не может проникать в сверхпроводник, точнее говоря, в глубь сверхпроводника, так как в действительности, как мы скоро увидим, в тонком поверхностном слое сверхпроводника оно может быть отлично от нуля.

До сих пор мы ничего не говорили о значении внешнего магнитного поля. Между тем существенным является, каково это поле. Дело в том, что только достаточно слабые поля не проникают внутрь сверхпроводников, сильное же поле разрушает сверхпроводимость, после чего, естественно, проникает в металл, перешедший в нормальное состояние. Граничное поле H_c , начиная с которого разрушается сверхпроводимость (оно носит название *критического*), зависит от температуры T и уменьшается с ее повышением. При температуре перехода T_c поле H_c обращается в нуль. Ход зависимости $\mu_0 H_c$ от температуры изображен на рис. 13.3. Кривая $\mu_0 H_c(T)$ разделяет плоскость (B, T) на две части, из которых S соответствует сверхпроводящему, а N — нормальному состоянию. Можно сказать, что S соответствует сверхпроводящей, а N — нормальной фазе и что вдоль кривой $B = \mu_0 H_c(T)$ обе фазы могут сосуществовать. Приблизительно зависимость H_c от T квадратична: $H_c(T) = H_c(0)(1 - T^2/T_c^2)$. Значения T_c и $H_c(0)$ для некоторых металлов приведены в таблице:

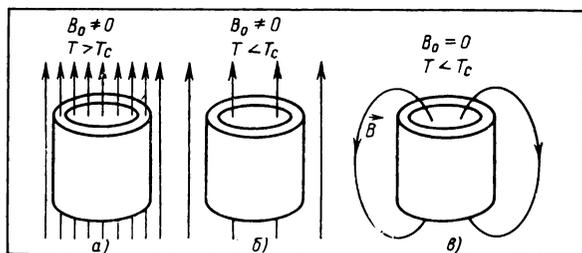


13.3. Зависимость критического поля от температуры

	Zn	Cd	Al	Ga	In	Ti	Sn	Pb
$H_c(0)$, Гс	53	30	99	51	283	162	306	803
T_c , К	0,88	0,56	1,19	1,09	3,41	1,37	3,72	7,18

Мы видим, что металлам с большей температурой сверхпроводящего перехода соответствует большее критическое магнитное поле.

Не следует думать, что соотношение $B = 0$ означает лишь невозможность проникновения достаточно слабого внешнего магнитного



13.4.
Эффект Мейсснера

поля в глубь сверхпроводника, т. е. проводника, перешедшего в состояние сверхпроводимости до включения внешнего магнитного поля. Можно поставить опыт и иначе, а именно взять нормальный проводник, создать в нем магнитное поле и затем, охладив проводник, перевести его в сверхпроводящее состояние. Тогда оказывается, что магнитное поле будет вытолкнуто из сверхпроводника, т. е. внутри сверхпроводника снова будет выполняться соотношение $B = 0$ (выталкивание магнитного поля из сверхпроводника носит название *эффекта Мейсснера*).

Разъясним, как проявляется этот эффект, на примере сверхпроводящего кольца. Поместим металлическое кольцо в нормальном состоянии во внешнее магнитное поле B_0 (рис. 13.4, а). Охладим затем кольцо до температуры ниже критической. Тогда магнитное поле будет вытолкнуто из толщи кольца и силовые линии будут иметь вид, изображенный на рис. 13.4, б. Они проходят вне кольца и пронизывают также отверстие кольца. Выключим, наконец, внешнее поле, не изменяя температуры кольца. Легко видеть, что при этом поток Φ магнитного поля через отверстие в кольце не может измениться. Действительно, вспомним, что производная по времени от потока Φ равна циркуляции электрического поля по контуру, пронизываемому потоком Φ . Расположим контур внутри сверхпроводящего кольца. Так как внутри кольца $B = 0$, то поток, пронизывающий контур, совпадает с потоком Φ . С другой стороны, электрическое поле на контуре, целиком находящемся в толще сверхпроводника, равно нулю. Поэтому производная от Φ по времени равна нулю, т. е. Φ является константой. Иными словами, поток магнитного поля через отверстие в сверхпроводящем кольце не изменится при выключении внешнего поля. Это значит, что магнитные силовые линии, которые шли через отверстие при включенном внешнем поле, «заморозятся» после выключения внешнего поля. Но «замороженные» линии должны быть замкнутыми. Поэтому картина поля после выключения внешнего поля имеет вид, изображенный на рис. 13.4, в.

Замкнутые силовые линии всегда связаны с токами. Поэтому по сверхпроводящему кольцу течет электрический ток, обеспечивающий постоянство магнитного потока через отверстие в сверхпроводящем кольце. Ток этот, как мы увидим далее, течет только в узком слое вблизи поверхности кольца. Замечательным является тот факт, что

магнитный поток Φ через отверстие в сверхпроводящем кольце может принимать только дискретные значения, а именно

$$\Phi = n\Phi_0, \quad \Phi_0 = \frac{\pi\hbar}{e} \approx 2,068 \cdot 10^{-15} \text{ Дж} \cdot \text{с} \cdot \text{Кл}^{-1} \quad (n = 0, 1, \dots).$$

13.3. Термодинамика сверхпроводников

Как мы видели, кривая $B = \mu_0 H_c(T)$ разделяет плоскость (B, T) на две части, из которых одна соответствует сверхпроводящей, а другая — нормальной фазе металла. Если поле и температура связаны между собой соотношением $B = \mu_0 H_c(T)$, то должны сосуществовать обе фазы, т. е. кривая $B = \mu_0 H_c(T)$ представляет собой кривую равновесия сверхпроводящей и нормальной фаз. По этой причине значение критического магнитного поля является важнейшей характеристикой сверхпроводника.

Покажем прежде всего, что разность свободных энергий нормальной (F_n) и сверхпроводящей (F_s) фаз определяется критическим полем:

$$F_n(T) - F_s(T) = \mu_0 H_c^2(T)/2$$

($F_{n,s}$ — плотности свободной энергии). Представим себе с этой целью, что цилиндр, сделанный из металла, способного находиться в сверхпроводящем состоянии, помещается в соленоид, по которому протекает постоянный ток I . Тогда свободная энергия системы цилиндр плюс поле, отнесенная к единичному объему, составляет $F^{(1)} = F_n + B^2/(2\mu_0)$, где $B = \mu_0 In$ — индукция в соленоиде (n — число витков на единичной длине соленоида). Пусть теперь цилиндр переходит в сверхпроводящее состояние. Тогда плотность свободной энергии системы $F^{(2)} = F_s$, так как поле в сверхпроводящем цилиндре равно нулю. Предположим, что переход из нормального в сверхпроводящее состояние происходит при постоянной температуре, так что $B = \mu_0 H_c(T)$. Кроме того, будем считать, что сила тока в катушке не изменяется. Из условия постоянства температуры следует, что $F^{(1)} - F^{(2)} = A_{1 \rightarrow 2}$, где $A_{1 \rightarrow 2}$ — работа, затрачиваемая при переходе $1 \rightarrow 2$, т. е. работа, совершаемая источником э. д. с., включенным в обмотку соленоида (работа должна быть отнесена к единичному объему). Так как ток поддерживается неизменным, то работа $A_{1 \rightarrow 2} = I(\Phi_1 - \Phi_2)$, где Φ_1 и Φ_2 — магнитные потоки, пронизывающие обмотку соленоида в состояниях 1 и 2. Поток в состоянии 2 равен нулю, так что $A_{1 \rightarrow 2} = I\Phi_1$. Учитывая, что $\mu_0 In = B$ и всю обмотку соленоида пронизывает поток $S B n l$ (l — длина соленоида), получим $A_{1 \rightarrow 2} = \mu_0 H_c^2$. Поэтому

$$F^{(1)} - F^{(2)} = F_n + \frac{\mu_0 H_c^2}{2} - F_s = \mu_0 H_c^2,$$

откуда и вытекает написанное выше соотношение для разности свободных энергий F_n и F_s .

С помощью этого соотношения можно найти разность энтропий нормальной и сверхпроводящей фаз, находящихся в равновесии. Ис-

пользуя общее соотношение $S = -(\partial F/\partial T)_B$ (индекс B означает, что дифференцировать следует при $B' = \text{const}$), связывающее энтропию S со свободной энергией F , получим

$$S_n(T) - S_s(T) = -\mu_0 H_c(T) \frac{dH_c(T)}{dT}$$

(энтропии S_n и S_s относятся к единичному объему). Как видно из рис. 13.3, $dH_c/dT < 0$, поэтому $S_n > S_s$. Этого и следовало ожидать, так как сверхпроводящая фаза отличается большим «порядком», чем нормальная. Умножим изменение энтропии на температуру, найдем скрытую теплоту перехода $Q_{s \rightarrow n}$ из сверхпроводящей в нормальную фазу:

$$Q_{s \rightarrow n} = T(S_n - S_s) = -\mu_0 T H_c \frac{dH_c}{dT}.$$

Правая часть этого равенства положительна вдоль всей кривой перехода $B = \mu_0 H_c(T)$, кроме точки $T = T_c$, где обращается в нуль критическое поле. Это значит, что

при переходе из сверхпроводящего состояния в нормальное в присутствии магнитного поля поглощается теплота, т. е. такой переход является фазовым переходом первого рода. В отсутствие же поля сверхпроводящий переход (при $T = T_c$) не сопровождается ни поглощением, ни выделением теплоты, т. е. является фазовым переходом второго рода.

Если при переходе из сверхпроводящего в нормальное состояние проводник теплоизолирован, то его температура понижается.

Как мы уже говорили, при переходе из нормального в сверхпроводящее состояние теплоемкость металла испытывает скачок. Этот скачок также может быть связан с критическим магнитным полем. Для этого следует лишь вспомнить соотношение $C = dQ/dT = T(dS/dT)$, связывающее теплоемкость C с энтропией S . Учитывая, что $S_n - S_s = -\mu_0 H_c dH_c/dT$, найдем скачок теплоемкости в точке перехода:

$$C_n - C_s = \left[T \frac{d}{dT} (S_n - S_s) \right]_{T=T_c} = -\mu_0 \left[T \left(\frac{dH_c}{dT} \right)^2 \right]_{T=T_c}.$$

Это соотношение (формула Рутгерса) подтверждается экспериментально с точностью, превышающей 1%.

13.4. Глубина проникновения магнитного поля

Хотя в толще сверхпроводника $B = 0$, не следует думать, что если сверхпроводник вносится во внешнее магнитное поле, то на границе сверхпроводника поле испытывает резкий скачок. В действительности поле по мере продвижения в глубь сверхпроводника уменьшается постепенно, но уже на сравнительно небольшой глубине становится очень малым. Эффективная глубина проникновения магнитного поля в сверхпроводник составляет примерно 50 нм для таких металлов, как Al, Sn, и достигает 200 нм для переходных металлов (Nb, V, Ta).

Покажем теперь, как происходит затухание поля в сверхпроводнике. Но предварительно выясним, чем определяется электрический ток в сверхпроводнике. В нормальном проводнике плотность электрического тока пропорциональна электрическому полю. Ясно, что в сверхпроводнике такой связи быть не может, так как электроны в сверхпроводнике движутся без сопротивления, поэтому даже бесконечно малое электрическое поле вызвало бы ток бесконечной силы. Это значит, что в сверхпроводнике плотность тока должна быть связана не с электрическим, а с магнитным полем. Но непосредственной линейной связи между плотностью электрического тока \mathbf{j} и магнитным полем \mathbf{B} также не может быть, так как эти векторы имеют разную природу — вектор \mathbf{j} является полярным, а вектор \mathbf{B} — аксиальным. Полярным вектором является, как мы знаем, векторный потенциал \mathbf{A} , связанный с \mathbf{B} соотношением $\mathbf{B} = \text{rot } \mathbf{A}$. Поэтому в сверхпроводниках следует ожидать связи между величинами \mathbf{j} и \mathbf{A} — связи, естественно, линейной. Но непосредственный физический смысл имеет не векторный потенциал, а магнитное поле. Это находит свое выражение в том, что к векторному потенциалу можно прибавить градиент любой скалярной функции, магнитное же поле, равное ротору векторного потенциала, при этом не изменяется. Поэтому можно с равным успехом пользоваться как векторным потенциалом \mathbf{A} , так и векторным потенциалом \mathbf{A}' , связанным с \mathbf{A} соотношением $\mathbf{A}' = \mathbf{A} + \text{grad } f$, где f — любая скалярная функция координат.

О какой же связи между вполне определенной плотностью сверхпроводящего тока \mathbf{j} и не полностью определенным векторным потенциалом \mathbf{A} может идти речь? Ответ гласит, что сверхпроводящий ток определяется не только векторным потенциалом, но еще одной удивительной величиной — фазой *потенциала спаривания* электронов в сверхпроводнике. Мы не будем давать здесь строгого определения потенциала спаривания. Заметим лишь, что существует такая величина — потенциал спаривания $\Delta(\mathbf{r})$, который определяет энергию связи пары электронов в сверхпроводнике. Эту величину можно записать в виде $\Delta(\mathbf{r}) = |\Delta(\mathbf{r})| \exp(i\varphi(\mathbf{r}))$, где $\varphi(\mathbf{r})$ и есть фаза потенциала спаривания.

Если перейти от векторного потенциала \mathbf{A} к векторному потенциалу $\mathbf{A}' = \mathbf{A} + \text{grad } f$, то фаза потенциала спаривания изменится, но таким образом, что величина $\mathbf{u} = (\hbar/2)\text{grad } \varphi - e\mathbf{A}$ останется неизменной, т. е.

$$(\hbar/2) \text{grad } \varphi' - e\mathbf{A}' = (\hbar/2) \text{grad } \varphi - e\mathbf{A}.$$

Поэтому естественно предположить, что именно этой величине и пропорциональна плотность сверхпроводящего тока \mathbf{j} . Так как плотность тока должна быть пропорциональна плотности носителей заряда и их заряду и обратно пропорциональна их массе, то мы приходим к соотношению

$$\mathbf{j} = \frac{en_s}{m_e} \left(\frac{\hbar}{2} \text{grad } \varphi - e\mathbf{A} \right),$$

где n_s — плотность сверхпроводящих электронов, e и m_e — заряд и масса электрона. Это соотношение представляет собой простейшую

линейную связь между векторами \mathbf{j} и \mathbf{A} . Простота заключается в том, что эта связь является локальной, т. е. векторы \mathbf{j} и \mathbf{A} , зависящие от положения, связываются между собой в одной и той же точке сверхпроводника.

Однако локальность связи сама по себе еще ниоткуда не вытекает. Из общих соображений можно лишь утверждать, что при достаточно медленном изменении поля в пространстве связь должна быть локальной, вообще же она может быть и нелокальной, т. е. ток в некоторой точке может определяться не только значением векторного потенциала в этой же точке, но и его значениями в соседних точках. Чтобы разобраться в этой ситуации, нужно понять, что означает медленность изменения поля. С этой целью мы сперва сделаем все выводы, вытекающие из предположения о локальной связи, и лишь потом вернемся к вопросу о нелокальности связи между векторами \mathbf{j} и \mathbf{A} .

Магнитное поле \mathbf{B} в сверхпроводнике связано с плотностью тока \mathbf{j} известным нам общим уравнением $\text{rot } \mathbf{B} = \mu_0 \mathbf{j}$, в которое нужно подставить написанное выше выражение для \mathbf{j} . Так как в это выражение входит фаза потенциала спаривания, то желательно избавиться от нее. Это легко сделать, замечая, что $\text{rot grad } \varphi = 0$. Поэтому следует взять ротор от обеих частей написанного уравнения $\text{rot rot } \mathbf{B} = \text{rot } \mu_0 \mathbf{j}$. Так как $\text{rot } \mathbf{A} = \mathbf{B}$, то мы приходим к уравнению, содержащему только магнитное поле:

$$\text{rot rot } \mathbf{B} = - \frac{\mu_0 e^2 n_s}{m_e} \mathbf{B}.$$

Но $\text{rot rot } \mathbf{B} = \text{grad div } \mathbf{B} - \Delta \mathbf{B}$, а $\text{div } \mathbf{B} = 0$. Поэтому уравнение для \mathbf{B} приобретает вид

$$\Delta \mathbf{B} = \frac{1}{\lambda_L^2} \mathbf{B}, \quad \lambda_L = \sqrt{\frac{m_e}{\mu_0 e^2 n_s}}.$$

Это уравнение было впервые установлено Ф. Лондоном и Г. Лондоном, его называют уравнением Лондонов.

Легко найти решение уравнения Лондонов для того случая, когда сверхпроводник занимает все пространство $z > 0$ и магнитное поле направлено вдоль границы сверхпроводника по оси x . Уравнение тогда приобретает вид $d^2 B/dz^2 = B/\lambda_L^2$ и имеет решение $B(z) = B(0)e^{-z/\lambda_L}$, где $B(0)$ — значение поля на границе сверхпроводника. (Уравнение допускает формально также решение, зависящее от z как $\exp(z/\lambda_L)$, но это решение должно быть отброшено, так как оно приводит к бесконечному полю при $z \rightarrow \infty$.)

Итак, мы пришли к выводу, что поле в сверхпроводнике экспоненциально убывает от границы и на глубине $z = \lambda_L$ уменьшается в e раз. Поэтому величину λ_L можно назвать *глубиной проникновения поля*. Так как ток пропорционален ротору поля, то на такую же глубину λ_L проникает и ток. Если поле направлено вдоль оси x , то плотность тока направлена вдоль оси y , причем $d\mathbf{B}/dz = \mu_0 \mathbf{j}$, откуда

$$\mathbf{j} = - \frac{B(0)}{\mu_0 \lambda_L} e^{-z/\lambda_L}.$$

13.5. Два рода сверхпроводников

Теперь мы можем вернуться к вопросу о том, что значит медленность изменения поля. Поле и ток в сверхпроводнике существенно изменяются на расстояниях порядка λ_L . Если эта величина (лондоновская глубина проникновения) велика по сравнению с другими характерными длинами, то поле можно считать изменяющимся медленно, если же длина λ_L мала по сравнению с другими длинами, возникающими в задаче, то считать поле изменяющимся медленно нельзя. Поэтому возникает вопрос: какие еще длины играют роль в проблеме сверхпроводимости?

Мы будем рассматривать только массивные сверхпроводники, размеры которых значительно больше лондоновской глубины проникновения λ_L . Тогда единственной величиной, имеющей размерность длины и играющей роль в проблеме сверхпроводимости, является размер пары электронов, связь которых между собой обусловлена обменом фононов. Эта величина, которую мы обозначим ξ , тем меньше, чем больше энергия связи пары, т. е. потенциал спаривания $|\Delta|$, и тем больше, чем больше граничная фермиевская скорость v_F сверхпроводника. Именно: оказывается, что

$$\xi = \hbar v_F / (\pi |\Delta|).$$

Понять происхождение этой формулы можно рассматривая, например, атом водорода. В этом случае уравнение движения электрона (если описывать движение классически) имеет вид $m_e v^2 / r = e^2 / (4\pi\epsilon_0 r^2)$, где v и r — скорость электрона и радиус его орбиты. Присоединяя сюда условие квантования $m_e v r = \hbar$ и учитывая, что энергия электрона

$$\Delta = \frac{m_e v^2}{2} - \frac{e^2}{4\pi\epsilon_0 r} = - \frac{e^2}{8\pi\epsilon_0 r},$$

получим

$$r = \frac{4\pi\hbar\epsilon_0}{m_e e^2}, \quad v = \frac{e^2}{4\pi\epsilon_0\hbar}, \quad \Delta = - \frac{m_e e^4}{2(4\pi\epsilon_0\hbar)^2},$$

откуда $r = \hbar v / (2|\Delta|)$. Заменяя в этом выражении v на v_F , мы и получим приведенную выше формулу для ξ .

Итак, мы должны сравнить величины λ_L и ξ . Если $\lambda_L \gg \xi$, то можно говорить о медленном изменении поля, если же $\lambda_L \leq \xi$, то поле нельзя считать медленно меняющимся. В первом случае ($\lambda_L \gg \xi$) можно исходить из предположения о локальности связи между векторами \mathbf{j} и \mathbf{A} . Поэтому в этом случае справедливо уравнение Лондонов. Если же $\lambda_L \leq \xi$, то уравнение Лондонов не справедливо и вместо него должно быть получено другое уравнение, основанное на нелокальной связи между током и векторным потенциалом. Конкретно, если взять простые (непереходные) металлы (In, Pb, Sn), то для них $\lambda_L \approx 40 \div 70$ нм, а $\xi \approx 10^3$ нм, т. е. к таким металлам уравнение Лондонов неприменимо и для описания эффекта Мейснера требуется другой подход, основанный на нелокальной связи \mathbf{j} с \mathbf{A} . Напротив, в переходных металлах (Nb, V) и в интерметаллических соединениях типа Nb_3Sn , V_3Ga лондоновская глубина велика ($\lambda_L \approx 200$ нм), а

параметр ξ мал ($\xi \approx 5$ нм), так что $\lambda_L \gg \xi$, поэтому полностью справедливо уравнение Лондонов.

Сверхпроводники, для которых $\lambda_L \gg \xi$, называют сверхпроводниками второго рода (или лондоновскими), а сверхпроводники, для которых $\lambda_L < \xi$, — сверхпроводниками первого рода (или пиппардовскими).

Разъясним характер нелокальной связи между плотностью тока и векторным потенциалом в сверхпроводниках первого рода. Будем рассматривать массивный односвязный сверхпроводник (т. е. сверхпроводник не в виде кольца) и будем пользоваться векторным потенциалом \mathbf{A} , удовлетворяющим условиям $\operatorname{div} \mathbf{A} = 0$, $A_n = 0$, где A_n — нормальная к поверхности сверхпроводника составляющая \mathbf{A} на границе сверхпроводника. Этими условиями векторный потенциал для односвязного сверхпроводника определяется однозначно по заданному магнитному полю, т. е. уравнение $\operatorname{rot} \mathbf{A} = \mathbf{B}$ для определения \mathbf{A} имеет одно-единственное решение.

Для массивного односвязного сверхпроводника можно считать фазу φ потенциала спаривания постоянной (см. ниже). Поэтому в выражении для тока исчезает слагаемое, содержащее $\operatorname{grad} \varphi$, и для лондоновских сверхпроводников плотность тока определяется формулой

$$\mathbf{j} = - (n_s e^2 / m_e) \mathbf{A}.$$

Здесь, как уже отмечалось, \mathbf{j} и \mathbf{A} относятся к одной и той же точке и соотношение это справедливо только в том случае, если \mathbf{j} и \mathbf{A} медленно изменяются на расстоянии порядка ξ ($\lambda_L \gg \xi$). Если это условие не выполняется, то ток в какой-либо точке \mathbf{r} зависит от значений векторного потенциала $\mathbf{A}(\mathbf{r}')$ во всех точках \mathbf{r}' , но в основном играют роль те точки \mathbf{r}' , расстояние которых до точки \mathbf{r} меньше или порядка ξ , т. е. $|\mathbf{r}' - \mathbf{r}| \leq \xi$. Формула, которая описывает этот эффект, имеет следующую структуру:

$$\mathbf{j}(\mathbf{r}) = C \int (\mathbf{A}(\mathbf{r}') \mathbf{R}) \mathbf{R} \frac{e^{-R/\xi}}{R^4} d^3 r',$$

где $\mathbf{R} = \mathbf{r}' - \mathbf{r}$ и C — некоторая константа. Для того чтобы определить ее, представим себе, что $\xi \rightarrow 0$. Тогда величину $\mathbf{A}(\mathbf{r}')$ можно вынести за знак интеграла в точке $\mathbf{r}' = \mathbf{r}$ и мы должны получить уравнение Лондонов. Отсюда вытекает, что

$$C = - 3e^2 n_s / (4\pi m_e \xi).$$

Входящий в формулу для тока (ее называют формулой Пиппарда) экспоненциальный фактор $\exp(-R/\xi)$ «срезает» вклад далеких точек, а множитель R^4 в знаменателе «компенсирует» влияние двух множителей R в числителе и множителя R^3 , соответствующего значению объема интегрирования (более высокую степень R вводить в знаменатель нельзя, так как интеграл начнет расходиться при $R \rightarrow 0$).

Покажем теперь, как с помощью формулы Пиппарда оценить глубину проникновения магнитного поля в сверхпроводники первого рода.

Обозначим эту глубину λ_p в отличие от λ_L . На расстояниях $|\mathbf{r}' - \mathbf{r}| \sim \lambda_p$ потенциал $\mathbf{A}(\mathbf{r}')$ мало отличается от $\mathbf{A}(\mathbf{r})$. Если бы потенциал $\mathbf{A}(\mathbf{r}')$ не менялся на больших расстояниях (порядка ξ , а не λ_p), то интеграл, определяющий плотность тока, сводился бы к формуле Лондонов $\mathbf{j} = -n_s e^2 / (m_e \mathbf{A})$. Но следует иметь в виду, что $\lambda_p \ll \xi$. Поэтому можно считать, что плотность тока уменьшится в ξ/λ_p раз. Таким образом, мы оценим плотность тока по формуле

$$\mathbf{j} = -\frac{n_s e^2}{m_e} \frac{\lambda_p}{\xi} \mathbf{A}; \quad \lambda_p \ll \xi.$$

Теперь мы должны использовать это соотношение вместе с уравнением $\text{rot } \mathbf{B} = \mu_0 \mathbf{j}$. Для геометрии, рассмотренной ранее, мы получим $B(z) = B(0) \exp(-z/\lambda_p)$, где λ_p теперь определяется из уравнения $\lambda_p^{-2} = e^2 \mu_0 n_s (\lambda_p / \xi)$, откуда

$$\lambda_p = \lambda_L \left(\frac{\xi}{\lambda_L} \right)^{1/3} = \xi \left(\frac{\lambda_L}{\xi} \right)^{2/3}.$$

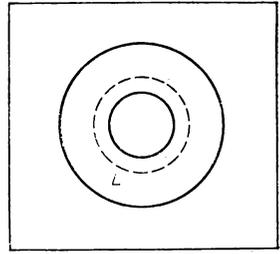
Мы видим, что $\lambda_L \ll \lambda_p \ll \xi$. Таким образом,

глубина проникновения поля в сверхпроводник первого рода значительно больше лондоновской глубины проникновения, но тем не менее значительно меньше размеров пары.

До сих пор мы рассматривали односвязные сверхпроводники. Для них, как уже указывалось, можно считать фазу φ потенциала спаривания постоянной. Но для неодносвязных сверхпроводников такое предположение уже незаконно. Действительно, возьмем сверхпроводящее кольцо и рассмотрим внутри него некоторый произвольный замкнутый контур L , лежащий далеко от поверхности кольца (рис. 13.5). На таком контуре плотность тока равна нулю, а так как плотность тока определяется величиной $(\hbar/2) \text{grad} \varphi - e \mathbf{A}$, то и она равна нулю, т. е. на L справедливо равенство $\mathbf{A} = [\hbar/(2e)] \text{grad} \varphi$. Возьмем теперь интеграл от этого выражения по контуру L . Вспоминая, что $\oint_L \mathbf{A} d\mathbf{l} = \Phi$ представляет собой магнитный поток через поверхность, опирающуюся на контур L , т. е. магнитный поток через отверстие в кольце, получим

$$\Phi = \frac{\hbar}{2e} \oint_L d\varphi.$$

Таким образом, магнитный поток через отверстие в сверхпроводящем кольце пропорционален изменению фазы φ потенциала спаривания при обходе замкнутого контура L . Если бы мы считали фазу постоянной, то поток Φ был бы равен нулю, что не соответствует, как мы знаем, действительности. Это значит, что фаза должна обязательно входить в выражение для сверхпроводящего тока.



13.5.
К объяснению квантования магнитного потока

С другой стороны, потенциал спаривания должен быть однозначной функцией координат. Отсюда следует, что изменение фазы при обходе замкнутого контура не может отличаться от кратного 2π , т. е. $\oint \mathbf{d}\varphi = 2\pi n$, где n — произвольное целое число. Таким образом, мы приходим к замечательному результату $\Phi = 2\hbar n / (2e)$. Иными словами,

магнитный поток через отверстие в сверхпроводящем кольце всегда кратен некоторому элементарному потоку Φ_0 :

$$\Phi = n\Phi_0, \quad \Phi_0 = \pi (\hbar/e).$$

Об этом факте мы уже говорили в § 13.2. На нем, в частности, основано очень точное измерение величины \hbar/e .

13.6. Промежуточное состояние

Сильное магнитное поле, как мы уже говорили, разрушает сверхпроводимость. Это разрушение имеет место независимо от происхождения поля, т. е. независимо от того, является ли поле внешним или создается самим током, текущим в сверхпроводнике. Для существования сверхпроводимости поле не должно превышать некоторого критического значения, и это условие должно выполняться повсюду. Но поле может быть разным в разных точках металлического проводника, и поэтому может сложиться ситуация, когда в одних точках металла поле окажется меньше, а в других — больше критического поля H_c . Что же тогда произойдет? Полностью сверхпроводящим металлический проводник стать не может, но он не станет и полностью нормальным. Проводник при этом разобьется на домены, одни из которых будут сверхпроводящими, а другие — нормальными. Такое состояние проводника называют *промежуточным*.

Промежуточное состояние существенно определяется геометрией проводника, так как от нее зависит характер распределения магнитного поля. Если проводник имеет форму длинного цилиндра, который помещен во внешнее магнитное поле \mathbf{H}_0 , ориентированное вдоль оси цилиндра, то на всей поверхности цилиндра поле имеет одно и то же значение (в пренебрежении влиянием концов цилиндра). Поэтому если поле \mathbf{H}_0 не превышает критического значения H_c (и температура ниже критической температуры), то весь цилиндр находится в сверхпроводящем состоянии.

Но ситуация существенно изменится, если проводник имеет, например, форму шара (рис. 13.6). Действительно, пусть сверхпроводящая сфера находится в достаточно слабом внешнем магнитном поле \mathbf{H}_0 (рис. 13.6, а). Тогда, как можно показать, компонента поля, тангенциальная к поверхности сферы, равна $H_\theta = \frac{3}{2}H_0 \sin \theta$, где θ — полярный угол. Таким образом, поле равно нулю на полюсах сферы и равно $\frac{3}{2}H_0$ на экваторе. Пусть теперь выполняется неравенство $H_c > H_0 > \frac{2}{3}H_c$ (рис. 13.6, б). Тогда у экватора и близлежащих областей поле превосходит критическое значение H_c , вблизи

же полюсов поле меньше, чем H_c . Поэтому у экватора сверхпроводимость разрушается, области же вблизи полюсов являются сверхпроводящими. Таким образом, в интервале полей $H_c > H_0 > \frac{2}{3}H_c$ сверхпроводящая сфера находится в промежуточном состоянии, в котором нормальные и сверхпроводящие области сосуществуют.

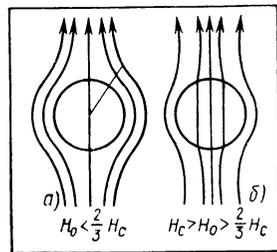
Еще более интересным примером является сверхпроводящая пластинка, находящаяся во внешнем магнитном поле, ориентированном перпендикулярно ее поверхности. В этом случае при любом поле H_0 , не превосходящем H_c , пластинка находится в промежуточном состоянии. Действительно, пластинка не может находиться полностью в сверхпроводящем состоянии, так как при этом нарушалось бы условие непрерывности нормальной составляющей поля на поверхности пластинки — внутри пластинки поле было бы равно нулю, а снаружи — H_0 . Не может пластинка находиться и полностью в нормальном состоянии — энергетически это невыгодно при $T < T_c$. Поэтому пластинка находится в промежуточном состоянии, т. е. состоит из нормальных и сверхпроводящих доменов.

Можно сделать предположение, что пластинка обладает так называемой *ламинарной структурой* типа изображенной на рис. 13.7. Здесь N и S обозначают нормальные и сверхпроводящие домены, представляющие собой слои, перпендикулярные плоскости рисунка. Магнитные силовые линии проходят только через слои N , а так как имеет место равновесие между областями N и S , то поле в областях N должно быть равно H_c . (В областях S поле, естественно, равно нулю.)

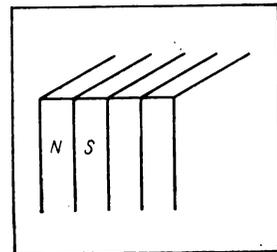
Учтем теперь, что магнитный поток должен быть непрерывен на поверхности пластинки. Снаружи пластинки он равен, очевидно, SH_0 , где S — площадь поверхности пластинки, внутри же пластинки равен $n(ad_n H_c + ad_s H_0)$, где a — размер пластинки в направлении, перпендикулярном рисунку, d_n и d_s — толщины слоев N и S , n — число слоев. Приравнявая эти выражения и учитывая, что $S = na(d_n + d_s)$, получим

$$\delta_s \equiv \frac{d_s}{d_n + d_s} = 1 - \frac{H_0}{H_c} \quad (H_0 \leq H_c).$$

Эта формула определяет относительный размер сверхпроводящих доменов, который оказывается зависящим только от отношения H_0/H_c . (Что касается самих величин d_s и d_n , то существуют экспериментальные методы их определения; эти опыты показывают, что толщина слоев значительно меньше толщины пластинки l .)



13.6. Сверхпроводник в форме шара



13.7. Сверхпроводник в форме пластинки

Для определения абсолютных значений толщин слоев d_n и d_s нужно, так же как и при определении размеров доменов в ферромагнетике, рассмотреть полную энергию образца (точнее говоря, свободную энергию) и выяснить, при каких условиях она имеет минимальное значение. При этом нужно учитывать как энергию образца, которой он обладал бы в отсутствие поля и доменов, так и энергию магнитного поля в присутствии доменов и, наконец, поверхностную энергию, связанную с существованием границы между сверхпроводящими и нормальными областями. Поверхностная энергия имеет особенно важное значение, так как она определяет толщину границы раздела между различными доменами, а эта толщина, в свою очередь, определяет размеры доменов.

Но имеется еще одно, более существенное обстоятельство. Дело в том, что наше предположение о ламинарной структуре сверхпроводника в промежуточном состоянии, в свою очередь, основывается на молчаливо сделанном предположении, о положительности поверхностной энергии между доменами. Между тем это предположение само по себе ни на чем не основано и требуется специальное исследование характера поверхностной энергии. Этот анализ показывает, что сверхпроводники первого и второго родов существенно отличаются друг от друга своей поверхностной энергией. Только

для сверхпроводников первого рода поверхностная энергия (как для ферромагнетика) положительна, для сверхпроводников же второго рода она отрицательна.

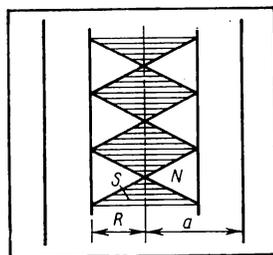
Если обозначить γ_i поверхностную энергию (отнесенную к единичной площади границы), где индекс $i = 1, 2$ обозначает род сверхпроводника, то оказывается, что

$$\gamma_1 = \frac{\mu_0 H_c^2}{2} \xi, \quad \gamma_2 = -\frac{\mu_0 H_c^2}{2} \lambda_L,$$

где ξ — размер электронной пары и λ_L — глубина проникновения поля в сверхпроводник второго рода. Таким образом, изображенная на рис. 13.7 ламинарная структура промежуточного состояния относится только к сверхпроводникам первого рода. Период этой структуры, т. е. величина $d \equiv d_s + d_n$, существенно определяется поверхностной энергией, по порядку величины $d \approx \sqrt{l\xi}$, где l — толщина пластинки.

Что касается сверхпроводников второго рода, то для них промежуточное состояние не обладает простой ламинарной структурой. Так как для них поверхностная энергия, связанная с границей раздела между нормальным и сверхпроводящим доменами, отрицательна, то эта граница должна быть максимально большой, т. е. разбиение на области N и S должно быть очень мелким (в противоположность сверхпроводникам первого рода, в которых граница между областями N и S минимальна). Если при этом лишь небольшая часть сверхпроводника второго рода находится в нормальном состоянии, большая же его часть — в сверхпроводящем, то области N имеют вид не слоев, а нитей, называемых *вихревыми нитями*.

До сих пор мы касались сверхпроводников, находящихся во внешнем магнитном поле. Если внешнего поля нет, а сверхпроводящий провод присоединен к генератору, вызывающему в нем некоторый ток I , то этот ток создает магнитное поле, которое на поверхности провода равно $H(a) = I/(2\pi a)$ (где a — радиус провода). При $H(a) < H_c$, т. е. $I < I_c$, где $I_c = 2\pi a H_c$, провод полностью находится в сверхпроводящем состоянии (речь идет о сверхпроводнике первого рода). Если же $I > I_c$, то на поверхности провода поле превышает H_c , и, следовательно, вблизи поверхности возникает слой, находящийся в нормальном состоянии. Внутренний радиус R этого слоя определяется из условия $H(R) = H_c$, откуда для тока, текущего по внутреннему сечению, получим $I_1 = 2\pi R H_c = I_c R/a$. По внешней (нормальной) части сечения, таким образом, течет ток $I - I_1$. В этой части сечения существует, очевидно, электрическое поле, направленное вдоль провода. Что касается внутренней части ($r < a$), то она не должна полностью находиться в сверхпроводящем состоянии, так как в этом случае генератор мог бы быть замкнут накоротко сверхпроводящей сердцевинкой. Поэтому внутренняя часть может находиться в промежуточном состоянии, структура которого изображена на рис. 13.8.



13.8. Сверхпроводящий провод с током, бóльшим критического

13.7. Контакт между сверхпроводниками

Как мы знаем (см. § 9.5), если привести в соприкосновение два разных металла, находящихся при одинаковой температуре, то электроны начнут переходить из одного металла в другой, но хотя между металлами появится контактная разность потенциалов, она не приведет к протеканию тока в цепи в отсутствие внешней э. д. с. Иная ситуация возникает при контакте между сверхпроводниками. Дело в том, что при контакте обычных металлов происходят переходы нормальных электронов, при контакте же сверхпроводников происходят также переходы сверхпроводящих электронов.

Если взять два сверхпроводника S и S' , разделенных тонким изолирующим слоем, то (при толщинах слоя порядка нанометров) сверхпроводящие электроны вследствие квантово-механического туннельного эффекта смогут переходить из одного сверхпроводника в другой, и так как эти электроны движутся беспрепятственно, то при включении контакта SS' в замкнутую цепь в последней возможно протекание тока без сторонней э. д. с. Мало того, контактирующие сверхпроводники могут быть сделаны из одного и того же вещества и тем не менее между ними возможен ток в отсутствие внешней э. д. с. Это явление носит название эффекта Джозефсона.

Так как эффект этот является квантовым, то для его описания нужна квантовая механика; фактически же нам требуется лишь тот факт, что существуют системы сверхпроводящих электронов (как и

любой квантово-механической системы) характеризуется некоторой функцией, называемой *волновой функцией*, которая подчиняется определенному уравнению — уравнению Шредингера (см. в этой связи § 21.2). В рассматриваемой задаче речь идет о совокупности двух сверхпроводников S и S' , между которыми возможен переход сверхпроводящих электронов. Если числа спаренных электронов в S и S' равны 2ν и $2N-2\nu$, где $2N$ — общее число спаренных электронов, которое считается заданным, то волновая функция системы $S + S'$ может быть представлена в виде $\psi_\nu = \psi_{2\nu}^S \psi_{2N-2\nu}^{S'}$, где ψ^S и $\psi^{S'}$ — волновые функции сверхпроводников S и S' порознь.

Уравнение Шредингера имеет вид

$$-\frac{\hbar}{i} \frac{\partial \psi}{\partial t} = H\psi,$$

где H — так называемый оператор Гамильтона, или гамильтониан системы, зависящий от тех динамических переменных, которые характеризуют систему.

В рассматриваемом случае простейшая структура гамильтониана такова:

$$H\psi_\nu = E_0\psi_\nu + K(\psi_{\nu+1} + \psi_{\nu-1}),$$

где E_0 и K — некоторые константы (зависящие от природы сверхпроводников S и S' и геометрии сверхпроводящего контакта, причем величина K может быть как положительной, так и отрицательной). Эта структура описывает возможность изменения числа ν , причем здесь учитываются только наиболее вероятные туннельные переходы, соответствующие изменению ν не более чем на единицу.

Нас интересует стационарный случай, когда система обладает определенной энергией E . В этом случае ψ зависит от времени как $\exp[-(i/\hbar)Et]$ и уравнение Шредингера приобретает вид $H\psi = E\psi$. Таким образом, мы приходим к уравнению

$$E_0\psi_\nu + K(\psi_{\nu+1} + \psi_{\nu-1}) = E\psi_\nu.$$

Решение этого уравнения имеет вид $\psi_\nu = A \exp\{ik\nu\}$, где A и k — константы; причем, как легко видеть,

$$E \equiv E(k) = E_0 + 2K \cos k.$$

Параметр k может принимать любые значения, так что энергия системы является функцией непрерывного параметра; при этом она заключена в интервале $(E_0 - 2K, E_0 + 2K)$. Параметр k играет важную роль; он аналогичен волновому вектору электромагнитной волны или квазимпульсу электрона в периодической решетке.

Производная от энергии по k определяет групповую скорость v_g волнового пакета, образованного суперпозицией волновых функций ψ_ν (см. в этой связи гл. 16):

$$v_g = \frac{1}{\hbar} \frac{\partial E}{\partial k} = -\frac{2K}{\hbar} \sin k.$$

Умножив групповую скорость на заряд пары $2e$ (e — заряд электрона), найдем ток I , который может течь между S и S' :

$$I = I_0 \sin k, \quad I_0 = -4eK/\hbar.$$

Это значит, что если систему SS' включить в замкнутую цепь, то в ней может в отсутствие сторонней э. д. с. протекать ток. Сила и направление этого тока определяются начальными условиями в момент замыкания цепи. При этом обязательно $|I| \leq |I_0|$, где, как мы видели, I_0 зависит только от K . Иными словами, возможное значение тока не может превосходить некоторого максимального значения, которое определяется лишь характером сверхпроводящего контакта и не зависит от начальных условий.

Пусть теперь между S и S' включена разность потенциалов V . Тогда «волновой вектор» k начнет зависеть от времени, причем, как показывает квантовомеханическая теория эффекта Джозефсона,

$$\frac{d}{dt} (\hbar k) = 2eV.$$

Отсюда следует, что если V есть некоторая заданная функция времени $V(t)$, то

$$k \equiv k(t) = k_0 + \frac{2e}{\hbar} \int_0^t V(t) dt,$$

где k_0 — значение k при $t = 0$.

Если, в частности, $V = V_0$ — константа, то k — линейная функция времени и, следовательно, ток — синусоидальная функция времени с частотой $\omega = 2eV_0/\hbar$:

$$I = I_0 \sin \left(k_0 + \frac{2e}{\hbar} V_0 t \right).$$

Частота ω очень велика, так что в среднем мы получим ток, равный нулю.

Таким образом, при $V_0 = 0$ ток наблюдается, а при $V_0 \neq 0$ — не наблюдается!

Пусть теперь наряду с постоянным напряжением V_0 приложено еще переменное напряжение $v \cos \omega t$, т. е. пусть $V(t) = V_0 + v \cos \omega t$. Тогда

$$k = k_0 + \frac{2e}{\hbar} V_0 t + \frac{2e}{\hbar \omega} v \sin \omega t.$$

Предположим, что $v \ll V_0$. В этом случае ток определится формулой

$$I = I_0 \left[\sin \left(k_0 + \frac{2eV_0}{\hbar} t \right) + \frac{2e}{\hbar \omega} v \sin \omega t \cos \left(k_0 + \frac{2eV_0}{\hbar} t \right) \right],$$

Мы должны теперь определить среднее значение этого тока I по времени. Ясно, что первое слагаемое в квадратных скобках даст нуль.

Второе же слагаемое может дать значение, отличное от нуля. Для этого необходимо, чтобы для частоты приложенного напряжения выполнялось равенство

$$\omega = (2e/\hbar)V_0$$

(условие резонанса). В этом случае среднее (по времени) значение тока

$$\langle I \rangle = -I_0 \frac{e\nu}{\hbar\omega} \sin k_0 = -\frac{1}{2} I_0 \frac{\nu}{V_0} \sin k_0.$$

Если условие резонанса не выполняется, то, очевидно, $\langle I \rangle = 0$.

Исследование этого резонанса позволяет с большой точностью находить величину e/\hbar . Дело в том, что существуют очень точные методы измерения частоты (например, с помощью интерференции), а также макроскопической величины — потенциала V_0 .

Рассмотрим, наконец, подробнее, что произойдет, если система SS' (джозефсоновский контакт) включается в замкнутую цепь, в которой действует постоянная э. д. с. U . Если R — сопротивление цепи, то напряжение на контакте SS'

$$V = U - RI = U - I_0 R \sin k.$$

Используя закон изменения k , имеем

$$\frac{dk}{dt} = \frac{2e}{\hbar} (U - I_0 R \sin k) = \omega_0 (1 - \lambda \sin k),$$

где $\omega_0 = 2eU/\hbar$ и $\lambda = RI_0/U$. Проинтегрировав это уравнение, получим

$$\omega_0 t = \int (1 - \lambda \sin k)^{-1} dk.$$

Если $\lambda < 0$, то знаменатель подынтегрального выражения не обращается в нуль и интеграл может быть легко вычислен. Полагая $\operatorname{tg} k/2 = \vartheta$, получим

$$I = I_0 \frac{2\vartheta}{1 + \vartheta^2}, \quad \vartheta = \lambda + \sqrt{1 - \lambda^2} \operatorname{tg} \left[\frac{1}{2} \sqrt{1 - \lambda^2} \omega_0 (t - t_0) \right].$$

Таким образом, при $\lambda < 1$ ток является периодической функцией времени с периодом $2\pi\omega_0^{-1}(1 - \lambda^2)^{-1/2}$; однако он (за исключением случая $\lambda = 0$) несинусоиден.

Пусть теперь $\lambda > 1$. Тогда подынтегральное выражение имеет полюс при $k = k_0$, где $\sin k_0 = \lambda^{-1}$. Это значит, что когда $t \rightarrow \infty$, то $k \rightarrow k_0$, т. е. ток стремится к своему предельному значению U/R , а напряжение на контакте — к нулю. Мы видим, что возможен режим, когда ток остается постоянным. При этом он определяется только э. д. с. и сопротивлением цепи ($I = U/R$) и не зависит от характеристик джозефсоновского контакта.

ОСНОВНЫЕ ФОРМУЛЫ

Эффект Мейсснера $B = 0$

Магнитный поток через отверстие в сверхпроводящем кольце $\Phi = n\Phi_0$, $\Phi_0 = \frac{\pi\hbar}{e} = 2,068 \cdot 10^{-15} \text{ Дж} \cdot \text{с} \cdot \text{Кл}^{-1}$
($n = 0, 1, 2, \dots$)

Скрытая теплота перехода из сверхпроводящей в нормальную фазу

$$Q_{s \rightarrow n} = -\mu_0 T H_c \frac{dH_c}{dT}$$

Скачок теплоемкости в точке сверхпроводящего перехода (формула Рутгерса)

$$C_n - C_s = -\mu_0 \left[T \left(\frac{dH_c}{dT} \right)^2 \right]_{T=T_c}$$

«Размеры» пары электронов (длина, корреляции)

$$\xi = \frac{\hbar v_F}{\pi |\Delta|}$$

Лондоновские сверхпроводники (сверхпроводники второго рода)

$$\lambda_L \gg \xi,$$

$$\lambda_L = \sqrt{\frac{m_e}{\mu_0 e^2 n_s}}$$

Уравнение Лондонов для поля

$$\Delta \mathbf{B} - \frac{1}{\lambda_L^2} \mathbf{B} = 0$$

Плотность сверхпроводящего тока в сверхпроводниках первого рода ($\xi \gg \lambda_L$)

$$\mathbf{j}_s(\mathbf{r}) = C \int (\mathbf{A}(\mathbf{r}') \cdot \mathbf{R}) \operatorname{Re} e^{-R/\xi} R^{-4} d^3 r',$$

$$C = -\frac{3e^2 n_s}{4\pi m_e \xi}$$

Глубина проникновения поля в сверхпроводник первого рода

$$\lambda_P = \lambda_L (\xi/\lambda_L)^{1/2}$$

$$\lambda_L \ll \lambda_P \ll \xi$$

Связь между частотой переменного напряжения и постоянным напряжением в эффекте Джозефсона

$$\omega = \frac{2e}{\hbar} V_0$$



III. ЭЛЕКТРОМАГНИТНЫЕ ВОЛНЫ И ОПТИКА

Глава 14. ЭЛЕКТРОМАГНИТНЫЕ ВОЛНЫ В ВАКУУМЕ

14.1. Плоская монохроматическая волна

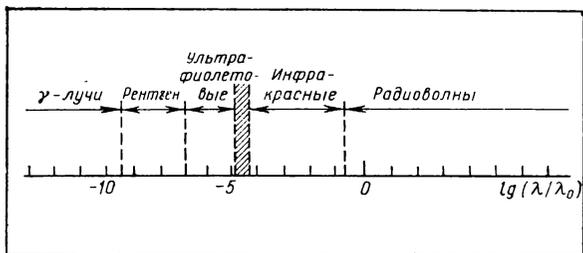
В гл. 6 было показано, что электромагнитное поле может существовать само по себе, т. е. без зарядов и токов. Такое поле, которое называют *свободным*, не может быть статическим и обязательно представляет собой волну. Это значит, что зависимость поля от времени и координат не произвольна, а, напротив, эти переменные входят только в определенных комбинациях. Проще всего обстоит дело в плоской волне, когда поля (электрическое и магнитное) зависят только от одной декартовой координаты, скажем x . Тогда если волна распространяется вдоль положительного направления x , то зависимость поля от координат и времени должна иметь вид функции от $(x - ct)$, или, что то же самое, от $(t - x/c)$. Таким образом, поле в момент $t + \Delta t$ в точке $x + \Delta x$, где $\Delta x = c\Delta t$, такое же, как поле в момент t в точке x , т. е. за время Δt поле «перемещается» на расстояние Δx . Это наглядно иллюстрирует волновой характер поля. Величину $c = 2,997925 \cdot 10^8$ м/с можно интерпретировать как скорость электромагнитной волны. Она не зависит от характера волны, т. е. носит универсальный характер и совпадает со скоростью света в вакууме — это не случайно, ибо свет есть не что иное, как электромагнитная волна.

Простейшая из плоских волн — это *плоская монохроматическая волна*, когда каждая из компонент поля f описывается гармонической функцией $f = A \cos \omega(t - \frac{x}{c})$, где A и ω — постоянные (амплитуда и частота). Важность этого типа волн обусловлена тем, что любую электромагнитную волну можно представить как суперпозицию плоских монохроматических волн. Напомним, что $T = 2\pi/\omega$ называют периодом, $\lambda = cT = 2\pi c/\omega$ — длиной волны.

На рис. 14.1 схематически представлена шкала электромагнитных волн (так как диапазон волн огромен, то использована логарифмиче-

14.1.

Шкала электромагнитных волн (диапазон видимого света отмечен штриховкой)



ская шкала, т. е. приведены значения $\lg(\lambda/\lambda_0)$, где в качестве λ_0 выбрано $\lambda_0 = 1$ см. Только ничтожная часть этого диапазона принадлежит к видимому свету (от 390 до 770 нм, на рис. 14.1 этот диапазон заштрихован). К электромагнитным волнам относятся и радиоволны (диапазон которых начинается с длин волн порядка 1 мм), и рентгеновые лучи (длины волн которых лежат в интервале 1,0—0,005 нм), и γ -лучи (длина волны которых в космическом излучении составляет 10^{-6} нм), а также ультрафиолетовые и инфракрасные лучи.

Так как уравнения Максвелла линейны, то для описания поля удобно пользоваться комплексными величинами, т. е. вначале считать, что поле плоской волны f имеет вид $f = Ae^{-i\omega(t-x/c)}$, где A — комплексная амплитуда, и только после того, как она будет найдена, перейти к вещественным величинам, полагая $f = \text{Re}[Ae^{-i\omega(t-x/c)}]$. Удобно также не выбирать ось x в качестве направления распространения волны, а ввести единичный вектор \mathbf{n} вдоль этого направления. Тогда поля в плоской монохроматической волне можно представить в виде $\mathbf{E} = \mathbf{E}_0 \exp\{i(\mathbf{k}\mathbf{r} - \omega t)\}$, $\mathbf{H} = \mathbf{H}_0 \exp\{i(\mathbf{k}\mathbf{r} - \omega t)\}$, где $\mathbf{k} = \mathbf{n}\omega/c$ — волновой вектор. Подставляя эти выражения в уравнения Максвелла

$$\text{rot } \mathbf{E} = -\partial \mathbf{B} / \partial t, \quad \text{div } \mathbf{B} = 0,$$

$$\mu_0^{-1} \text{rot } \mathbf{B} = \epsilon_0 \partial \mathbf{E} / \partial t, \quad \text{div } \mathbf{E} = 0$$

и учитывая, что $\epsilon_0 \mu_0 = c^{-2}$ и что для плоской волны $\text{rot } \mathbf{A} = i[\mathbf{k}\mathbf{A}]$, $\text{div } \mathbf{A} = i\mathbf{k}\mathbf{A}$, получим

$$\mathbf{B}_0 = \frac{1}{c} [\mathbf{n}\mathbf{E}_0], \quad \mathbf{E}_0 = -c [\mathbf{n}\mathbf{B}_0], \quad \mathbf{n}\mathbf{E}_0 = \mathbf{n}\mathbf{B}_0 = 0.$$

Таким образом, комплексные амплитуды полей взаимно ортогональны и, кроме того, ортогональны к направлению распространения волны.

Как мы знаем, плотности электрической и магнитной энергии пропорциональны квадратам полей: $\omega_e = \epsilon_0 E^2/2$, $\omega_m = B^2/(2\mu_0)$ (E и B здесь вещественны!). Средние значения квадратов синуса и косинуса равны $1/2$, поэтому $\langle \omega_e \rangle = \epsilon_0 |\mathbf{E}_0|^2/4$, $\langle \omega_m \rangle = |\mathbf{B}_0|^2/(4\mu_0)$. Мы видим, что средние значения плотностей электрической и магнитной энергии равны между собой. Для суммарной плотности электромагнитной энергии имеем

$$\omega = \langle \omega_e \rangle + \langle \omega_m \rangle = \epsilon_0 |\mathbf{E}_0|^2/2 = |\mathbf{B}_0|^2/(2\mu_0).$$

Вычислим плотность потока электромагнитной энергии в плоской монохроматической волне: $\mathbf{S} = [\mathbf{E}\mathbf{H}]$. Так как $\mathbf{E} \perp \mathbf{H}$ и $E = Hc\mu_0$, то $\langle \mathbf{S} \rangle = c\omega\mu_0$. Этот результат имеет наглядный смысл: он означает, что

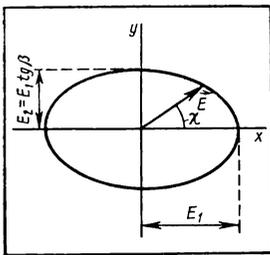
энергия переносится вместе с волной со скоростью света c .

14.2. Поляризация электромагнитной волны

В предыдущем параграфе мы видели, что в плоской монохроматической волне электрическое и магнитное поля направлены перпендикулярно направлению распространения волны. В таких случаях говорят о волне с *поперечной поляризацией*, или просто о поперечной волне. Заметим, что иначе обстоит дело со звуковой волной в жидкости или газе (но не в кристалле!), в которых вектор скорости частиц среды направлен вдоль направления распространения волны (продольная волна, или волна с *продольной поляризацией*).

Остановимся подробнее на поляризации плоской монохроматической волны. Ее комплексную амплитуду можно представить в виде $\mathbf{E}_0 = \mathbf{E}_1 + i\mathbf{E}_2$, где $\mathbf{E}_{1,2}$ — два вещественных вектора, ортогональные (как и вектор \mathbf{E}_0) к направлению распространения волны. Выбор \mathbf{E}_1 и \mathbf{E}_2 , однако, неоднозначен, так как можно произвольно выбрать фазу \mathbf{E}_0 (изменение фазы \mathbf{E}_0 равнозначно изменению начала отсчета времени). Поэтому можно выбрать векторы \mathbf{E}_1 и \mathbf{E}_2 таким образом, чтобы они были взаимно ортогональны. Формула $\mathbf{E} = \text{Re}\{\mathbf{E}_0 \exp(i\mathbf{k}\mathbf{r} - \omega t)\}$ означает тогда, что электрическое поле плоской монохроматической волны всегда можно представить в виде суперпозиции двух полей, направленных вдоль двух взаимно перпендикулярных постоянных векторов и изменяющихся, как $\cos(\mathbf{k}\mathbf{r} - \omega t)$ и $\sin(\mathbf{k}\mathbf{r} - \omega t)$, т. е. со сдвигом по фазе на $\pi/2$. В случае вещественного \mathbf{E}_0 (т. е. $\mathbf{E}_2 = 0$) говорят о *линейной поляризации* волны. Таким образом, можно сказать, что

в общем случае поле плоской монохроматической волны представляет собой суперпозицию двух взаимно перпендикулярных полей, соответствующих двум линейно поляризованным волнам, причем разность фаз полей равна $\pi/2$.



14.2.

Эллипс, описываемый вектором напряженности электрического поля

Наличие трех взаимно перпендикулярных векторов \mathbf{k} , \mathbf{E}_1 , \mathbf{E}_2 позволяет естественным образом выбрать декартову систему координат, орты которой направлены вдоль этих векторов. А именно: направим ось z вдоль \mathbf{k} , ось x — вдоль \mathbf{E}_1 и ось y — вдоль \mathbf{E}_2 (либо — \mathbf{E}_2 , так чтобы система координат была праввинтовой). Тогда $(E_x/E_1)^2 + (E_y/E_2)^2 = 1$. Мы видим, что конец вектора \mathbf{E} в каждой точке пространства описывает с течением времени эллипс, полуоси которого равны E_1 и E_2 (рис. 14.2). Вращение этого вектора, образующего угол $\chi = \mathbf{k}\mathbf{r} - \omega t$ с осью x , происходит

либо по часовой стрелке (если $\mathbf{E}_2 \parallel y$), либо против нее (если $-\mathbf{E}_2 \parallel y$). В первом случае говорят о *правополяризованной*, а во втором — о *левополяризованной* эллиптической волне. В частном случае, когда $E_1 = E_2$ и эллипс вырождается в окружность, говорят о *круговой поляризации* волны.

Легко показать, что вектор \mathbf{B} , как и вектор \mathbf{E} , описывает эллипс, причем оба эллипса равны и ориентированы взаимно перпендикулярно.

14.3. Частично поляризованная волна

Наличие определенной поляризации не есть необходимое свойство электромагнитной волны с определенными частотой и направлением распространения. Если такая волна излучается одним атомом, то она обладает строго определенной поляризацией. Но если плоские монохроматические волны, пусть даже с одинаковыми ω и \mathbf{k} , излучаются рядом идентичных атомов, то результирующая волна, оставаясь плоской монохроматической, уже не обладает определенной поляризацией. Дело в том, что ее комплексная амплитуда складывается из комплексных амплитуд отдельных волн, фазы же этих амплитуд могут быть самыми различными и никак не связанными между собой, в силу чего результирующая комплексная амплитуда — случайная, а не строго определенная величина. Чтобы разъяснить это, введем поле, возбуждаемое j -м атомом: $\mathbf{E}^{(j)} = \text{Re } \mathcal{E}^{(j)} e^{i[\mathbf{k}\mathbf{r} - \omega(t-t_j)]}$, где $\mathcal{E}^{(j)}$ — комплексная амплитуда (векторная) и t_j — начало отсчета времени для j -го атома (скажем, время начала излучения). Поле $\mathbf{E}^{(j)}$ будем считать эллиптически поляризованным. Результирующее поле $\mathbf{E} = \text{Re} \mathcal{E} \exp(i\mathbf{k}\mathbf{r} - i\omega t)$, создаваемое всеми атомами, является суммой полей $\mathbf{E}^{(j)}$; здесь \mathcal{E} — комплексная амплитуда результирующего поля;

$$\mathcal{E} = \sum_j \mathcal{E}^{(j)} \exp(i\omega t_j).$$

Входящие сюда величины t_j являются, очевидно, случайными; кроме того, случайны и углы между осями эллипсов поляризации волн с разными j . Поэтому амплитуда \mathcal{E} является случайной величиной.

Отсюда следует, что при исследовании поляризационных свойств волны, возникающей в результате суперпозиции многих плоских монохроматических волн, излучаемых различными независимыми источниками, должно быть использовано понятие вероятности. А именно: если разложить результирующую амплитуду \mathcal{E} по закрепленным осям $\mathbf{e}_1, \mathbf{e}_2$, $\mathcal{E} = \mathcal{E}_1 \mathbf{e}_1 + \mathcal{E}_2 \mathbf{e}_2$, то комплексные величины \mathcal{E}_1 и \mathcal{E}_2 (скалярные) будут случайными и должна быть задана вероятность распределения этих величин.

Структура величин $\mathcal{E}_1, \mathcal{E}_2$ такова, что их средние значения $\langle \mathcal{E}_1 \rangle$ и $\langle \mathcal{E}_2 \rangle$ по этому распределению равны нулю. Действительно, так как моменты времени t_j в суперпозиционных множителях $\exp(i\omega t_j)$ в выражении для \mathcal{E} с равной вероятностью принимают любые значения, то среднее значение этих множителей равно нулю: $\langle \exp(i\omega t_j) \rangle = 0$. По той же причине $\langle \mathcal{E}_\alpha \mathcal{E}_\beta \rangle = 0$ (здесь $\alpha, \beta = 1, 2$) и только ве-

личины $\langle \mathcal{E}_\alpha \mathcal{E}_\beta^* \rangle$ отличны от нуля (угловые скобки служат для обозначения усреднения). Эти величины и характеризуют плоскую монохроматическую волну со случайной комплексной амплитудой. При этом $\langle |\mathcal{E}|^2 \rangle \equiv \langle \sum_\alpha \mathcal{E}_\alpha \mathcal{E}_\alpha^* \rangle$ определяет интенсивность волны, а четыре величины

$$\rho_{\alpha\beta} = \langle |\mathcal{E}|^2 \rangle^{-1} \langle \mathcal{E}_\alpha \mathcal{E}_\beta^* \rangle$$

— поляризацию волны. Величины $\rho_{\alpha\beta}$ образуют двухрядную матрицу, которая является эрмитовой: $\rho_{\alpha\beta} = \rho_{\beta\alpha}^*$. Сумма ее диагональных элементов (называемая следом или шпуром) равна единице: $\rho_{11} + \rho_{22} \equiv \text{Sp } \rho = 1$. Матрицу ρ называют *поляризационной матрицей*. Разумеется, значения ее элементов зависят от выбора базиса, т. е. ортов $\mathbf{e}_1, \mathbf{e}_2$.

Для эллиптически поляризованной волны поляризационная матрица имеет вид $\rho_{\alpha\beta} = |\mathcal{E}|^{-2} \mathcal{E}_\alpha \mathcal{E}_\beta^*$; здесь, естественно, нет никакого усреднения. Эта матрица представляет собой просто произведения компонента вектора $\mathcal{E} / |\mathcal{E}|$ и комплексно-сопряженного вектора $\mathcal{E}^* / |\mathcal{E}|$. Справедливо и обратное утверждение: для того чтобы волна обладала определенной поляризацией, необходимо, чтобы ее поляризационная матрица была представима в таком виде.

Поляризационную матрицу для эллиптически поляризованной волны можно представить в виде

$$\rho = \frac{1}{2} \begin{pmatrix} 1 + \cos 2\beta \cos 2\chi, & \cos 2\beta \sin 2\chi - i \sin 2\beta \\ \cos 2\beta \sin 2\chi + i \sin 2\beta, & 1 - \cos 2\beta \cos 2\chi \end{pmatrix},$$

где χ — угол между большой полуосью и ортом \mathbf{e}_1 и β — степень эллиптичности поляризации (отношение полуосей эллипса поляризации, рис. 14.2). Положив здесь $\chi = 0$, получим ρ в системе координат, оси которой направлены вдоль главных осей эллипса поляризации.

Если $\rho_{\alpha\beta} = \frac{1}{2} \delta_{\alpha\beta}$, то волну называют *неполяризованной* (в случае света говорят о *естественном свете*).

В общем случае частично поляризованной волны поляризационную матрицу можно представить в виде

$$\rho_{\alpha\beta} = \frac{1}{2} (1 - P) \delta_{\alpha\beta} + P e_\alpha e_\beta^*,$$

где e_α — компоненты некоторого единичного вектора и P — неотрицательное число, меньшее единицы. При $P = 1$ мы получаем поляризованную волну, если же $P = 0$, то волна является неполяризованной. Поэтому величину P называют *степенью поляризации*.

Часто поляризационную матрицу записывают в виде

$$\rho = \frac{1}{2} \begin{pmatrix} 1 + \xi_3, & \xi_1 - i\xi_2 \\ \xi_1 + i\xi_2, & 1 - \xi_3 \end{pmatrix},$$

где ξ_1, ξ_2, ξ_3 — вещественные величины, не превосходящие единицы. Их называют *параметрами Стокса*. Для полностью поляризованной волны

$$\xi_1 = \cos 2\beta \sin 2\chi, \quad \xi_2 = \sin 2\beta, \quad \xi_3 = \cos 2\beta \cos 2\chi,$$

для частично поляризованной

$$\xi_1 = P \cos 2\beta \sin 2\chi, \quad \xi_2 = P \sin 2\beta, \quad \xi_3 = P \cos 2\beta \cos 2\chi.$$

Мы видим, что

$$\xi_1^2 + \xi_2^2 + \xi_3^2 = P^2 \leq 1.$$

Существенно, что параметры Стокса для света могут быть непосредственно определены экспериментально. Для этого пропускают исследуемый свет через некоторые кристаллы и определяют интенсивность прошедшего света, который оказывается линейно поляризованным. Его относительная интенсивность и определяет величины ξ_1, ξ_2, ξ_3 . Так, если свет линейно поляризован вдоль оси x_1 , то относительная интенсивность прошедшего света ω_1 непосредственно определяет $\rho_{11} = \frac{1}{2}(1 + \xi_3) = \omega_1$. Если далее получен свет, линейно поляризованный вдоль оси, составляющей с осью x_1 угол $\pi/4$, то его относительная интенсивность ω_2 определит параметр ξ_1 , т. е. $\omega_2 = \frac{1}{2}(1 + \xi_1)$. Наконец, чтобы найти ξ_2 , нужно получить из исследуемого света не линейно поляризованный, а поляризованный по кругу свет, что также достигается с помощью кристаллов. Таким образом, для полного исследования поляризационных свойств света необходимы измерения интенсивностей, соответствующих двум линейным поляризациям (удобнее всего — под углом $\pi/4$ одна к другой) и одной круговой поляризации.

Отметим в заключение этого параграфа один любопытный факт. Глаз человека, как мы знаем, различает длину волны (цвет) и интенсивность (яркость) видимого света (большинство млекопитающих не различает даже длины волны), но не различает (как правило) его поляризации. Глаза же некоторых насекомых, в том числе пчелы, устроены так, что кроме цвета и яркости они различают и поляризацию света. Именно эта способность позволяет пчеле находить направление на солнце даже в облачный день.

14.4. Волны в пространстве между металлическими поверхностями

Рассмотрев электромагнитные волны в безграничном пространстве, перейдем к изучению волн в ограниченном пространстве. Начнем с полупространства, ограниченного металлом.

Как мы видели в гл. 6, на границе двух сред должны быть непрерывны величины E_t и B_n (индексы t и n означают тангенциальную и нормальную к границе составляющие поля). Если одна из рассматриваемых сред — хороший проводник, в частности металл, то можно считать, что высокочастотное поле в глубь него не проникает. Физически это связано с тем, что в среде с бесконечной проводимостью даже слабые поля создавали бы бесконечные токи; мы вернемся к этому вопросу подробнее в гл. 16. Таким образом, на поверхности проводника

переменное электромагнитное поле должно удовлетворять граничным условиям $E_t = 0$, $B_n = 0$.

Однако если рассмотреть одну отдельно взятую электромагнитную волну, падающую на поверхность металла, то эти граничные условия удовлетворяются, вообще говоря, не будут. Отсюда вытекает, что

на границе с металлом должна возникнуть еще одна — отраженная электромагнитная волна, так что электромагнитное поле в вакууме представляет собой суперпозицию двух волн — падающей и отраженной и эта суперпозиция удовлетворяет нужным граничным условиям.

Чтобы разъяснить это, предположим, что на металлическую плоскость $x = 0$ падает плоская монохроматическая волна с волновым вектором \mathbf{k} и частотой ω . Выберем ось y перпендикулярно \mathbf{k} ; тогда вектор \mathbf{k} имеет только две составляющие: k_x и k_z . Будем для простоты считать, что электрическое поле падающей волны $\mathbf{E}^{(i)}$ лежит в плоскости (x, z) :

$$E_x^{(i)} = A \exp \{ik_x x + ik_z z - i\omega t\},$$

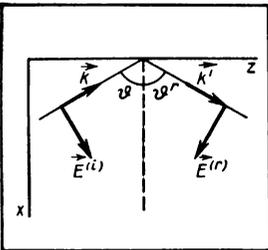
$$E_z^{(i)} = B \exp \{ik_x x + ik_z z - i\omega t\}, \quad E_y^{(i)} = 0,$$

где A и B — амплитуды компонент поля. Эти амплитуды не независимы, а связаны условием поперечности электромагнитной волны, поэтому $B = -Ak_x/k_z$.

Обозначим далее $\mathbf{E}^{(r)}$ напряженность поля отраженной волны и потребуем, чтобы для суммарной напряженности $\mathbf{E}_t = \mathbf{E}^{(i)} + \mathbf{E}^{(r)}$ выполнялись граничные условия на поверхности $x = 0$. Так как $E_y^{(i)} = 0$, то, очевидно, и $E_y^{(r)} = 0$, т. е. отличны от нуля только x - и z -компоненты поля $\mathbf{E}^{(r)}$: $E_x^{(r)} = A' \exp \{ik'_x x + ik'_z z - i\omega' t\}$,

$$E_z^{(r)} = B' \exp \{ik'_x x + ik'_z z - i\omega' t\},$$

где \mathbf{k}' и ω' — волновой вектор и частота отраженной волны, A' , B' — амплитуды, связанные (в силу поперечности последней) соотношением $B' = -A'k'_x/k'_z$. Граничное условие $E_t^{(i)} + E_t^{(r)} = 0$ сведется к соотношению $B \exp \{ik_z z - i\omega t\} + B' \exp \{ik'_z z - i\omega' t\} = 0$, которое должно выполняться тождественно при всех значениях z и t . А это может быть только в том случае, если $k'_z = k_z$, $\omega' = \omega$ и $B' = -B$. Но $\omega = ck$, $\omega' = ck'$, поэтому модуль волнового вектора не изменяется, а меняется только знак его x -составляющей: $k'_x = -k_x$. Отсюда следует, что угол падения θ равен углу отражения θ' (углы отсчитываются от нормали к поверхности, рис. 14.3).



14.3. Векторы напряженностей электрического поля падающей и отраженной волн

Учтем, наконец, связь между амплитудами A , B и A' , B' , тогда получим $A' = A$. Заметим, что в рассматриваемой задаче, как легко убедиться, граничное условие $B_n = 0$ выполняется тождественно.

Сложив $E^{(i)}$ и $E^{(r)}$, получим суммарное электрическое поле, возникающее при падении плоской монохроматической волны на металлическую поверхность:

$$E_x = 2A \cos k_x x \cdot \exp(ik_z z - i\omega t),$$

$$E_z = 2iB \sin k_x x \cdot \exp(ik_z z - i\omega t), \quad E_y = 0.$$

Мы видим, что бегущая волна сохраняется только в тангенциальном к поверхности направлении, в нормальном же направлении возникает стоячая волна. При этом компонента поля E_z обращается в нуль на плоскостях $x = 0$, $x = \pi/k_x$, $x = 2\pi/k_x$, ... (которые можно назвать *плоскостями узлов E_z*), максимума же эта компонента достигает на плоскости $x = \pi/(2k_x)$, $x = 3\pi/(2k_x)$ (плоскости пучностей этой компоненты). Там, где находятся узлы E_z , у компоненты E_x — пучности, и наоборот. Обратим внимание на множитель i в выражении для E_z . Он означает, что величины E_z и E_x сдвинуты по фазе на $\pi/2$, т. е. на четверть периода. Аналогичную структуру имеет и магнитное поле.

Посмотрим теперь, что получится, если параллельно первой на расстоянии a от нее поместить вторую металлическую поверхность. Компоненты поля будут определяться прежними формулами, но теперь поле E_z должно обращаться в нуль еще и на второй поверхности, т. е. при $x = a$. Отсюда следует, что $\sin ak_x = 0$, т. е. $k_x = \pi n/a$, где n — целое число. Иными словами, составляющая волнового вектора, нормальная к обоим металлическим плоскостям, должна быть дискретной, а именно кратной π/a .

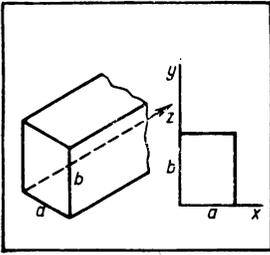
Заметим еще, что $k_x = k \cos\theta = 2\pi \cos\theta/\lambda$, где θ — угол падения и λ — длина волны, откуда $\cos\theta = \pi\lambda/(2a)$. Но $\cos\theta$ не может быть больше единицы и, следовательно, λ не может быть больше $2a$. Можно сказать, что

более длинные волны «не входят» в пространство между металлическими поверхностями.

При $\lambda < 2a$ угол θ , как мы видим, тоже не может быть произвольным: он может принимать лишь n_m дискретных значений, где n_m — целая часть отношения $2a/\lambda$ (наибольшее положительное число, меньшее $2a/\lambda$).

14.5. Волноводы

Итак, поле в пространстве между двумя металлическими плоскостями представляет собой волну, бегущую вдоль этих поверхностей и стоячую в направлении нормали к поверхностям. Ясно, что если, добавив еще две металлические плоскости, рассмотрим бесконечно длинный прямоугольный параллелепипед, то поле окажется стоячей волной уже в двух направлениях и бегущей только в одном — вдоль оси параллелепипеда. Такой бесконечный параллелепипед, или в общем случае цилиндр произвольного сечения,



14.4.
Прямоугольный волно-
вод и его сечение

может использоваться для передачи волны в определенном направлении, поэтому его называют *волноводом*.

Рассмотрим простейший прямоугольный волновод, сечение которого имеет вид прямоугольника со сторонами a и b (рис. 14.4). Легко понять, какую структуру должна иметь z -компонента электрического поля в волноводе. Так как она должна обращаться в нуль на всей поверхности волновода, то, очевидно, $E_z = C \sin(k_x x) \sin(k_y y) \exp(ik_z z - i\omega t)$, где C — постоянная и k_x, k_y — компоненты волнового вектора, могущие принимать только дискретные значения $k_x = \pi m/a,$

$k_y = \pi n/b$ ($m, n = 1, 2, \dots$). Так как $\omega = ck$ и $k = \sqrt{k_x^2 + k_y^2 + k_z^2}$, то

$$\omega = c \sqrt{\left(\frac{\pi m}{a}\right)^2 + \left(\frac{\pi n}{b}\right)^2 + k_z^2}.$$

Целые числа m и n имеют простой физический смысл: они определяют числа узловых точек компоненты поля E_z , а именно: вдоль оси x имеется $m + 1$, а вдоль оси y — $n + 1$ узловых точек. Волну с заданными значениями m и n называют *mn -гармоникой E -поля* и обозначают E_{mn} .

Распространению волны соответствуют вещественные значения k_z . Поэтому распространяться вдоль волновода могут лишь такие волны, частота которых превосходит $\pi c \sqrt{(m/a)^2 + (n/a)^2}$. Эту частоту называют *граничной* и обозначают ω_{mn} . Очевидно, $\omega = \sqrt{\omega_{mn}^2 + c^2 k_z^2}$.

Найдем скорость волны вдоль оси волновода. Она определяется из условия постоянства фазового множителя $\exp\{ik_z z - i\omega t\}$, а этот множитель постоянен для наблюдателя, движущегося вдоль оси z со скоростью $dz/dt = \omega/k_z$. Для этой скорости, называемой *фазовой скоростью* волны, получим $v_\phi = c(1 - \omega_{mn}^2/\omega^2)^{-1/2}$; она, очевидно, больше скорости света и при $\omega = \omega_{mn}$ обращается даже в бесконечность.

Выясним теперь структуру различных компонент поля в волноводе с поперечным сечением произвольной формы. Обратимся для этого к уравнениям Максвелла, которые запишем в виде

$$[\nabla \mathbf{E}] = - \frac{\partial \mathbf{B}}{\partial t}, \quad [\nabla \mathbf{H}] = \frac{\partial \mathbf{D}}{\partial t},$$

где $\nabla = \mathbf{i}_x \partial/\partial x + \mathbf{i}_y \partial/\partial y + \mathbf{i}_z \partial/\partial z$ — оператор набла ($\mathbf{i}_x, \mathbf{i}_y, \mathbf{i}_z$ — орты вдоль осей x, y, z), \mathbf{B} и \mathbf{D} — индукции, связанные с напряженностями \mathbf{H} и \mathbf{E} соотношениями $\mathbf{B} = \mu_0 \mu \mathbf{H}$, $\mathbf{D} = \epsilon_0 \epsilon \mathbf{E}$. Здесь ϵ и μ — проницаемости «наполнителя» волновода, которые равны единице

в случае вакуума. Все компоненты полей мы считаем пропорциональными $\exp\{ik_z z - i\omega t\}$. Поэтому в уравнениях Максвелла

$$[\nabla\mathbf{E}] = i\omega\mathbf{B}, \quad [\nabla\mathbf{B}] = -i\omega\epsilon_0\mu_0\epsilon\mu\mathbf{E} \quad (\epsilon_0\mu_0 = c^{-2})$$

мы можем положить

$$\nabla = \nabla_{\perp} + i\mathbf{i}_z k_z, \quad \nabla_{\perp} = \mathbf{i}_x \frac{\partial}{\partial x} + \mathbf{i}_y \frac{\partial}{\partial y}.$$

Заменяя \mathbf{E} и \mathbf{B} на $\mathbf{E} = \mathbf{E}_{\perp} + \mathbf{i}_z E_z$, $\mathbf{B} = \mathbf{B}_{\perp} + \mathbf{i}_z B_z$ (\mathbf{E}_{\perp} , \mathbf{B}_{\perp} и E_z , B_z — поперечные и продольные составляющие полей относительно направления распространения z), имеем

$$[\nabla\mathbf{E}] = [\nabla_{\perp} \mathbf{i}_z] E_z + [\nabla_{\perp} \mathbf{E}_{\perp}] + ik_z [\mathbf{i}_z \mathbf{E}_{\perp}]$$

и аналогичное соотношение для $[\nabla\mathbf{B}]$. Подставим эти выражения в уравнения Максвелла и умножим их векторно на \mathbf{i}_z . Используя формулу $[\mathbf{a}[\mathbf{b}\mathbf{c}]] = \mathbf{b}(\mathbf{a}\mathbf{c}) - \mathbf{c}(\mathbf{a}\mathbf{b})$, получим

$$[\mathbf{i}_z [\nabla \mathbf{i}_z]] E_z = \nabla_{\perp} E_z, \quad [\mathbf{i}_z [\nabla_{\perp} \mathbf{E}_{\perp}]] = 0.$$

Поэтому уравнения Максвелла примут вид

$$\nabla_{\perp} E_z - ik_z E_{\perp} = i\omega [\mathbf{i}_z \mathbf{B}_{\perp}],$$

$$\nabla_{\perp} B_z - ik_z \mathbf{B}_{\perp} = -i\omega c^{-2} \epsilon\mu [\mathbf{i}_z \mathbf{E}_{\perp}].$$

Эти уравнения можно решить относительно \mathbf{E}_{\perp} и \mathbf{B}_{\perp} :

$$i(k_z^2 - \epsilon\mu\omega^2 c^{-2}) \mathbf{E}_{\perp} = k_z \nabla_{\perp} E_z - \omega [\mathbf{i}_z \nabla_{\perp}] B_z.$$

$$i(k_z^2 - \epsilon\mu\omega^2 c^{-2}) \mathbf{B}_{\perp} = k_z \nabla_{\perp} B_z + \epsilon\mu\omega c^{-2} [\mathbf{i}_z \nabla_{\perp}] E_z.$$

Мы видим, что поперечные составляющие полей однозначно выражаются через продольные составляющие, которые входят в правые части этих равенств независимо. Поэтому поля в волноводах можно представить как суперпозиции полей двух независимых типов. В полях одного типа $E_z \neq 0$, а $B_z = 0$; в полях другого типа $B_z \neq 0$, а $E_z = 0$. Поля первого типа называют *E-волнами* или *поперечно-магнитными* (сокращенно *ТМ-волнами*). Поля второго типа называют *H-волнами* или *поперечно-электрическими* (сокращенно *ТЕ-волнами*).

Все компоненты полей *E-волн* выражаются через E_z , а *H-волн* — через H_z ; сами же эти величины удовлетворяют одинаковым уравнениям

$$\Delta_{\perp} E_z + k_{\perp}^2 E_z = 0, \quad \Delta_{\perp} H_z + k_{\perp}^2 H_z = 0,$$

где $k_{\perp}^2 = \epsilon\mu\omega^2 c^{-2} - k_z^2$ и $\Delta_{\perp} = \partial^2/\partial x^2 + \partial^2/\partial y^2$, но с разными граничными условиями. Для *E-волн* на контуре S поперечного сечения волновода должно выполняться условие $E_z = 0$. Для *H-волн* в нуль должно обращаться \mathbf{E}_t , а это условие, как легко убедиться, эквивалентно обращению в нуль нормальной к контуру производной от H_z . Таким образом, должны выполняться граничные условия

$$E_z \Big|_C = 0, \quad \frac{\partial H_z}{\partial n} \Big|_C = 0.$$

Вернемся к прямоугольному вакуумному волноводу и выпишем выражение для H_z в H -волне:

$$H_z = C \cos \frac{\pi m x}{a} \cos \frac{\pi n y}{b} e^{i(k_z z - \omega t)} \quad (m, n = 0, 1, \dots),$$

где C — постоянная. Оно, как мы видим, отличается от полученного выше выражения для E_z в E -волне заменой синусов на косинусы. Граничная частота для волн обоих типов есть

$$\omega_{mn} = \pi c \sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2}.$$

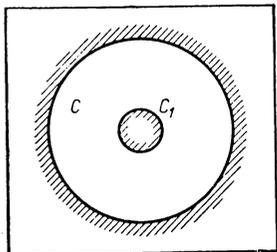
Если для определенности считать, что $a < b$, то наименьшей граничной частотой $\omega_{01} = \pi c n / b$ обладает H -волна с $m = 0, n = 1$, или H_{01} -волна (для E -волн наименьшей граничной частотой является ω_{11}).

Можно показать, что

у волновода произвольного (односвязного) сечения тоже обязательно существует граничная частота $\omega_{гр}$, причем $\omega = \sqrt{\omega_{гр}^2 + c^2 k_z^2}$ и, следовательно, $v_{ф} > c$.

Как и в частном случае прямоугольного волновода, граничная частота определяется двумя целыми числами m и n и может быть обозначена ω_{mn} (хотя в общем случае зависимость ω_{mn} от m и n сложнее, чем в случае прямоугольного волновода).

Зададимся теперь вопросом: могут ли в волноводе распространяться волны с такой же структурой полей, как в неограниченном пространстве, т. е. волны чисто поперечные с $E_z = H_z = 0$? Полученная выше неоднородная (содержащая в правых частях E_z, H_z) система уравнений для $\mathbf{E}_\perp, \mathbf{H}_\perp$ превращается для поперечных волн в однородную; условие существования у нее нетривиальных решений дает $\omega = c k_z$. Иными словами, фазовая скорость такой волны равна скорости света и, следовательно, не существует граничной частоты: $\omega_{mn} = 0$. Но если $\omega = c k_z$, то $k_\perp = 0$ и поэтому уравнения для полей $\mathbf{E}_\perp, \mathbf{H}_\perp$ перестают быть волновыми и превращаются в двумерные уравнения Лапласа — уже знакомые нам уравнения, описывающие электростатическое поле. Но электростатическое поле внутри односвязной области, ограниченной металлической поверхностью, равно нулю — уравнение Лапласа с нулевыми граничными условиями в односвязной области не имеет решений. Мы приходим к выводу, что



14.5.

Поперечное сечение коаксиального кабеля

в волноводе с односвязным сечением поперечные электромагнитные волны распространяться не могут; в нем всегда $\omega_{mn} > 0$ и $v_{ф} > c$.

Иная ситуация возникает для волновода с двусвязным сечением, например если внутри него находится цилиндрический проводник (рис. 14.5, на котором C_1 — контур этого проводника). В этом случае уравнение Лапласа в пространстве между поверхностями C и C_1 имеет нетривиальные решения, так что

в волноводе возможно распространение поперечных волн со скоростью c . Важнейшим с практической точки зрения волноводом такого типа является коаксиальный кабель.

Подчеркнем, что распределение поля в двусвязной области между поверхностями C и C_1 носит электростатический характер.

14.6. Резонаторы

В волноводах происходит свободное распространение волн вдоль оси волновода — в этом направлении возникает бегущая волна, в перпендикулярных же направлениях из-за отражения волн от металлической поверхности волновода возникают стоячие волны.

Посмотрим теперь, что получится, если перегородить волновод с двух сторон металлическими перегородками, например металлическими плоскостями $z = 0$ и $z = d$. Волны отразятся от этих торцовых поверхностей, в результате чего и вдоль оси z вместо бегущей возникнет стоячая волна. Граничные условия для полей на торцовых поверхностях, разумеется, те же, что на стенках цилиндра, а именно должны обращаться в нуль тангенциальная составляющая электрического и нормальная составляющая магнитного поля.

Мы видели, что продольные компоненты полей в волноводе имеют структуру $C\psi(x, y) \exp(ik_z z - i\omega t)$, где вещественной функцией ψ определяется зависимость от x, y (она различна для E - и H -волн) и C — произвольная комплексная постоянная (комплексная амплитуда). Поэтому поле бегущей в положительном направлении оси z волны пропорционально $C \exp(ik_z z - i\omega t)$, а поле отраженной, т. е. бегущей в отрицательном направлении этой оси, волны пропорционально $C' \exp(-ik_z z - i\omega t)$ (C' — комплексная амплитуда отраженной волны).

Для H -волны граничное условие $H_z = 0$ при $z = 0$ дает $H_z \sim \sin k_z z$. Добавив граничное условие $H_z = 0$ при $z = d$, получим $k_z = \pi p/d$, где p — целое число. Иными словами,

$$H_z = \psi_H(x, y) \sin \frac{\pi p z}{d} e^{-i\omega t} \quad (p = 1, 2, \dots),$$

где $\psi_H(x, y)$ — некоторая функция от x и y .

В случае E -волны, как нетрудно убедиться из общих формул для полей в волноводе, условие обращения в нуль E_t на торцах волновода приводит к требованию обращения на них в нуль производной $\partial E_z / \partial z$. Поэтому

$$E_z = \psi_E(x, y) \cos \frac{\pi p z}{d} e^{-i\omega t} \quad (p = 0, 1, \dots),$$

где $\psi_E(x, y)$ — функция от x и y .

Мы видим, что в перегороженном волноводе величина k_z вместо непрерывной становится дискретной. Отсюда вытекает важный вывод относительно частоты колебаний ω . Так как $\omega = ck$, а $k^2 = k_t^2 + k_z^2$ и $k_t^2 = \omega_{mn}^2/c^2$, где ω_{mn} — граничная частота mn -колебания, то

$$\omega = \sqrt{\omega_{mn}^2 + \frac{\pi^2 c^2 p^2}{d^2}}.$$

Эта величина определяется тремя целыми числами m , n , p (при заданной форме волновода и заданном расстоянии между торцами) и ее можно обозначить ω_{mnp} . В частности, для прямоугольного волновода

$$\omega_{mnp} = \pi c \sqrt{\frac{m^2}{a^2} + \frac{n^2}{b^2} + \frac{p^2}{d^2}}.$$

Иными словами,

перегороженный волновод имеет дискретный спектр частот.

Перегороженный волновод — это частный случай полости, окруженной со всех сторон металлической оболочкой, и полученный результат является частным случаем общего результата, заключающегося в том, что электромагнитное поле в такой полости, если полость полностью изолирована от окружающего пространства, обладает дискретным спектром частот, зависящим от геометрии полости, причем каждая частота определяется тремя целыми числами. Полость эту называют *резонатором*, а частоты — *собственными частотами резонатора*. Резонаторы широко применяются в качестве колебательных контуров в технике СВЧ.

Если бы стенки полости обладали бесконечной проводимостью, то поле не проникало бы в них и электромагнитная энергия в резонаторе не убывала бы, а сохранялась. Иными словами, поле, раз попав в полость, колебалось бы с одной из собственных частот резонатора, не испытывая никакого затухания. Отсюда, в частности, даже не решая уравнений Максвелла для полости, можно сделать вывод, что амплитуды колебаний электрического и магнитного полей должны быть сдвинутыми по фазе на $\pi/2$ и относиться как $\sqrt{\mu_0/\epsilon_0}$.

В действительности, однако, электромагнитная волна проникает вглубь стенок резонатора, и здесь ее энергия превращается в джоулеву теплоту. По этой причине электромагнитная энергия, запасенная в резонаторе, постепенно уменьшается. Отношение энергии, запасенной в резонаторе, к энергии, теряемой на период колебаний, называют *добротностью резонатора*. Если $W(t)$ — энергия резонатора в момент времени t и Q — его добротность, то W изменяется по закону $dW/dt = -\omega_0 W/(2\pi Q)$, откуда

$$W(t) = W_0 \exp[-\omega_0 t/(2\pi Q)],$$

где ω_0 — резонансная частота и W_0 — начальное значение энергии резонатора. Отсюда далее можно сделать заключение, что поля в ре-

зонаторе изменяются с течением времени не по закону $\exp(-i\omega_0 t)$, а по закону

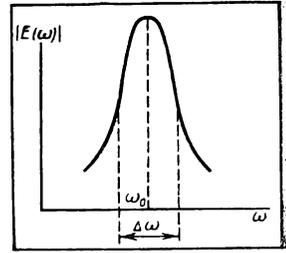
$$E(t) = E_0 \exp[-i\omega_0 t - \omega_0 t / (4\pi Q)].$$

Такое поле, не будучи строго гармоническим, может быть представлено в виде суперпозиции отдельных гармоник

$$E(t) = \int_{-\infty}^{\infty} E(\omega) \exp(-i\omega t) d\omega,$$

где амплитуда поля $E(\omega)$ частоты ω определяется формулой

$$E(\omega) = \int_{-\infty}^{\infty} E(t) \exp(i\omega t) dt / (2\pi).$$



14.6. Резонансная кривая

Зная $E(\omega)$, можно найти частотное распределение энергии — энергия в интервале частот $(\omega, \omega + d\omega)$ пропорциональна $|E(\omega)|^2 d\omega$. Легко убедиться, что $|E(\omega)|^2 \sim \{(\omega - \omega_0)^2 + [\omega_0 / (4\pi Q)]^2\}^{-1}$.

Таким образом, в поле резонатора вместо дискретной частоты ω_0 присутствует целый спектр частот. Интенсивность поля как функция частоты представлена на рис. 14.6. Эта величина максимальна при резонансной частоте ω_0 и убывает при увеличении $|\omega - \omega_0|$. Интенсивность поля вдвое меньше максимальной, если $\omega = \omega_0 \pm 1/2 \Delta\omega$, где $\Delta\omega = \omega_0 / (2\pi Q)$. Величина $\Delta\omega$ определяет эффективную ширину резонансной кривой. Таким образом,

добротность пропорциональна отношению резонансной частоты к ширине резонансной кривой.

Выясним теперь, чем определяется добротность резонатора. Так как потери энергии обусловлены проникновением поля в толщу стенки резонатора, то существенной величиной должна быть глубина δ проникновения поля в стенку резонатора. Эту величину называют *глубиной скин-слоя* (она определена в гл. 16). При известной δ легко оценить Q . Действительно, энергия, запасенная в резонаторе, пропорциональна его объему V , теряемая же энергия пропорциональна объему слоя стенок толщины δ , т. е. пропорциональна $S\delta$, где S — площадь поверхности стенок резонатора. Поэтому добротность должна определяться формулой $Q = GV / (S\delta)$, где G — безразмерный фактор, зависящий от геометрии резонатора и структуры резонансного поля. Мы не будем здесь приводить конкретных выражений для G .

ОСНОВНЫЕ ФОРМУЛЫ

Плотность энергии волны $\omega = \frac{1}{2} \epsilon_0 |E_0|^2 = \frac{1}{2\mu_0} |B_0|^2$

Эллиптическая полярзация $\left(\frac{E_x}{E_1}\right)^2 + \left(\frac{E_y}{E_2}\right)^2 = 1$

Поляризационная матрица и параметры Стокса
Полностью поляризованная волна

$$\rho = \frac{1}{2} \begin{pmatrix} 1 + \xi_3, & \xi_1 - i\xi_2 \\ \xi_1 + i\xi_2, & 1 - \xi_3 \end{pmatrix}$$

$$\xi_1 = \cos 2\beta \sin 2\gamma,$$

$$\xi_2 = \sin 2\beta, \quad \xi_3 = \cos 2\beta \cos 2\gamma.$$

Частично поляризованная волна

$$\xi_1^2 + \xi_2^2 + \xi_3^2 = P^2 \leq 1$$

Уравнение для продольных составляющих полей

$$\Delta_{\perp} E_z + k_{\perp}^2 E_z = 0,$$

$$\Delta_{\perp} B_z + k_{\perp}^2 B_z = 0,$$

$$E_z|_C = 0, \quad \partial B_z / \partial n|_C = 0$$

Собственные частоты прямоугольного резонатора

$$\omega_{mnp} = \pi c \sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2 + \left(\frac{p}{d}\right)^2}$$

Затухание энергии в резонаторе

$$W(t) = W(0) \exp[-\omega_0 t / (2\pi Q)]$$



Глава 15. ИНТЕРФЕРЕНЦИЯ И КОГЕРЕНТНОСТЬ

15.1. Интенсивность суперпозиции монохроматических волн

В предыдущей главе было показано, что при суперпозиции двух бегущих плоских монохроматических волн, обладающих одинаковыми амплитудой и частотой, но распространяющихся в противоположных направлениях, возникает стоячая волна, интенсивность которой в отличие от интенсивностей исходных волн меняется от точки к точке, обращаясь в нуль в узлах и достигая максимума (в два раза превосходящего суммарную интенсивность исходных волн) в пучностях. Такое поведение интенсивности связано с тем, что

складываются напряженности полей, а не интенсивности.

Дело в том, что под интенсивностью понимается средняя по времени плотность потока энергии волны, а так как эта величина отличается от среднего (по времени) значения квадрата электрического поля $\langle \mathbf{E}^2 \rangle$ только множителем, то под интенсивностью волны можно понимать величину $J = \langle \mathbf{E}^2 \rangle$.

Если происходит суперпозиция двух монохроматических волн с напряженностями полей в некоторой точке $\mathbf{E}_1 = \mathbf{A}_1 \cos(\omega_1 t + \varphi_1)$ и $\mathbf{E}_2 = \mathbf{A}_2 \cos(\omega_2 t + \varphi_2)$, где $\mathbf{A}_1, \mathbf{A}_2$ — амплитуды и φ_1, φ_2 — начальные фазы (для простоты рассматриваются линейно поляризованные волны), то для интенсивностей обеих волн имеем $J_1 = 1/2 A_1^2$ и $J_2 =$

$= 1/2 A_2^2$. Суммарное поле в рассматриваемой точке есть $\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2$, и поэтому интенсивность результирующей волны

$$J = \langle \mathbf{E}^2 \rangle = \langle \mathbf{E}_1^2 \rangle + \langle \mathbf{E}_2^2 \rangle + 2 \langle \mathbf{E}_1 \mathbf{E}_2 \rangle = J_1 + J_2 + 2 \langle \mathbf{E}_1 \mathbf{E}_2 \rangle.$$

Мы видим, что она отличается от суммы интенсивностей исходных волн на величину

$$2 \langle \mathbf{E}_1 \mathbf{E}_2 \rangle = \mathbf{A}_1 \mathbf{A}_2 [\langle \cos(\psi_1 - \psi_2) \rangle + \langle \cos(\psi_1 + \psi_2) \rangle],$$

где $\psi_1 = \omega_1 t + \varphi_1$ и $\psi_2 = \omega_2 t + \varphi_2$. Ясно, что $\langle \cos(\psi_1 + \psi_2) \rangle = 0$. Что же касается величины $\langle \cos(\psi_1 - \psi_2) \rangle$, то она также равна нулю, если $\omega_1 \neq \omega_2$, и может отличаться от нуля только в случае равенства частот ω_1 и ω_2 . Мы приходим к выводу, что

при суперпозиции двух волн с неравными частотами интенсивности волн складываются:

$$J = J_1 + J_2 \quad (\omega_1 \neq \omega_2).$$

Ясно, что это утверждение справедливо и для большего числа накладываются волн. Если же волны имеют одинаковую частоту и разность фаз ($\varphi_1 - \varphi_2$) представляет собой строго определенную (а не случайную) величину, то интенсивность результирующей волны

$$J = J_1 + J_2 + J_{12}; \quad J_{12} = \mathbf{A}_1 \mathbf{A}_2 \cos(\varphi_1 - \varphi_2) = \\ = 2 \sqrt{J_1 J_2} \cos \chi \cos(\varphi_1 - \varphi_2),$$

где χ — угол между \mathbf{A}_1 и \mathbf{A}_2 . Таким образом, если $\chi \neq \pi/2$ и $\varphi = \varphi_1 - \varphi_2 \neq \pi/2$, то $J_{12} \neq 0$, так что интенсивность результирующей волны отличается от суммы интенсивностей исходных волн. Это явление называют *интерференцией* волн, а слагаемое J_{12} — *интерференционным членом*. Оно может меняться от $-2\sqrt{J_1 J_2}$ до $2\sqrt{J_1 J_2}$, поэтому интенсивность результирующей волны может меняться в пределах, определяемых неравенством

$$J_{\min} \leq J \leq J_{\max}, \quad J_{\min} = (\sqrt{J_1} - \sqrt{J_2})^2, \quad J_{\max} = (\sqrt{J_1} + \sqrt{J_2})^2.$$

Если $J_1 = J_2$, то $J_{\min} = 0$, $J_{\max} = 4J_1$.

Заметим, что если две волны (даже с одинаковой частотой) линейно поляризованы в двух взаимно перпендикулярных направлениях, т. е. $\chi = \pi/2$, то для них явление интерференции наблюдаться не будет. Это было экспериментально установлено Френелем и Араго, которые исследовали поляризационные свойства света, прошедшего через кристалл. Они сделали вывод, что *световые волны должны быть поперечными и не могут быть продольными*. Действительно, у поперечной волны могут быть два независимых состояния с взаимно перпендикулярными поляризациями, для продольной же волны возможно только одно состояние.

Мы для простоты рассматривали суперпозицию линейно поляризованных волн. Но интерференция возникает и при суперпозиции волн любых поляризаций. Более того, волны не должны быть обяза-

тельно плоскими. Чтобы пояснить это, вспомним, что в самом общем случае напряженность поля монохроматической волны может быть записана в форме

$$\mathbf{E}(\mathbf{r}, t) = \operatorname{Re} \{ \mathcal{E}(\mathbf{r}) e^{-i\omega t} \},$$

где $\mathcal{E}(\mathbf{r})$ — некоторый комплексный вектор, зависящий от координаты \mathbf{r} точки наблюдения и удовлетворяющий волновому уравнению $\Delta \mathcal{E}(\mathbf{r}) + (\omega^2/c^2)\mathcal{E}(\mathbf{r}) = 0$ с определенными граничными условиями. Интенсивность волны в точке \mathbf{r} определяется формулой

$$J(\mathbf{r}) = \langle (\mathbf{E}(\mathbf{r}, t))^2 \rangle = \frac{1}{4} \langle (\mathcal{E}^2 e^{-2i\omega t} + \mathcal{E}^{*2} e^{2i\omega t} + 2\mathcal{E}\mathcal{E}^*) \rangle,$$

а так как средние по времени значения первых двух слагаемых равны нулю, то

$$J(\mathbf{r}) = 1/2 \mathcal{E}(\mathbf{r}) \mathcal{E}^*(\mathbf{r}).$$

Пусть теперь происходит наложение двух волн, поля которых в точке \mathbf{r} равны $\operatorname{Re} \mathcal{E}_1 e^{-i\omega t}$ и $\operatorname{Re} \mathcal{E}_2 e^{-i\omega t}$. Тогда результирующее поле $\mathbf{E}(\mathbf{r}, t) = \operatorname{Re} \mathcal{E}(\mathbf{r}) e^{-i\omega t}$, где $\mathcal{E}(\mathbf{r}) = \mathcal{E}_1(\mathbf{r}) + \mathcal{E}_2(\mathbf{r})$, а интенсивность результирующей волны

$$J(\mathbf{r}) = 1/2 \mathcal{E}(\mathbf{r}) \mathcal{E}^*(\mathbf{r}) = 1/2 \mathcal{E}_1(\mathbf{r}) \mathcal{E}_1^*(\mathbf{r}) + 1/2 \mathcal{E}_2(\mathbf{r}) \mathcal{E}_2^*(\mathbf{r}) + 1/2 \{ \mathcal{E}_1(\mathbf{r}) \mathcal{E}_2^*(\mathbf{r}) + \mathcal{E}_1^*(\mathbf{r}) \mathcal{E}_2(\mathbf{r}) \}.$$

Если ввести интерференционный член

$$J_{12}(\mathbf{r}) = 1/2 \{ \mathcal{E}_1(\mathbf{r}) \mathcal{E}_2^*(\mathbf{r}) + \mathcal{E}_1^*(\mathbf{r}) \mathcal{E}_2(\mathbf{r}) \},$$

то формулу для $J(\mathbf{r})$ можно переписать в виде

$$J(\mathbf{r}) = J_1(\mathbf{r}) + J_2(\mathbf{r}) + J_{12}(\mathbf{r}).$$

Если функции $\mathcal{E}_1(\mathbf{r})$ и $\mathcal{E}_2(\mathbf{r})$ не являются полностью случайными, то интерференционный член, вообще говоря, отличен от нуля, так что имеет место интерференция.

Совокупность частот складывающихся волн называют спектром результирующей волны. Спектр не обязательно должен состоять из отдельных частот — быть дискретным. Он может быть и непрерывным, т. е. частоты могут непрерывно заполнять некоторый интервал.

В этом случае вместо суммы по частотам результирующая волна описывается интегралом по частотам — *интегралом Фурье*. Его удобно записывать в комплексной форме: $\mathbf{E} = \int_{-\infty}^{\infty} \mathbf{E}_\omega \exp(-i\omega t) d\omega$, где \mathbf{E}_ω — комплексная амплитуда волны с частотой ω . Формально здесь интегрирование производится по всем частотам, как положительным (которые только и имеют физический смысл), так и отрицательным. Для того чтобы поле \mathbf{E} было вещественным, должно выполняться условие, $\mathbf{E}_{-\omega} = \mathbf{E}_\omega^*$. Комплексная амплитуда \mathbf{E}_ω (ее называют компонентой

Фурье) связана с полем $\mathbf{E} = \mathbf{E}(t)$ соотношением $\mathbf{E}_\omega = \int_{-\infty}^{\infty} \mathbf{E}(t) \exp(i\omega t) \times dt/(2\pi)$. Интенсивность волн в интервале частот $(\omega, \omega + d\omega)$ пропорциональна $|\mathbf{E}_\omega|^2 d\omega$, полная же интенсивность (определяемая как интеграл по времени от квадрата поля) связана с \mathbf{E}_ω соотношением

$$\int_{-\infty}^{\infty} \mathbf{E}^2 dt = 4\pi \int_0^{\infty} |\mathbf{E}_\omega|^2 d\omega$$

(формула Парсеваля).

15.2. Некогерентные и когерентные пучки

В предыдущем параграфе мы показали, что волны с различными частотами не интерферируют между собой, монохроматические же волны с одинаковыми частотами могут интерферировать или, как иначе говорят, могут быть *когерентными*. Но для этого косинус разности фаз между волнами $\delta\varphi = \varphi_1 - \varphi_2$ должен быть отличен от нуля. Мы рассматривали тот случай, когда эта величина была строго определенной. Между тем фазы волн, испускаемых отдельными атомами, представляют собой случайные величины и никак не скоррелированы между собой. Поэтому и фазы волн, испускаемых реальными макроскопическими источниками, также являются случайными величинами. Поэтому возникает вопрос, как сделать, чтобы $\cos(\varphi_1 - \varphi_2)$ или, точнее говоря, среднее значение этой величины по всем возможным значениям фаз φ_1 и φ_2 не было тождественно равным нулю. До последнего времени считалось, что для этого обе волны должны создаваться не разными источниками, а одним общим источником: в этом случае стохастические (случайные) слагаемые в фазах φ_1 и φ_2 , имея общее происхождение, сократятся и разность фаз $\varphi_1 - \varphi_2$ будет определенной, детерминированной (или, точнее говоря, частично детерминированной) величиной. Но источники могут быть и разными, например если это лазеры одинаковой частоты.

Таким образом, мы приходим к важному выводу: чтобы наблюдать явление интерференции, следует пользоваться двумя волнами, создаваемыми общим источником (либо двумя однотипными лазерами). В случае света от одного источника это значит, что нужно взять пучок света искусственно разделить его на две части и заставить разделенные пучки приобрести в результате их движения различные фазы. Производя затем суперпозицию этих пучков, мы сможем наблюдать явление интерференции.

Пучки эти, однако, когерентны не полностью, т. е. интенсивность света при их суперпозиции не полностью определяется формулами предыдущего параграфа, относящимися к строго монохроматическим волнам с абсолютно детерминированной разностью фаз. Дело в том, что пучки создаются протяженными источниками и поэтому обязательно несут в себе элемент стохастичности. По этой причине их назы-

вают не когерентными, а *частично когерентными*. Кроме того, необходимо учитывать, что пучки никогда не могут быть строго монохроматическими. Это связано с тем, что электромагнитные волны, даже излученные каким-либо одним атомом, никогда не могут быть абсолютно монохроматическими. Таково принципиальное свойство материи.

Возьмем, например, отдельный атом, который, находясь в некотором состоянии, излучает свет. Если время жизни этого состояния определяется величиной τ , то степень немонохроматичности, т. е. разброс по частотам $\Delta\omega$ относительно некоторой средней частоты ω_0 , имеет порядок $\Delta\omega \sim 1/\tau$. Только при бесконечном времени жизни состояния свет был бы монохроматическим, при конечных же временах жизни, что всегда имеет место, он не является монохроматическим, причем степень немонохроматичности тем больше, чем меньше время жизни состояния. Поэтому электромагнитные волны не могут быть строго монохроматическими; в лучшем случае они могут быть квази-монохроматическими, для которых $\Delta\omega \ll \omega_0$.

Если рассматривать не отдельный атом, а реальный источник квазимонохроматических волн, то и для него всегда выполняется соотношение $\Delta\omega \sim 1/\tau$, где τ определяет длительность цуга волн (*волнового пакета*), испускаемых источником (см. гл. 16).

Величина τ играет важную роль в наблюдении явления интерференции. Дело в том, что разность фаз между разделенными пучками достигается вследствие их движения, т. е. определяется разностью их путей Δs . С другой стороны, свет за время существования цуга волн τ проходит расстояние ct . Это значит, что большие расстояния в рассматриваемой задаче не имеют смысла. Поэтому интерференцию разделенных и затем собранных пучков можно наблюдать только в том случае, если $\Delta s \ll ct$ или $\Delta s \ll l_c$, где $l_c = ct = c/\Delta\omega$. На этом основании величину l_c называют *длиной когерентности*.

Так как длина когерентности обратно пропорциональна степени немонохроматичности, то в опытах по исследованию интерференции допустимые разности хода пучков тем больше, чем монохроматичнее свет.

Разделение пучка на два когерентных (точнее говоря, частично когерентных) пучка производится двумя методами — либо исходный пучок делится на два, проходя через близко расположенные друг к другу отверстия в экране, либо исходный пучок делится на два на частично отражающих, частично пропускающих свет поверхностях. Рассмотрим оба эти метода.

15.3. Опыт Юнга

Первый метод был впервые применен Юнгом в его историческом опыте (1802), сыгравшем важную роль в доказательстве волновой природы света. В этом опыте свет от источника S (рис. 15.1) падал на экран, в котором было сделано два небольших отверстия S_1 и S_2 . Эти отверстия можно рассматривать как вторичные источники света, освещающие пространство за экраном. (Такой подход основан на общем принципе Гюйгенса—Френеля, который будет

сформулирован в гл. 21.) Вторичные источники света S_1 и S_2 являются когерентными, так как они происходят от общего источника (точнее говоря, они частично когерентны). Поэтому в пространственной области I , где пересекаются пучки света, создаваемые обоими вторичными источниками, наблюдается явление интерференции.

Чтобы выяснить вид интерференционной картины, предположим, что как источник S , так и отверстия S_1, S_2 — точечные и свет монохроматичен. При этих предположениях вопрос сводится к вычислению разности хода лучей от вторичных источников S_1, S_2 до точки наблюдения в области I . Рис. 15.2 разъясняет расчет разности хода. Плоскость Oxy , в которой расположена точка наблюдения $P(x, y)$, параллельна линии, соединяющей отверстия S_1 и S_2 . Отверстия предполагаются находящимися на одинаковых расстояниях от источника S . Поэтому интенсивность света в точке P определяется только разностью расстояний s_2 и s_1 между точками S_2, P и S_1, P . Эти расстояния составляют

$$s_1 = \sqrt{a^2 + y^2 + \left(x - \frac{d}{2}\right)^2},$$

$$s_2 = \sqrt{a^2 + y^2 + \left(x + \frac{d}{2}\right)^2},$$

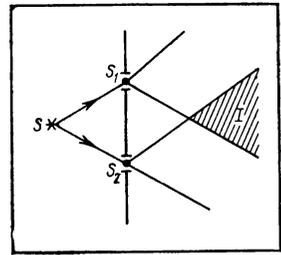
где d — расстояние между S_1 и S_2 , a — расстояние между линией S_1S_2 и плоскостью Oxy . Поэтому $s_2^2 - s_1^2 = 2xd$ и разность хода лучей $\Delta s = s_2 - s_1 = 2xd/(s_1 + s_2)$.

Интерференционная картина наблюдается только в том случае, когда отверстия S_1 и S_2 находятся близко друг от друга, так что выполняется неравенство $d \ll a$ (это связано с неточностью источников S_1 и S_2 , см. ниже). Кроме того, интенсивность света велика только в окрестности точки O . Поэтому s_1 и s_2 мало отличаются от a , т. е. $s_1 + s_2 \sim 2a$ и $\Delta s = xd/a$.

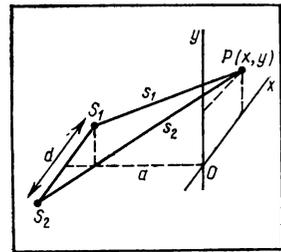
Мы можем теперь воспользоваться результатами предыдущего параграфа, считая, что интенсивности обоих суперпонируемых пучков J_1 и J_2 в точке $P(x, y)$ не отличаются от интенсивностей в точке O и притом равны между собой: $J_1 = J_2$. Разность фаз

$$\varphi \equiv \varphi_2 - \varphi_1 = 2\pi \frac{\Delta s}{\lambda} = 2\pi \frac{xd}{\lambda a},$$

где λ — длина волны света. Интенсивность света J в области I определяется формулой



15.1.
Схема опыта Юнга с точечным источником



15.2.
Расчет разности хода лучей в опыте Юнга

$$J = 2J_1(1 + \cos \varphi) = 4J_1 \cos^2 \left(\pi \frac{xd}{\lambda a} \right).$$

Мы видим, что интенсивность имеет одно и то же значение при $x = \text{const}$, т. е. вдоль прямой, параллельной оси y . Иными словами, прямые перпендикулярные плоскости SS_1S_2 , представляют собой линии одинаковой интенсивности. При этом максимум интенсивности наблюдается на прямых $x = m\lambda/d$ ($m = 0, \pm 1, \pm 2, \dots$), а минимум — на прямых $x = (m + 1/2)\lambda/d$. Таким образом, интерференционная картина в непосредственной близости от точки O имеет вид чередующихся светлых и темных полос — их называют *интерференционными полосами*.

В нашем расчете мы предполагали, что свет является монохроматическим и что размеры источника S и размеры отверстий в экране S_1 и S_2 очень малы. Посмотрим теперь, что произойдет с интерференцией при нарушении этих условий. Если свет немонохроматичен, то его можно представить в виде суперпозиции монохроматических компонент. Каждая такая компонента дает интерференционную картину рассмотренного только что вида, и все эти картины накладываются друг на друга, т. е. интенсивности отдельных картин складываются. Но расположение максимумов и минимумов зависит от длины волны, поэтому для отдельных компонент они не совпадают. В результате интерференционная картина для немонохроматического света будет не столь резкой, как для света монохроматического, а в большей или в меньшей мере смазанной.

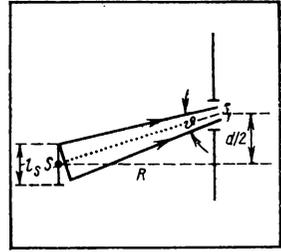
Чтобы яснее понять это, обратимся к формуле $x_m = m\lambda/d$, определяющей положение m -го максимума. Мы видим, что только положение центрального максимума $m = 0$ не зависит от длины волны, координаты же других максимумов пропорциональны λ . Поэтому если источник излучает в диапазоне $(\lambda - 1/2\Delta\lambda, \lambda + 1/2\Delta\lambda)$, то максимум с $m \neq 0$ размоется в полосу шириной $\Delta x_m = m\Delta\lambda/d$. Четкая интерференционная картина возникает, если ширина полосы мала по сравнению с расстоянием между m -м максимумом и $(m + 1)$ -м минимумом, т. е. для $\Delta\lambda/\lambda \ll 1/(2m)$. Мы видим, что условие квази-хроматичности света $\Delta\lambda/\lambda \ll 1$ является необходимым, но не достаточным условием наблюдения интерференции. В самом деле, при выполнении этого условия видны интерференционные полосы с малыми m ; общее число наблюдаемых полос m_{max} тем больше, чем меньше $\Delta\lambda/\lambda$ и $m_{\text{max}} \sim \lambda/(2\Delta\lambda)$.

Рассмотрим теперь условие точности источников, с которым также связана возможность получения четкой интерференционной картины. Если l_S — размеры источника S , то разность хода лучей от его крайних точек до отверстия S_1 составит $l_S \theta$ (рис. 15.3; угол θ предполагается малым). Для наблюдения четкой интерференционной картины эта величина должна быть малой по сравнению с длиной волны: $l_S \theta \ll \lambda$; тогда интерференционная картина мало отличается от интерференционной картины, соответствующей точечному источнику. Так как $\theta \sim d/R$, то это условие (его можно назвать условием квазиточечности источника S) может быть переписано в виде $l_S \ll \ll$

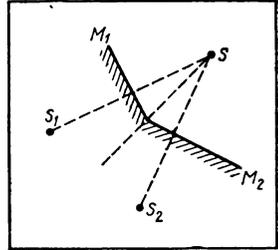
$\ll \lambda R/d$. Аналогичное условие должно выполняться и для вторичных источников. Именно: если l_1 — размер отверстия, то вторичный источник ведет себя как точечный, если $l_1 \theta_1 \ll \lambda$, где $\theta_1 \sim d/a$, т. е. если $l_1 \ll \lambda a/d$.

Так как для видимого света $\lambda \ll l_1$ и $\lambda \ll l_s$, то обязательным (хотя и далеко не достаточным) условием интерференции являются неравенства $R \gg d$ и $a \gg d$ (вторым из этих неравенств мы уже пользовались ранее).

В опыте Юнга когерентными источниками света являются два отверстия в непрозрачном экране, освещенные общим источником света. Существуют и другие методы получения когерентных источников, фактически эквивалентные методу Юнга. Один из них основан на использовании так называемых *зеркал Френеля* (рис. 15.4). В этом методе свет от точечного источника S падает на два плоских зеркала M_1 и M_2 , расположенных под небольшим углом друг к другу. Возникающие мнимые изображения S_1 и S_2 источника действуют как два когерентных источника.



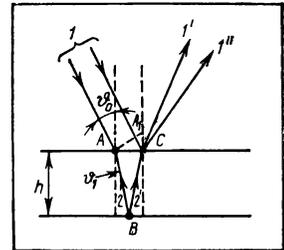
15.3. Роль размеров источника в опыте Юнга



15.4. Схема зеркал Френеля

15.4. Двухлучевая интерференция при отражении

Рассмотрим теперь интерференцию, возникающую при отражении светового пучка от границ плоскопараллельной стеклянной пластины (рис. 15.5). Падающий пучок I частично отражается от верхней поверхности пластины (луч I') и частично преломляется (луч 2). После отражения луча 2 от нижней поверхности пластины возникает луч $2'$, который далее, преломляясь на верхней поверхности, переходит в луч I'' . Интерференция происходит между лучами I' и I'' . Лучи эти, очевидно, когерентны, и интерференционная картина определяется разностью фаз между ними.



15.5. Отражение луча от границ плоскопараллельной пластины

При нахождении разности фаз следует учитывать два обстоятельства: во-первых, отличие скорости распространения света в среде (в данном случае — в стекле) от скорости света в вакууме и, во-вторых, изменения фазы на π при каждом отражении. В гл. 16 мы определим скорость света в среде и покажем, что она равна $v = c/N$, где N — так называемый показатель преломления, связанный с углами

падения θ_0 и преломления θ_1 соотношением $N = \sin\theta_0/\sin\theta_1$ (возникновение добавочной разности фаз π при отражении будет также разъяснено в гл. 16).

Зависимость полей плоской волны от координаты определяется экспоненциальным множителем $\exp(i\mathbf{k}\mathbf{r})$, причем в вакууме, как мы знаем, волновой вектор равен $\mathbf{k} = \mathbf{n}\omega/c$ (\mathbf{n} — единичный вектор в направлении распространения). Если волна распространяется в среде, то вместо c в эту формулу нужно подставить $v = c/N$, поэтому фаза волны в точке \mathbf{r} составит $\omega\mathbf{n}\mathbf{r}N/c$. Отсюда следует, что разность фаз между двумя лучами I и II определяется формулой

$$\varphi_I - \varphi_{II} = \frac{\omega}{c} \left(\sum_{(I)} \mathbf{n}\mathbf{r}N - \sum_{(II)} \mathbf{n}\mathbf{r}N \right) = \frac{\omega}{c} \left(\sum_{(I)} s_i N_i \right) - \frac{\omega}{c} \left(\sum_{(II)} s_i N_i \right),$$

где $s_i = (\mathbf{n}\mathbf{r})_i$ — длина пути луча на участке i , индекс i нумерует отдельные участки, проходимые лучом, и первая сумма распространяется на все участки i , проходимые лучом I (вообще говоря, в различных средах), а вторая сумма имеет аналогичный смысл для луча II . Величину $s^{\text{opt}} = \sum s_i N_i$ называют *оптической длиной пути* пучка. Учитывая, что $\omega/c = 2\pi/\lambda$ (λ — длина волны в вакууме), можно записать разность фаз между двумя не испытывавшими отражения лучами в виде $\varphi_I - \varphi_{II} = 2\pi/(s_I^{\text{opt}} - s_{II}^{\text{opt}})/\lambda$.

Но это еще не вся разность фаз, определяющая интерференцию пучков. Для того чтобы найти полную разность фаз, нужно, как говорилось выше, учесть еще изменение фазы на π при каждом отражении. Таким образом, если суммарное число отражений для обоих пучков есть p , то окончательная формула для разности фаз имеет вид

$$\delta\varphi = \varphi_I - \varphi_{II} + \pi p = \frac{2\pi}{\lambda} (s_I^{\text{opt}} - s_{II}^{\text{opt}}) + \pi p.$$

Применим теперь ту формулу для нахождения разности фаз лучей I' и I'' , изображенных на рис. 15.5. Разность оптических длин путей равна, очевидно, $N(ABC) - A_1C$, откуда

$$\varphi_{I'} - \varphi_{I''} = \frac{2\pi}{\lambda} \left(\frac{2hN}{\cos\theta_1} - 2h \operatorname{tg}\theta_1 \sin\theta_0 \right).$$

Но $\sin\theta_0/\sin\theta_1 = N$, поэтому

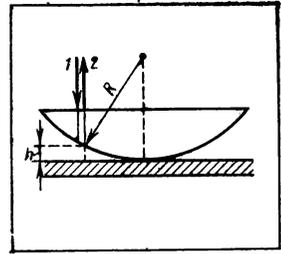
$$\varphi_{I'} - \varphi_{I''} = \frac{4\pi}{\lambda} hN \cos\theta_1.$$

Это и есть полная разность фаз, поскольку имеют место два отражения (по одному на каждой из поверхностей), а кратное 2π слагаемое в разности фаз можно отбросить. Отсюда легко найти условия максимума и минимума интенсивности: если $\delta\varphi = 2\pi n$ ($n = 0, 1, 2, \dots$), то интенсивность максимальна, если же $\delta\varphi = 2\pi(n + 1/2)$, то она будет минимальна.

Итак, условия экстремумов в интенсивности гласят:

$$2hN \cos \theta_1 = \begin{cases} n\lambda & (\text{максимум}) \\ (n + 1/2)\lambda & (\text{минимум}). \end{cases}$$

Они означают, что (для определенной частоты света и определенной толщины пластины) экстремумы возникают при вполне определенных значениях угла θ_1 , т. е. вполне определенных значениях угла падения θ_0 (в этом случае говорят об *интерференционных линиях*, или *полосах равного наклона*). Если же толщина h для разных точек пластины разная (пластина не строго плоскопараллельная), то экстремумы возникают при довольно произвольных углах падения; в интерференции при этом играют роль отражения от различных участков пластины, имеющих разную толщину, но удовлетворяющих условиям интерференции (в этом случае говорят об *интерференционных линиях*, или *полосах равной толщины*).



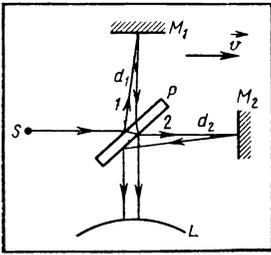
15.6. Когерентные лучи при наблюдении колец Ньютона

Интерференционные линии равной толщины можно наблюдать при падении света на сферическую линзу, соприкасающуюся в воздухе с плоской поверхностью стекла (рис. 15.6). В этом случае речь идет об интерференции между лучами 1 и 2. Разность хода лучей равна $2h$, где h — толщина воздушного слоя в месте падения луча, связанная с радиусом линзы R и расстоянием r между местом падения луча и точкой соприкосновения линзы с пластиной соотношением $r = \sqrt{h(2R-h)} \approx \sqrt{2hR}$ (предполагается, что $h \ll R$).

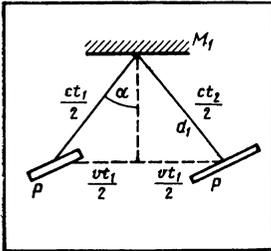
Разность фаз рассматриваемых лучей равна $\delta\varphi = 4\pi h/\lambda + \pi$ (последнее слагаемое возникает из-за того, что имеет место одно отражение от стеклянной плоскости). Максимумы интенсивности появятся при $\delta\varphi = 2\pi n$ ($n = 0, 1, 2, \dots$), а минимумы — при $\delta\varphi = 2\pi(n + 1/2)$; в первом случае $h = (2n + 1)\lambda/4$, а во втором $h = n\lambda/2$. Соответствующие радиусы r составляют для минимумов $r = \sqrt{nR\lambda}$, а для максимумов $r = \sqrt{(n + 1/2)R\lambda}$. Возникающие интерференционные линии называют *кольцами Ньютона*. Ньютон наблюдал их около 1670 г. и с их помощью впервые определил длину волны света. Заметим, что точка соприкосновения линзы с пластиной кажется темной, что связано с изменением фазы волны при отражении на π .

15.5. Опыт Майкельсона

Явление интерференции находит себе многочисленные практические применения: при сравнении эталонов, при проверке качества поверхностей и т. п. Мы не будем на этом останавливаться, а рассмотрим лишь одно применение принципиального характера, а именно покажем, как можно было бы, используя это явление, обнаружить движение Земли относительно «светоносного эфира»,



15.7.
Схема опыта Майкельсона



15.8.
К расчету времени t_1 в опыте Майкельсона

а точнее, выяснить, зависит ли скорость света от скорости движения Земли.

С этой целью Майкельсоном в 1881 г. был поставлен следующий опыт. Свет от источника S (рис. 15.7) фокусируется на полупрозрачную посеребренную стеклянную пластину P , которая делит пучок на два луча 1 и 2 , перпендикулярные друг другу и идущие к зеркалам M_1 и M_2 . После отражения лучи попадают в линзу L , где наблюдается их интерференция.

Будем исходить из предположения, что существует преимущественная система отсчета, связанная с эфиром, в котором свет распространяется со скоростью c . Если Земля вместе с прибором движется со скоростью v , то скорость света относительно прибора на пути от P к M_2 составит $c + v$, а на обратном пути будет $c - v$. Поэтому свет проходит путь от P к M_2 и обратно за время

$$t_2 = d_2 \left[(c - v)^{-1} + (c + v)^{-1} \right] = \frac{2d_2}{c} \left(1 - \frac{v^2}{c^2} \right)^{-1}.$$

Определим теперь время t_1 , необходимое свету, чтобы пройти путь от P к M_1 и обратно. В системе координат, связанной с эфиром, имеют место геометрические соотношения, изображенные на рис. 15.8. Ясно, что $t_1 = 2d_1 / (c \cos \alpha)$, где $\sin \alpha = v/c$. Поэтому $t_1 = (2d_1/c)(1 - v^2/c^2)^{-1/2}$. Разность времен прохождения всего пути до L составит

$$\begin{aligned} \Delta t = t_2 - t_1 &= \frac{2}{c} \left[d_2 \left(1 - \frac{v^2}{c^2} \right)^{-1} - d_1 \left(1 - \frac{v^2}{c^2} \right)^{-1/2} \right] \approx \\ &\approx \frac{2}{c} \left[d_2 - d_1 + \frac{v^2}{c^2} \left(d_2 - \frac{1}{2} d_1 \right) \right] \end{aligned}$$

(мы учли, что $v \ll c$).

Повернем теперь прибор на 90° . Тогда времена прохождения обоих лучей составят $t'_2 = (2d_2/c)(1 - v^2/c^2)^{-1/2}$, $t'_1 = (2d_1/c)(1 - v^2/c^2)^{-1}$, а разность времен

$$\Delta t' = \frac{2}{c} \left[d_2 - d_1 + \frac{v^2}{c^2} \left(\frac{d_2}{2} - d_1 \right) \right].$$

Так как $\Delta t' \neq \Delta t$, то при повороте прибора должен возникнуть сдвиг интерференционных полос

$$\delta = c(\Delta t' - \Delta t) = (d_1 + d_2)v^2/c^2.$$

Скорость Земли составляет $v \sim 3 \cdot 10^4$ м/с, так что $v^2/c^2 \sim 10^{-8}$. Для $d_1 + d_2 \sim 3$ м получим $\delta \sim 3 \cdot 10^{-8}$ м = 30 нм, что в 10 раз

меньше длины волны видимого света. Такое смещение интерференционных полос прибор (называемый *интерферометром Майкельсона*) обнаружить позволял, однако никакого смещения обнаружено не было. В дальнейшем были построены усовершенствованные приборы типа интерферометра Майкельсона и с их помощью был подтвержден прежний результат — сдвига интерференционных полос обнаружено не было. Объяснение этому удалось получить только в рамках специальной теории относительности, согласно которой не существует избранной системы отсчета, связанной с эфиром, и скорость света во всех инерциальных системах отсчета одинакова и равна c (см. гл. 7).

15.6. Корреляционная функция поля

В предыдущих параграфах мы видели, что резкая интерференционная картина в интерференционных опытах возникает только при строго определенной разности фаз обоих световых пучков. Размытость картины свидетельствует о неопределенности разности фаз, т. е. о случайном, стохастическом характере этой величины, а следовательно, и о стохастическом характере поля в целом. Такая ситуация отражает статистические свойства источника поля, а именно неопределенность и отсутствие корреляции в поведении зарядов, создающих поле (а также квантовые свойства самого поля, см. ниже). Поэтому изучение интерференции света приводит к более общей проблеме статистических свойств электромагнитного поля. Немного мы уже сталкивались с этой проблемой в гл. 14 при изучении частично поляризованного света. Мы рассматривали там произведения различных компонент вектора поляризации волны и усредняли их по тем параметрам, которые характеризуют случайный характер вектора поляризации. В результате усреднения мы получили поляризационную матрицу, которая может служить для описания свойств частично поляризованного света.

Эту идею мы можем теперь обобщить, рассматривая не вектор поляризации, а сам вектор поля, например электрического, и усредняя произведения компонент поля по параметрам, характеризующим его стохастичность. Такого рода усредненные произведения носят название корреляционных функций. Удобно пользоваться полями в комплексной форме и определять корреляционные функции следующим образом:

$$G_{\mu\nu}(\mathbf{r}_1 t_1, \mathbf{r}_2 t_2) = \langle E_{\mu}(\mathbf{r}_1 t_1) E_{\nu}^*(\mathbf{r}_2 t_2) \rangle,$$

где угловые скобки служат для обозначения усреднения и индексы μ и ν — для обозначения компонент поля ($\mu, \nu = 1, 2, 3$). Корреляционные функции зависят от двух координат $\mathbf{r}_1, \mathbf{r}_2$ и двух моментов времени t_1, t_2 . Обратим внимание на то, что в корреляционную функцию входит произведение компонент поля и комплексно-сопряженного поля. При таком определении оказывается, что счетчик фотонов реагирует именно на корреляционную функцию. Дело в том, что работа счетчика фотонов основана на фотоэффекте (см. гл. 24), т. е. счетчик регистрирует фотоэлектроны: число же последних (зарегистрирован-

ное между моментами t и $t + dt$) пропорционально $G(\mathbf{r}t, \mathbf{r}t)$, где \mathbf{r} — место расположения счетчика.

Разъясним теперь смысл усреднения, содержащегося в определении корреляционных функций. Напомним для этого, что поле можно представить в виде суперпозиции плоских волн $\mathbf{e}_k \exp(i\mathbf{k}\mathbf{r} - i\omega t)$ или какой-либо иной системы волн $\mathbf{u}_p(\mathbf{r}) \exp(-i\omega_p t)$ (могущих, например, существовать в каком-либо резонаторе или волноводе):

$$\mathbf{E}(\mathbf{r}, t) = \sum_p C_p \mathbf{u}_p(\mathbf{r}) \exp(-i\omega_p t),$$

где C_p — комплексная амплитуда p -волны. Эта величина может быть и строго определенной и случайной в зависимости от характера источника поля. В первом случае, т. е. в случае полностью детерминированного поля, введенная выше корреляционная функция фактически не содержит никакого усреднения, т. е. $G_{\mu\nu}(\mathbf{r}_1 t_1, \mathbf{r}_2 t_2) = E_\mu(\mathbf{r}_1, t_1) E_\nu^*(\mathbf{r}_2, t_2)$, и является некоторой функцией суперпозиционных коэффициентов, т. е. $G_{\mu\nu}(\mathbf{r}_1 t_1, \mathbf{r}_2 t_2) \equiv F(C_1, C_2, \dots)$. Во втором случае, т. е. в случае случайного поля, суперпозиционные коэффициенты являются случайными величинами и функция F должна быть усреднена по возможным значениям C_1, C_2, \dots с некоторой вероятностью $\omega(C_1, C_2, \dots)$. Таким образом, фактическое вычисление величины $G_{\mu\nu}(\mathbf{r}_1 t_1, \mathbf{r}_2 t_2)$ становится возможным только тогда, когда известна вероятность $\omega(C_1, C_2, \dots)$ различных значений C_1, C_2, \dots .

В ряде случаев случайное поле представляет собой стационарную стохастическую величину, удовлетворяющую так называемому *условию эргодичности*. Корреляционные функции тогда зависят от разности времен $t_2 - t_1 \equiv \tau$ и могут быть определены как средние по времени:

$$G_{\mu\nu}(\mathbf{r}_1, \mathbf{r}_2, \tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T E_\mu(\mathbf{r}_1 t_1) E_\nu^*(\mathbf{r}_2 t_1 + \tau) dt_1.$$

Подчеркнем, что условие эргодичности, т. е. равенство среднего по времени и среднего, вычисленного с помощью вероятностной функции $\omega(C_1, C_2, \dots)$ (последнее называют *средним по ансамблю полей*), выполняется далеко не для всех полей. Если распределение случайных коэффициентов разложения поля C_1, C_2 является гауссовым, то поле эргодическое, т. е. для него среднее по ансамблю совпадает со средним по времени. Напомним, что гауссово распределение — это распределение, для которого вероятность $\omega(x)$ того, что случайная величина x заключена в интервале $(x, x + dx)$, есть

$$\omega(x) dx = \text{const} \exp[-x^2/(2\langle x^2 \rangle)] dx,$$

где $\langle x^2 \rangle$ — среднее значение x^2 (const определяется из условия $\int \omega(x) dx = 1$, она равна $(2\pi\langle x^2 \rangle)^{-1/2}$).

До сих пор при анализе понятия корреляционной функции мы исходили из представления о классическом характере электромагнитного поля. Между тем свет имеет квантовую природу (см. гл. 24) и

нам необходимо отразить квантование поля в корреляционной функции. Заметим для этого, что корреляционная функция квадратична по полю, т. е. линейна по энергии поля; поэтому она должна быть линейной функцией чисел фотонов. Кроме того, корреляционная функция квадратична относительно собственных функций поля $u_k(\mathbf{r})$. В силу этих причин вид ее определяется фактически однозначно:

$$G_{\mu\nu}(\mathbf{r}_1 t_1, \mathbf{r}_2 t_2) = \frac{1}{2} \sum_k \hbar \omega_k \langle n_k \rangle u_{k\mu}(\mathbf{r}_1) u_{k\nu}^*(\mathbf{r}_2) \exp\{-i\omega_k(t_1 - t_2)\},$$

где $\langle n_k \rangle$ — среднее число фотонов сорта k , $\hbar \omega_k \langle n_k \rangle$ — средняя энергия k -й моды поля и $u_{k\mu}(\mathbf{r})$ — собственные функции поля, удовлетворяющие условию нормировки

$$\int u_{k\mu}^*(\mathbf{r}) u_{k'\mu'}(\mathbf{r}) dV = \delta_{kk'} \delta_{\mu\mu'}$$

($\delta_{pp'} = 1$ при $p = p'$ и $\delta_{pp'} = 0$ при $p \neq p'$; множитель $1/2$ возникает из-за того, что мы пользуемся комплексной формой поля). Усреднение в этой формуле содержится в величине $\langle n_k \rangle$: число фотонов в каждой моде поля может быть любым целым числом, в формулу же входит среднее число фотонов.

Если поле является термодинамически равновесным (см. гл. 24), что соответствует равновесному состоянию источника, то $\langle n_k \rangle$ имеет вид **р а с п р е д е л е н и я П л а н к а**:

$$\langle n_k \rangle = \{\exp[\hbar\omega_k/(k_B T)] - 1\}^{-1}$$

(T — температура газа фотонов, k_B — постоянная Больцмана).

В оптике мы имеем обычно дело с непрерывным спектром излучения и длинами, много большими длины волны. В этом случае $u_k(\mathbf{r}) = V^{-1/2} \mathbf{e} \exp(i\mathbf{k}\mathbf{r})$ — плоские волны, где \mathbf{e} — вектор поляризации волны и V — занимаемый полем объем. Суммирование по различным модам в формуле для G может быть при этом заменено интегрированием: $\sum_k \rightarrow \int V d^3k / (2\pi)^3$, так что в результате получается

$$G_{\mu\nu}(\mathbf{r}_1 t_1, \mathbf{r}_2 t_2) = \frac{1}{2(2\pi)^3} \sum_{\lambda} \int d^3k e_{\mu}^{(\lambda)} e_{\nu}^{(\lambda)*} \langle n_{k\lambda} \rangle \hbar \omega \exp\{i\mathbf{k}(\mathbf{r}_1 - \mathbf{r}_2) - i\omega(t_1 - t_2)\},$$

где $e^{(\lambda)}$ — вектор поляризации, ω — частота и $\langle n_{k\lambda} \rangle$ — среднее число фотонов моды $\mathbf{k}\lambda$.

Величина $(2\pi)^{-3} \sum_{\lambda} \langle n_{k\lambda} \rangle \hbar \omega d^3k$ имеет простой физический смысл — она представляет собой плотность энергии поля в интервале волновых векторов $(\mathbf{k}, \mathbf{k} + d\mathbf{k})$. Если, выполнив интегрирование по углам вектора \mathbf{k} , перейти к *спектральной плотности энергии* $\omega(\omega)$ (т. е. плотности энергии, отнесенной к единичному интервалу частот вблизи частоты ω), то из формулы для $G_{\mu\nu}$ мы получим

$$\sum_{\mu} G_{\mu\mu}(\mathbf{r}t_1, \mathbf{r}t_2) = \frac{1}{2} \int_0^{\infty} \omega(\omega) \exp\{-i\omega(t_1 - t_2)\} d\omega.$$

Последнюю формулу можно обратить и выразить спектральную плотность излучения через корреляционную функцию:

$$\omega(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} \sum_{\mu} G_{\mu\mu}(\mathbf{r}0, \mathbf{r}t) \exp(i\omega t) dt.$$

Рассмотрим теперь в качестве примера плоскую световую волну, распространяющуюся вдоль оси y и обладающую определенной поляризацией. Пусть поле не зависит от координат x, z . Тогда корреляционная функция зависит только от координаты y и ее можно обозначить $G(y_1 t_1, y_2 t_2)$. В интегральном представлении G фактически содержится только интегрирование по k_y ; интегрирование же по k_x и k_z сводится просто к умножению на $(2\pi)^2 L^{-2}$ (L — поперечный размер светового пучка). Поэтому

$$\begin{aligned} G(y_1 t_1, y_2 t_2) &= \\ &= \frac{1}{4\pi L^2} \int \langle n_k \rangle \hbar \omega \exp\{ik_y(y_1 - y_2) - i\omega(t_1 - t_2)\} dk_y, \end{aligned}$$

где $\omega = ck_y$. Входящая сюда величина $\langle n_k \rangle \hbar \omega$ представляет собой энергию k_y -моды поля. Она есть некоторая функция $W(\omega)$ частоты, и мы предположим, что эта функция имеет острый максимум при определенной частоте ω_0 и определяется формулой

$$W(\omega) = \frac{1}{\pi} 2\gamma [(\omega - \omega_0)^2 + \gamma^2]^{-1}$$

(в этом случае говорят, что спектральная линия имеет *лоренцеву форму*). Максимум функции W равен $2(\pi\gamma)^{-1}$, в точках же $\omega = \omega_0 \pm \gamma$ эта функция вдвое меньше максимального значения (поэтому величину 2γ условно считают *эффективной шириной спектральной линии*). Функция W нормирована на единицу, т. е.

$$\int_0^{\infty} W(\omega) d\omega = 1.$$

Положив теперь

$$\frac{\langle n_k \rangle \hbar \omega_k}{cL^2} = U \frac{2\gamma}{(\omega - \omega_0)^2 + \gamma^2}$$

(U — некоторая постоянная), можно сказать, что U — характеризует интенсивность рассматриваемого светового луча. Подставляя это выражение в формулу для G и считая, что $\gamma \ll \omega_0$, получим

$$G(y_1 t_1, y_2 t_2) = \frac{\gamma}{2\pi} U e^{i\omega_0 s} \int_{-\infty}^{\infty} \frac{e^{i\omega' s}}{\omega'^2 + \gamma^2} d\omega',$$

где $s = t_2 - t_1 - (y_2 - y_1)/c$ и $\omega' = \omega - \omega_0$ (при $\gamma \ll \omega_0$ пределы интегрирования можно считать бесконечными). Выполнив интегрирование, найдем

$$G(y_1 t_1, y_2 t_2) = \frac{1}{2} U \exp(i\omega_0 s - \gamma |s|).$$

Мы видим, что G зависит только от одной переменной s . При $s = 0$ корреляционная функция достигает максимума, равного $\frac{1}{2}U$. Эта величина есть среднее значение квадрата комплексного поля $\mathbf{E}(\mathbf{r}, t)$.

В заключение этого параграфа приведем некоторые общие свойства корреляционной функции. Прежде всего это эрмитовость: $G_{\mu\nu}(x_1, x_2)^* = G_{\nu\mu}(x_2, x_1)$ (x — сокращенное обозначение пространственно-временной точки: $x \equiv \mathbf{r}, t$; поэтому величины $G_{\mu\mu}(x, x)$ вещественны). Далее,

$$|G_{\mu\nu}(x_1, x_2)|^2 \leq G_{\mu\mu}(x_1, x_1) G_{\nu\nu}(x_2, x_2).$$

15.7. Степень когерентности

Покажем теперь, что понятие когерентности теснейшим образом связано с корреляционной функцией. Мы говорили, что лучи света когерентны, если они интерферируют. Поэтому нужно разъяснить, какое отношение имеет интерференция к корреляционной функции. Для этого проще всего обратиться к опыту Юнга (см. рис. 15.2). Если размеры отверстий S_1 и S_2 малы, а расстояния от S_1 и S_2 до точки наблюдения P велики по сравнению с длиной волны, то поле $\mathbf{E}(\mathbf{r}, t)$ в точке P в момент времени t может быть представлено в виде суперпозиции полей $\mathbf{E}(\mathbf{r}_1, t_1)$ и $\mathbf{E}(\mathbf{r}_2, t_2)$ в точках S_1 и S_2 в более ранние моменты времени $t_1 = t - s_1/c$, $t_2 = t - s_2/c$ (см. гл. 21):

$$\mathbf{E}(\mathbf{r}, t) = K_1 \mathbf{E}(\mathbf{r}_1, t_1) + K_2 \mathbf{E}(\mathbf{r}_2, t_2),$$

где K_1 и K_2 — постоянные, зависящие от геометрии установки.

Интенсивность света в точке P определяется квадратом модуля поля $\mathbf{E}(\mathbf{r}, t)$. Если в этой точке находится идеальный детектор фотонов, реагирующий, как мы уже говорили, на корреляционную функцию поля $G(\mathbf{r}t, \mathbf{r}t)$, то скорость счета этого прибора

$$G(\mathbf{r}t, \mathbf{r}t) = |K_1|^2 G(\mathbf{r}_1 t_1, \mathbf{r}_1 t_1) + |K_2|^2 G(\mathbf{r}_2 t_2, \mathbf{r}_2 t_2) + 2\text{Re}[K_1 K_2^* G(\mathbf{r}_1 t_1, \mathbf{r}_2 t_2)]$$

(мы полагаем для простоты, что поле характеризуется определенной поляризацией, и не выписываем у G индексов μ, ν). Первое слагаемое здесь пропорционально числу фотонов, регистрируемых счетчиком в том случае, когда открыто только отверстие S_1 , отверстие же S_2 закрыто; аналогичный смысл имеет второе слагаемое — оно определяет скорость счета фотонов, если открыто только второе отверстие. Наконец, третье слагаемое, определяющее отличие в счете детектора по сравнению с тем, что было бы, если бы числа фотонов от обоих отверстий просто суммировались, соответствует интерференции волн,

идущих от обоих отверстий. Как мы видели, это слагаемое (которое можно назвать интерференционным) определяется корреляционной функцией $G(\mathbf{r}_1 t_1, \mathbf{r}_2 t_2)$. Поэтому именно эта функция описывает когерентность.

Мы можем теперь забыть о световых пучках — ведь корреляционная функция $G(\mathbf{r}_1 t_1, \mathbf{r}_2 t_2)$ определяется только полем и его статистическими свойствами. Поэтому можно говорить о когерентности самого поля и характеризовать его функцией $G(x_1, x_2)$, где $x \equiv (\mathbf{r}, t)$. Более точно, мы введем понятие *степени когерентности поля*, понимая под этим величину

$$g(x_1, x_2) = \frac{G(x_1, x_2)}{\sqrt{G(x_1, x_1) G(x_2, x_2)}}.$$

Эта величина, как и $G(x_1, x_2)$, является, вообще говоря, комплексной, и ее можно представить в виде $g(x_1, x_2) = |g(x_1, x_2)| \exp[i\varphi(x_1, x_2)]$, где φ — фаза степени когерентности. Так как $G(x_1, x_2) \leq \sqrt{G(x_1, x_1) G(x_2, x_2)}$, то $g(x_1, x_2) \leq 1$. Величину $g(x_1, x_2)$ называют еще *комплексной степенью когерентности*, а также *нормированной корреляционной функцией*.

Чтобы разъяснить, почему функцию $g(x_1, x_2)$ называют степенью когерентности, свяжем с ней *резкость*, или *контрастность*, *интерференционной картины*. Для простоты будем считать, что коэффициенты K_1 и K_2 одинаковы. Тогда скорость счета фотонов пропорциональна

$$I = G(x_1, x_1) + G(x_2, x_2) + 2 \sqrt{G(x_1, x_1) G(x_2, x_2)} |g(x_1, x_2)| \cos \varphi(x_1, x_2).$$

Величины $G(x_1, x_1)$, $G(x_2, x_2)$ и $g(x_1, x_2)$ медленно изменяются при смещении счетчика, и их можно считать просто постоянными. Поэтому максимумы и минимумы в показаниях счетчика соответствуют значениям $\cos \varphi(x_1, x_2) = \pm 1$:

$$I_{\max} = G(x_1, x_1) + G(x_2, x_2) + 2 \sqrt{G(x_1, x_1) G(x_2, x_2)} |g(x_1, x_2)|,$$

$$I_{\min} = G(x_1, x_1) + G(x_2, x_2) - 2 \sqrt{G(x_1, x_1) G(x_2, x_2)} |g(x_1, x_2)|.$$

Параметр, который характеризовал бы контрастность интерференционной картины, можно определить как

$$v = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}$$

(его называют *видностью* полос). В простейшем случае, когда $G(x_1, x_1) = G(x_2, x_2)$,

$$v = \frac{2 |g(x_1, x_2)|}{1 + |g(x_1, x_2)|}.$$

Таким образом, v определяется модулем степени когерентности. Если последний равен единице, то $v = 1$. В этом случае интенсивность

(т. е. число фотонов, регистрируемых счетчиком) изменяется в пределах от нуля до $4G(x_1, x_1)$. Именно в этом случае мы говорили о когерентности лучей, но правильнее говорить, что при $|g(x_1, x_2)| = 1$ имеет место полная когерентность, и количественно характеризовать когерентность степенью когерентности.

Корреляционная функция $G(x_1, x_2)$, а следовательно, и степень когерентности $g(x_1, x_2)$ отражают статистические свойства поля.

Для вполне детерминированного поля усреднение в определении корреляционной функции отпадает: $G(x_1, x_2) = E(x_1)E^*(x_2)$. В этом случае, очевидно, степень когерентности по модулю равна единице, т. е. поле является полностью когерентным. Справедливо и обратное утверждение: если $|g(x_1, x_2)| = 1$, то корреляционная функция факторизуется, т. е. может быть представлена в виде $G(x_1, x_2) = A(x_1)A^*(x_2)$. Величины $A(x_1)$, $A(x_2)$ в этом случае и представляют собой значения поля соответственно в точках x_1 и x_2 .

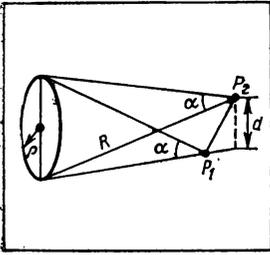
Говоря в этом параграфе о когерентности поля, мы не считали его монохроматическим. Между тем, рассматривая когерентные пучки в предшествующих параграфах, мы подчеркивали необходимость их монохроматичности. Противоречия здесь, однако, нет. Дело в том, что обычно рассматриваются стационарные источники поля. В этом случае $G(x_1, x_2)$ является функцией разности времен $t_1 - t_2$, а не времен t_1 и t_2 порознь: $G(x_1, x_2) = G(\mathbf{r}_1, \mathbf{r}_2, t_1 - t_2)$. Пусть теперь имеет место факторизация, т. е. $G(t_1 - t_2) = E(t_1)E^*(t_2)$ (пространственные координаты не выписаны). Легко показать, что единственной функцией, удовлетворяющей этому функциональному соотношению, является монохроматическая волна: $E(t) \sim \exp(-i\omega t)$.

Чтобы яснее представить себе структуру величины $g(x_1, x_2)$, определим степень когерентности для луча с лоренцевой формой спектральной линии. В предыдущем параграфе мы нашли корреляционную функцию в этом случае: $G(x_1, x_2) = \frac{1}{2} U \exp(i\omega_0 s - \gamma|s|)$, где $s = t_2 - t_1 - (y_2 - y_1)/c$, (напомним, что свет предполагался распространяющимся вдоль оси y). Так как $G(x_1, x_1) = G(x_2, x_2) = \frac{1}{2} U$, то степень когерентности

$$g(x_1, x_2) = \exp(i\omega_0 s - \gamma|s|),$$

так что $|g(x_1, x_2)| = \exp(-\gamma|s|)$. Таким образом, $|g|$ близко к единице, если $\gamma|s| \ll 1$; именно при этом условии наблюдается резкая интерференционная картина. Но условие $\gamma|s| \ll 1$ нам уже известно. Действительно, величина γ представляет собой ширину спектральной линии, т. е. степень немонохроматичности пучка, поэтому для введенной в § 15.2 длины когерентности l_c имеем $l_c \sim c/\gamma$. Но величина s представляет собой в рассматриваемом случае разность оптических путей двух интерферирующих лучей, поэтому условие $\gamma|s| \ll 1$ совпадает с полученным в § 15.2 условием наблюдения контрастной интерференционной картины.

Заметим, что для источников света, использующих газовый разряд, $\gamma \sim 10^9 \text{ с}^{-1}$, чему соответствует длина когерентности $l_c \sim 0,3 \text{ м}$. Для лазеров может быть $\gamma \sim 10 \text{ с}^{-1}$, чему соответствует длина когерентности $l_c \sim 30\,000 \text{ км}$ ($1/10$ расстояния от Земли до Луны!).



15.9.

К анализу когерентности поля от светящегося диска

Если $\gamma|s| \ll 1$, то $g \simeq \exp(i\omega_0 s)$, т. е. $|g| = 1$ при этом функция G факторизуется:

$$G(x_1, x_2) = 1/2 U f(x_1) f^*(x_2),$$

$$f(x) = \exp\{-i\omega_0(t - y/c)\}.$$

Этот пример показывает, что оба условия когерентности: $|g| = 1$ и факторизуемость корреляционной функции — эквивалентны. Но условие $|g| = 1$ выполняется здесь только приближенно (когда $\gamma|s| \ll 1$) и фактически всегда может выполняться только приближенно, так же как и условие факторизуемости корреляционной функции. Это значит, что полная когерентность является понятием идеализированным и о ней можно говорить только приближенно.

Рассмотрим еще один пример — интерференцию света, излучаемого протяженным источником. Пусть источником света является светящийся диск радиуса ρ . Будем интересоваться, в какой мере когерентности поля в двух близких точках P_1 и P_2 , находящихся на большом расстоянии R от источника (рис. 15.9; здесь α — угол, под которым радиус источника виден из точек P_1 и P_2). С этой целью нам нужно найти напряженности полей, создаваемые в точках P_1 и P_2 каждым элементом, т. е. каждым атомом диска. Эти поля можно, очевидно, записать в виде сферических волн:

$$E_M(\mathbf{r}, t) = A_M \frac{\exp\{i\omega(t - |\mathbf{r} - \rho|/c)\}}{|\mathbf{r} - \rho|},$$

где A_M определяет амплитуду волны, $\mathbf{r} = \mathbf{r}_1, \mathbf{r}_2$ — координаты точек P_1, P_2 и ρ — координаты точки M , в которой находится излучающий атом (начало координат выбираем в центре диска).

Учтем теперь, что радиус диска мал по сравнению с расстояниями r_1, r_2 . Поэтому необходимая для вычисления корреляционной функции квадратичная комбинация полей приводится к виду

$$E_M(\mathbf{r}_1, t) E_{M'}^*(\mathbf{r}_2 t) = A_M A_{M'}^* \frac{\exp\left\{i \frac{\omega}{c} \left(|\mathbf{r}_2 - \rho| - |\mathbf{r}_1 - \rho| \right)\right\}}{r_1 r_2}.$$

Поле, создаваемое в точке \mathbf{r} всем светящимся диском, является суммой полей, создаваемых всеми элементами последнего. Поэтому для получения корреляционной функции поля мы должны просуммировать $E_M E_{M'}^*$ по M, M' и усреднить по случайным параметрам (например, случайным начальным фазам), характеризующим излучение различных атомов. При усреднении воспользуемся тем, что разные атомы излучают некогерентно и в силу этого для $M \neq M'$ $\langle A_M A_{M'}^* \rangle = 0$. Если же $M = M'$, то, очевидно, величина $\langle A_M A_M^* \rangle$ дает интенсивность I_M излучения M -го атома. Так как атомов в источнике очень много, то суммирование по M можно заменить интегрированием, т. е. вместо суммы $\sum_M I_M$ писать $\int I(\rho) d^2\rho$,

где $d^2\rho$ — элемент поверхности источника.

Таким образом, для корреляционной функции поля получим окончательно

$$G(\mathbf{r}_1 t; \mathbf{r}_2 t) = \int_S I(\rho) \frac{\exp \left\{ i \frac{\omega}{c} (|\mathbf{r}_2 - \rho| - |\mathbf{r}_1 - \rho|) \right\}}{r_1 r_2} d^2\rho,$$

где интегрирование производится по поверхности источника. Комплексную степень когерентности можно поэтому представить в виде

$$g(\mathbf{r}_1 t; \mathbf{r}_2 t) = (I_1 I_2)^{-1/2} \int_S I(\rho) \frac{\exp \left\{ i \frac{\omega}{c} (|\mathbf{r}_2 - \rho| - |\mathbf{r}_1 - \rho|) \right\}}{r_1 r_2} d^2\rho,$$

где

$$I_{1,2} = \int_S I(\rho) r_{1,2}^2 d^2\rho.$$

Величины $I_{1,2}$ представляют собой, очевидно, интенсивности в точках P_1 и P_2 .

Формулы для функций G , g имеют ту же структуру, что и формулы для поля, дифрагированного от отверстия (см. гл. 21). Таким образом,

задача о нахождении корреляционной функции в случае протяженного источника сводится к задаче о дифракции от тела, имеющего форму и размеры источника (теорема ван Циттерта — Цернике).

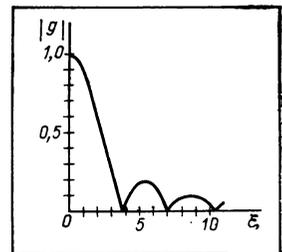
Вычисления приводят к следующему результату:

$$g(\mathbf{r}_1 t; \mathbf{r}_2 t) = \frac{2J_1(\xi)}{\xi} e^{i\psi}, \quad \xi = \frac{2\pi d}{\lambda R}, \quad \psi = \frac{\pi d^2}{\lambda R},$$

где J_1 — функция Бесселя, λ — длина волны и d — проекция отрезка $P_1 P_2$ на плоскость диска. Формула эта справедлива, если $d^2/R \ll \ll c/\Delta\omega$, где $\Delta\omega$ — некогерентность света.

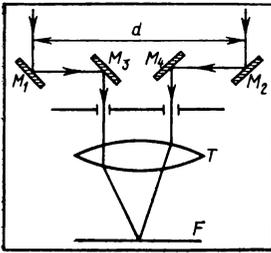
На рис. 15.10 показана зависимость модуля степени когерентности от параметра ξ . При $\xi = 3,83$ степень когерентности обращается в нуль, т. е. поля в точках P_1 и P_2 полностью некогерентны. Это соответствует $d = 3,83\lambda/(2\pi\alpha)$, где $\alpha = \rho/R$ — угол, под которым радиус источника виден из точек P_1 или P_2 .

Эта формула показывает, что, наблюдая интерференционную картину от протяженного источника, можно в принципе определить его угловые размеры, в частности можно определить угловые размеры звезд (а также расстояния между двойными звездами и т. п.). Это было впервые сделано Майкельсоном с помощью так называемого *звездного интерферометра*, схема которого изображена на рис. 15.11. Интерферометр состоит из двух пар зеркал — подвижной (M_1 , M_2) и непод-



15.10.

Зависимость модуля степени когерентности от параметра ξ



15.11
 Схема звездного интерферометра Майкельсона

вижной (M_3, M_4), — отражаясь от которых свет далекой звезды попадает в телескоп T . Интерференционная картина наблюдается в фокальной плоскости F телескопа. Изменяя расстояние d между зеркалами M_1 и M_2 , находят угловой радиус звезды.

На практике этот метод определения угловых размеров звезд имеет ограниченную применимость из-за трудностей работы с большими интерферометрами. Дело в том, что для измерения очень малых углов требуются очень большая разность хода лучей и, следовательно, большие расстояния d между зеркалами.

Между тем практически это расстояние не может превосходить примерно 2 м; при больших размерах интерферометра интерференционная картина оказывается смазанной из-за случайных изменений коэффициента преломления вдоль оптического пути.

15.8. Интерференция интенсивностей

Корреляционная функция $G(x_1, x_2)$, как и определяемая ею степень когерентности $g(x_1, x_2)$, отражает статистические свойства излучения, но для их полного описания ее, вообще говоря, недостаточно. Чтобы понять это, сравним хаотическое поле, которое можно рассматривать как некоторую динамическую и притом очень сложную систему, с другой динамической системой — идеальным газом. Корреляционной функции поля, т. е. среднему значению произведения компонент поля, соответствует, очевидно, среднее значение произведений компонент скорости молекулы или, проще говоря, среднее значение квадрата скорости молекулы. Последняя величина пропорциональна температуре газа T , поэтому корреляционной функции отвечает температура газа. Но температура не определяет полностью статистических свойств газа — для этого нужна функция распределения $f(\mathbf{v})$ молекул по их скоростям \mathbf{v} . Знания среднего значения квадрата скорости $\langle v^2 \rangle = \int v^2 f(\mathbf{v}) d^3v$ достаточно в том случае, если известно, что распределение частиц является максвелловским: $f(\mathbf{v}) \sim \exp[-mv^2/(2T)]$, где m — масса частицы, а температура выражается в энергетических единицах. Тогда $m\langle v^2 \rangle/2 = 3T/2$, так что при данном $\langle v^2 \rangle$ задана и температура T , полностью определяющая распределение Максвелла.

Но в общем случае $\langle v^2 \rangle$ еще не определяет вида функции $f(\mathbf{v})$ — нужно знать всю совокупность величин

$$\langle v^n \rangle = \int v^n f(\mathbf{v}) d^3v \quad (n = 1, 2, 3, \dots),$$

которые называют *моментами функции распределения*. Для максвелловского распределения все они выражаются через T и, следовательно, через $\langle v^2 \rangle$, для других же распределений это независимые величины.

Аналогично максвелловскому распределению ведет себя общее *распределение Гаусса* для случайных вещественных величин x_1, x_2, \dots, x_n . Вероятность того, что эти величины лежат в интервалах $(x_1, x_1 + dx_1), (x_2, x_2 + dx_2), \dots$, определяется формулой

$$p(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n = \\ = \sqrt{\frac{|A|}{(2\pi)^n}} \exp \left\{ -\frac{1}{2} \sum_{i,j} A_{ij} x_i x_j \right\} dx_1 dx_2 \dots dx_n,$$

где A_{ij} — совокупность вещественных величин: матрица, для которой $A_{ij} x_i x_j \geq 0$, и $|A|$ — определитель этой матрицы (предполагается, что средние значения x_i равны нулю). Усредненные произведения величин x определяются формулой

$$\langle x_1 x_2 \dots \rangle = \int x_1 x_2 \dots p(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n.$$

Отсюда следует, в частности, что $\langle x_i x_j \rangle = (A^{-1})_{ij}$, где A^{-1} — матрица, обратная A , и $(A^{-1})_{ij}$ — ее ij -элемент.

Гауссово распределение обладает тем замечательным свойством, что

$$\langle x_1 x_2 \dots x_{2k} \rangle = \sum \langle x_{i_1} x_{i_2} \rangle \langle x_{i_3} x_{i_4} \rangle \dots \langle x_{i_{(2k-1)}} x_{i_{2k}} \rangle,$$

где i_1, i_2, \dots, i_{2k} — любая перестановка чисел $1, 2, \dots, 2k$ и сумма распространяется на все возможные комбинации пар этих чисел. Таким образом, все моменты для распределения Гаусса выражаются через квадратичные моменты $\langle x_i x_j \rangle$ (средние значения произведений любого нечетного числа величин x_i равны нулю). Приведенная формула для $\langle x_1 \dots x_{2k} \rangle$ имеет место только для распределения Гаусса. Для других распределений различные моменты (моменты различных порядков) независимы, и все они необходимы для характеристики распределения.

Вернемся теперь к статистическим свойствам электромагнитного поля. Корреляционную функцию поля $\langle E_\mu(x_1) E_\nu^*(x_2) \rangle$ можно уподобить квадратичному моменту $\langle x_i x_j \rangle$ совокупности случайных величин x_1, \dots, x_n . Но, как только что было разъяснено, для полной характеристики свойств ансамбля случайных величин нужно знать и моменты более высших порядков. Подобно этому для характеристики статистических свойств поля необходимы помимо величин $\langle E_\mu(x_1) E_\nu^*(x_2) \rangle$ еще и величины $\langle E_{\mu_1}(x_1) \dots E_{\mu_n}(x_n) E_{\nu_1}^*(x'_1) \dots E_{\nu_n}^*(x'_n) \rangle$ для $n > 1$. Все они носят общее название корреляционных функций, и для них применяется обозначение $G_{\mu_1 \dots \mu_n \nu_1 \dots \nu_n}^{(n)}(x_1, \dots, x_n, x'_1 \dots x'_n)$. Функции $G^{(n)}$ называют *корреляционными функциями n -го порядка*. Корреляционную функцию, которую мы изучали до сих пор, следует теперь обозначать $G_{\mu, \nu}^{(1)}(x_1, x'_1)$ и называть корреляционной функцией первого порядка. Если поляризация поля задана точно, то в поляризационных индексах $\mu_1, \dots, \nu_1, \dots$ нет нужды и корреляционную функцию n -го порядка можно определить просто как

$$G^{(n)}(x_1 \dots x_n; x'_1 \dots x'_n) = \langle E(x_1) \dots E(x_n) E^*(x'_1) \dots E^*(x'_n) \rangle.$$

Корреляционные функции высших порядков позволяют обобщить понятие когерентности.

Дело в том, что это понятие связано с регулярностью излучения. Если поле регулярно, то имеет место факторизуемость корреляционной функции первого порядка $G^1(x_1, x_2) = E(x_1)E^*(x_2)$. Аналогичная ситуация характерна и для корреляционных функций более высоких порядков:

$$G^{(n)}(x_1 \dots x_n; x'_1 \dots x'_n) = E(x_1) \dots E(x_n) E^*(x'_1) \dots E^*(x'_n).$$

Если такая факторизация (с одной и той же функцией $E(x)$) имеет место для всех сколь угодно больших n , то можно считать поле идеально регулярным, или полностью детерминированным и соответственно полнотой когерентным, но может быть и так, что факторизуемость выполняется только при $n \leq M$; если же $n > M$, то факторизуемости нет. В этом случае говорят, что излучение обладает когерентностью M -го порядка.

Таким образом, с этой точки зрения полная когерентность, о которой шла речь в предыдущем параграфе, следует называть когерентностью первого порядка. Для описания элементарных интерференционных явлений в оптике этого понятия достаточно, но для описания более тонких интерференционных явлений, о которых пойдет речь далее, необходимо использовать понятие когерентности более высоких порядков.

Факторизуемость корреляционной функции эквивалентна тому, что величина

$$g^{(n)}(x_1 \dots x_n; x'_1 \dots x'_n) = \frac{G^{(n)}(x_1 \dots x_n; x'_1 \dots x'_n)}{\sqrt{G^{(1)}(x_1, x_1) G^{(1)}(x_2, x_2) \dots G^{(1)}(x'_n, x'_n)}}$$

равна по модулю единице. Эту величину называют *степенью когерентности n -го порядка*. Она является обобщением величины $g(x_1, x_2)$, которую мы рассматривали в предыдущем параграфе и которую должны теперь называть степенью когерентности первого порядка. Таким образом, если поле обладает когерентностью n -го порядка, то

$$|G^{(n)}(x_1 \dots x_n; x'_1 \dots x'_n)|^2 = G^{(1)}(x_1, x_1) \dots G^{(1)}(x'_n, x'_n).$$

Рассмотрим теперь достаточно хаотическое поле, скажем, поле, которое можно трактовать как гауссово. Тогда в соответствии со сказанным выше о моментах гауссова распределения все корреляционные функции выражаются через корреляционную функцию первого порядка. Например, корреляционная функция второго порядка имеет вид

$$G^{(2)}(x_1, x_2; x'_1, x'_2) = G^{(1)}(x_1, x'_1) G^{(1)}(x_2, x'_2) + G^{(1)}(x_1, x'_2) G^{(1)}(x'_1, x_2).$$

Это значит, что если даже гауссово поле обладает когерентностью первого порядка, то когерентностью второго порядка оно обладать уже

не может. В самом деле, в этом случае корреляционная функция второго порядка должна была бы иметь вид $G^{(2)}(x_1 x_2; x'_1 x'_2) = E(x_1)E(x_2)E^*(x'_1)E^*(x'_2)$, т. е. в ней не возникало бы множителя 2. Тем более гауссово поле не может обладать когерентностью третьего и более высоких порядков.

Нам остается разъяснить, как можно наблюдать наличие или отсутствие у поля когерентности второго и следующих порядков. Для этого нужно иметь прибор, реагирующий на корреляционную функцию соответствующего порядка. Выше мы уже говорили, что счетчик фотонов, основанный на фотоэффекте, реагирует на корреляционную функцию $G^{(1)}(x_1, x_2)$. Более точно это означает следующее. Если представить себе идеализированный прибор, имеющий в качестве чувствительного элемента один атом, то при падении света на этот атом из него может вырваться электрон, вероятность чего пропорциональна

$$W^{(1)}(t_0, t) = \int_{t_0}^t G^{(1)}(\mathbf{r} t', \mathbf{r} t') dt',$$

где t_0 и t — моменты начала и конца освещения атома светом и \mathbf{r} — место расположения атома.

Рассмотрим теперь более сложное, но тоже идеализированное устройство, содержащее два чувствительных элемента в виде двух атомов, расположенных в точках \mathbf{r}_1 и \mathbf{r}_2 . Ставится вопрос: какова будет вероятность того, что в течение интервала времени (t_0, t) каждый из атомов поглотит фотон? Ответ гласит, что вероятность пропорциональна

$$W^{(2)}(t_0, t) = \int_{t_0}^t \int_{t_0}^t G^{(2)}(\mathbf{r}_1 t'_1, \mathbf{r}_2 t'_2; \mathbf{r}_2 t'_2, \mathbf{r}_1 t'_1) dt'_1 dt'_2.$$

Если бы мы интересовались вероятностью того, что первый атом поглотит фотон в интервале времени (t_0, t_1) , а второй атом — в интервале времени (t_0, t_2) , то мы получили бы формулу

$$W^{(2)}(t_0, t_1; t_0, t_2) = \int_{t_0}^{t_1} dt'_1 \int_{t_0}^{t_2} dt'_2 G^{(2)}(\mathbf{r}_1 t'_1, \mathbf{r}_2 t'_2; \mathbf{r}_2 t'_2, \mathbf{r}_1 t'_1).$$

Таким образом, корреляционная функция $G^{(2)}$ определяет вероятность процесса поглощения двух фотонов двумя атомами в один и тот же момент времени либо в течение заданных интервалов времени. Следовательно, для того чтобы измерять корреляционную функцию $G^{(2)}$, нужно построить прибор, состоящий из двух детекторов фотонов и регистрирующий только такие события, когда оба фотона поглощаются в заданные моменты времени (точнее, интервалы времени). Такие приборы можно назвать *счетчиками фотонных (или задержанных) совпадений*. Аналогичным образом, если построить прибор, состоящий из n детекторов фотонов и регистрирующий совпадение процессов поглощения n фотонов, то скорость счета прибора будет измерять корреляционную функцию $G^{(n)}$.

Рассчитаем, например, скорость счета задержанных совпадений для двухфотонного счетчика, предполагая, что излучение представляет собой гауссов процесс. Если свет распространяется вдоль оси x , то скорость счета пропорциональна величине $G^{(2)}(x_1 t_1, x_2 t_2; x_2 t_2, x_1 t_1)$, которая в силу гауссовости процесса составляет

$$G^{(2)} \equiv G^{(2)}(x_1 t_1, x_2 t_2; x_2 t_2, x_1 t_1) = \\ = G^{(1)}(x_1 t_1, x_1 t_1) G^{(1)}(x_2 t_2, x_2 t_2) + |G^{(1)}(x_1 t_1, x_2 t_2)|^2.$$

Но согласно определению степени когерентности первого порядка

$$|G^{(1)}(x_1 t_1, x_2 t_2)|^2 = \\ = G^{(1)}(x_1 t_1, x_1 t_1) G^{(1)}(x_2 t_2, x_2 t_2) |g^{(1)}(x_1 t_1, x_2 t_2)|^2.$$

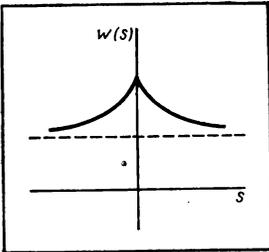
Поэтому $G^{(2)} = G^{(1)}(1)G^{(1)}(2) [1 + |g^{(1)}(x_1 t_1, x_2 t_2)|^2]$, где $1 \equiv x_1 t_1$, $2 \equiv x_2 t_2$. Обратим внимание на то, что рассматриваемое излучение не обладает когерентностью второго порядка, так как при когерентности второго порядка в выражении для $G^{(2)}$ не было бы множителя в квадратных скобках.

Мы видим, что при измерении величины $G^{(2)}$, по существу, находится модуль степени когерентности первого порядка. В частности, как мы знаем, при лоренцевой форме линии излучения $|g| = e^{-\tau|s|}$, где $s = t_1 - t_2 - (x_1 - x_2)/c$, и

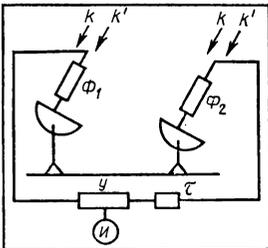
$$G^{(2)} = \frac{U^2}{4} (1 + e^{-2\tau|s|}).$$

Поэтому вероятность или частота $W(s)$ регистрации двух фотонов в точках x_1 и x_2 в моменты времени t_1 и t_2 пропорциональна $1 + e^{-2\tau|s|}$. Зависимость функции $W(s)$ от s изображена на рис. 15.12.

Таким образом, для излучения, имеющего статистические свойства гауссова процесса, существует корреляция между временами регистрации фотонов. Это явление называют эффектом Хэнбери—Брауна и Твисса. Наблюдая его, можно найти степень когерентности первого порядка, зная которую можно, в свою очередь, как мы видели в предыдущем параграфе, найти угловые размеры источника излучения. Такая программа была выполнена Хэнбери—Брауном и Твиссом, построившими с этой целью специальный прибор, который может быть назван *интерферометром интенсивностей* (рис. 15.13). В нем свет (точнее, радиоизлучение) от далекой звезды (радиозвезды) фокусируется на два фотодетектора Φ_1 и Φ_2 , сигналы



15.12.
К объяснению эффекта Хэнбери—Брауна и Твисса



15.13.
Интерферометр интенсивностей

от которых поступают, пройдя линию задержки τ и усилитель \mathcal{U} , в интегратор I . Здесь фототоки перемножаются и их произведение, усредненное по времени, регистрируется в зависимости от времени задержки (т. е. величины s/c). Эта величина пропорциональна $|g|$, и поэтому, определив $|g|$ как функцию углового размера звезды (конкретнее, радиозвезды), можно найти этот угловой размер. Так как в этом приборе усредняется произведение токов, т. е. интенсивностей, а не полей, как, например, в звездном интерферометре Майкельсона, то установка Хэнбери—Брауна и Твисса может быть названа, как уже упоминалось, интерферометром интенсивностей.

Интерферометр интенсивностей имеет большие преимущества по сравнению со звездным интерферометром Майкельсона — они связаны с тем, что в установке Хэнбери—Брауна и Твисса, по сути дела, нет высокочастотных колебаний, а есть только сравнительно низкочастотные фототоки. Поэтому их удается передать без искажений на расстояние, гораздо большее предельных размеров интерферометра Майкельсона. Это, в свою очередь, позволяет существенно увеличить разность хода лучей s без того, чтобы смазалась интерференционная картина.

ОСНОВНЫЕ ФОРМУЛЫ

Связь степени немонахроматичности излучения с временем жизни состояния

$$\Delta\omega \sim \frac{1}{\tau}$$

Длина когерентности

$$l_c = \frac{c}{\Delta\omega}$$

Разность фаз между лучами

$$\varphi_I - \varphi_{II} = \frac{\omega}{c} \left(\sum_{(I)} n r N - \sum_{(II)} n r N \right)$$

Корреляционные функции

$$G_{\mu\nu}(r_1 t_1, r_2 t_2) = \langle E_\mu(r_1 t_1) E_\nu^*(r_2 t_2) \rangle$$

Корреляционная функция n -го порядка

$$\begin{aligned} G^{(n)}(x_1 x_2 \dots x_n; x'_1 x'_2 \dots x'_n) = \\ = \langle E(x_1) \dots E(x_n) E^*(x'_1) \dots E^*(x'_n) \rangle \end{aligned}$$

Поле полностью когерентно при условии факторизации

$$\begin{aligned} G^{(n)}(x_1 x_2 \dots x_n; x'_1 x'_2 \dots x'_n) = \\ = E(x_1) \dots E(x_n) E^*(x'_1) \dots E^*(x'_n) \end{aligned}$$

Глава 16. ЭЛЕКТРОМАГНИТНЫЕ ВОЛНЫ В ВЕЩЕСТВЕ

16.1. Плоские волны в однородном изотропном диэлектрике

В гл. 6 и 14 мы показали, что в вакууме возможно распространение поперечных электромагнитных волн и что их скорость не зависит от частоты и амплитуды волны. Переходя теперь к изучению электромагнитных волн в материальных средах, начнем с волн в однородном изотропном диэлектрике.

Будем исходить из уравнений Максвелла для среды в отсутствие внешних зарядов и токов:

$$\operatorname{rot} \mathbf{E} = - \frac{\partial \mathbf{B}}{\partial t}, \quad \operatorname{div} \mathbf{D} = 0,$$

$$\operatorname{rot} \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t}, \quad \operatorname{div} \mathbf{B} = 0.$$

Как мы знаем, индукции \mathbf{D} и \mathbf{B} связаны с напряженностями полей \mathbf{E} и \mathbf{H} в случае статических полей и изотропных сред соотношениями $\mathbf{D} = \epsilon_0 \epsilon \mathbf{E}$, $\mathbf{B} = \mu_0 \mu \mathbf{H}$, где ϵ и μ — электрическая и магнитная проницаемости среды. Эти величины представляют собой материальные константы, которые могут зависеть от температуры, а также (как это имеет место, например, в ферромагнетиках и сегнетоэлектриках) и от амплитуд полей. Мы будем рассматривать здесь простейший случай и не учитывать амплитудных зависимостей ϵ и μ ; такие зависимости приводят к нелинейности уравнений поля, и этому вопросу посвящена гл. 22. Но главное, мы будем считать, что и в случае переменных полей связи между \mathbf{D} , \mathbf{E} и \mathbf{B} , \mathbf{H} такие же, как и в случае постоянных полей, т. е. что свойства переменных полей в среде характеризуются теми же материальными константами, что и свойства постоянных полей. Как мы покажем далее, это предположение, вообще говоря, не соответствует действительности, и если им и можно пользоваться, то только в ограниченной области частот, которую каждый раз нужно специально находить. Тем не менее имеет смысл изучить сперва волны в гипотетической среде с материальными константами ϵ и μ , не зависящими ни от каких параметров волны, ибо при этом будут выяснены многие закономерности, относящиеся и к реальным средам.

Итак, сперва будем считать, что и в случае переменных полей справедливы соотношения $\mathbf{D} = \epsilon_0 \epsilon \mathbf{E}$, $\mathbf{B} = \mu_0 \mu \mathbf{H}$, где ϵ и μ не зависят ни от частоты, ни от амплитуды волны. Кроме того, среду будем считать однородной, а в этом случае величины ϵ и μ не зависят и от пространственных координат. Наконец, предполагая среду изотропной, мы будем считать, что \mathbf{D} параллельно \mathbf{E} , а \mathbf{B} параллельно \mathbf{H} .

Используя эти соотношения, можно исключить из уравнений Максвелла векторы индукций и пользоваться только напряженностями полей:

$$\operatorname{rot} \mathbf{E} = - \frac{\mu_0 \mu}{\partial t} \frac{\partial \mathbf{H}}{\partial t}, \operatorname{div} \mathbf{E} = 0,$$

$$\operatorname{rot} \mathbf{H} = \frac{\epsilon_0 \epsilon}{\partial t} \frac{\partial \mathbf{E}}{\partial t}, \operatorname{div} \mathbf{H} = 0.$$

Применим к первому из этих уравнений операцию rot . Используя формулу

$$\operatorname{rot} \operatorname{rot} \mathbf{E} = \operatorname{grad} \operatorname{div} \mathbf{E} - \Delta \mathbf{E}$$

и учитывая, что $\operatorname{div} \mathbf{E} = 0$, получим

$$- \Delta \mathbf{E} = - \mu_0 \mu \frac{\partial}{\partial t} \operatorname{rot} \mathbf{H}.$$

Выразив далее с помощью второго уравнения $\operatorname{rot} \mathbf{H}$ через $\partial \mathbf{E} / \partial t$, мы придем к уравнению, содержащему только \mathbf{E} :

$$\Delta \mathbf{E} - \frac{\epsilon \mu}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0$$

(мы учли, что $\epsilon_0 \mu_0 = c^{-2}$). Если бы мы применили сперва ко второму уравнению операцию rot и затем использовали первое уравнение, то пришли бы к такому же уравнению для напряженности магнитного поля:

$$\Delta \mathbf{H} - \frac{\epsilon \mu}{c^2} \frac{\partial^2 \mathbf{H}}{\partial t^2} = 0.$$

Мы получили для полей волновые уравнения, отличающиеся от волновых уравнений для полей в вакууме только тем, что в них вместо c^2 входит величина $c^2 / (\epsilon \mu)$. Отсюда мы приходим к выводу, что

электромагнитные волны возможны и в средах, причем скорость их равна не c , а $v_\phi = c / \sqrt{\epsilon \mu}$, и волны, как и в вакууме, являются поперечными.

Рассмотрим подробнее простейший класс электромагнитных волн — *плоские монохроматические волны*, для которых пространственная и временная зависимости полей имеют вид

$$\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k}\mathbf{r} - \omega t)}; \quad \mathbf{H} = \mathbf{H}_0 e^{i(\mathbf{k}\mathbf{r} - \omega t)},$$

где ω , \mathbf{k} — частота и волновой вектор, а \mathbf{E}_0 , \mathbf{H}_0 — комплексные амплитуды. Для плоской волны, как мы знаем,

$$\operatorname{rot} \mathbf{A} = i[\mathbf{k}\mathbf{A}], \operatorname{div} \mathbf{A} = i\mathbf{k}\mathbf{A}, \Delta \mathbf{A} = -k^2 \mathbf{A};$$

далее, для монохроматической волны $\partial \mathbf{A} / \partial t = -i\omega \mathbf{A}$. Поэтому волновые уравнения для полей позволяют связать волновое число и частоту волны:

$$k = \omega / v_\phi, \quad v_\phi = c / \sqrt{\epsilon \mu}.$$

Записав \mathbf{k} в виде $\mathbf{k} = nk = \mathbf{n}\omega / v_\phi$, где \mathbf{n} — единичный вектор в направлении распространения волны, и подставив это выражение в

формулы, связывающие \mathbf{E}_0 , \mathbf{H}_0 и \mathbf{k} , получим $(\omega/v_\Phi)[\mathbf{nE}_0] = \omega\mu_0\mu\mathbf{H}_0$, а так как $v_\Phi = c/\sqrt{\epsilon\mu}$, то

$$[\mathbf{nE}_0] = c\mu_0 \sqrt{\frac{\mu}{\epsilon}} \mathbf{H}_0.$$

Аналогично,

$$[\mathbf{nH}_0] = -c\epsilon_0 \sqrt{\frac{\epsilon}{\mu}} \mathbf{E}_0.$$

Мы видим, что оба поля взаимно перпендикулярны и перпендикулярны направлению распространения волны, т. е. волна является поперечной. Если \mathbf{A} перпендикулярно \mathbf{n} , то $|\mathbf{nA}| = A$, поэтому

$$\sqrt{\epsilon_0\epsilon} E_0 = \sqrt{\mu_0\mu} H_0.$$

Соотношение, связывающее амплитуды полей, имеет простой физический смысл. При постоянных и не зависящих от параметров полей проницаемостях ϵ и μ плотности электрической и магнитной энергии равны $\epsilon_0\epsilon E^2/2$ и $\mu_0\mu H^2/2$. Поэтому это соотношение означает, что в плоской монохроматической волне обе плотности энергии одинаковы: $\epsilon_0\epsilon E^2/2 = \mu_0\mu H^2/2$.

Как мы знаем, вектор плотности потока электромагнитной энергии определяется формулой $\mathbf{S} = [\mathbf{EH}]$. Для плоской волны это выражение может быть представлено в виде

$$\mathbf{S} = EH\mathbf{n},$$

а так как $H = c\epsilon_0\sqrt{\epsilon/\mu} E$, то

$$\mathbf{S} = c\epsilon_0 \sqrt{\frac{\epsilon}{\mu}} E^2 \mathbf{n} = c\mu_0 \sqrt{\frac{\mu}{\epsilon}} H^2 \mathbf{n}.$$

Учитывая, наконец, что плотность электромагнитной энергии

$$w = \frac{1}{2} (\epsilon_0\epsilon E^2 + \mu_0\mu H^2),$$

можно переписать \mathbf{S} в виде

$$\mathbf{S} = v_\Phi w \mathbf{n},$$

где $v_\Phi = c/\sqrt{\epsilon\mu}$ — скорость электромагнитной волны. Эта формула имеет наглядный физический смысл:

плотность потока энергии равна произведению плотности энергии на скорость распространения волны.

16.2. Дисперсия проницаемостей

В предыдущем параграфе мы рассматривали электромагнитные волны в среде с не зависящими от параметров волны

проницаемостями ϵ и μ . В действительности, однако, как уже говорилось, таких сред не существует и проницаемости оказываются зависящими прежде всего от частоты волны, а также во многих случаях и от волнового вектора. Эти зависимости объединяют общим названием *дисперсии проницаемостей*. При этом если проницаемость (безразлично какая — диэлектрическая или магнитная) зависит от частоты, то говорят о *временной дисперсии*; если проницаемость зависит от волнового вектора, то говорят о *пространственной дисперсии*.

Математически это означает следующее. Векторы напряженностей \mathbf{E} и \mathbf{H} и индукций \mathbf{D} и \mathbf{B} , зависящие от пространственных координат \mathbf{r} и времени t , могут быть представлены в виде суперпозиции плоских монохроматических волн типа $\exp(i\mathbf{k}\mathbf{r} - i\omega t)$, т. е. в виде интегралов фурье-типа:

$$\mathbf{f}(\mathbf{r}, t) = \int \mathbf{f}_{\mathbf{k}\omega} \exp(i\mathbf{k}\mathbf{r} - i\omega t) d^3k d\omega,$$

где $f_{\mathbf{k}\omega}$ — амплитуды соответствующих волн, зависящие от частоты ω и волнового вектора \mathbf{k} (эти амплитуды носят название *фурье-компонент* функции $\mathbf{f}(\mathbf{r}, t)$). Для изотропных сред фурье-компоненты индукций и полей связаны между собой соотношениями

$$\mathbf{D}_{\mathbf{k}\omega} = \epsilon_0 \epsilon(\omega, \mathbf{k}) \mathbf{E}_{\mathbf{k}\omega}, \quad \mathbf{B}_{\mathbf{k}\omega} = \mu_0 \mu(\omega, \mathbf{k}) \mathbf{H}_{\mathbf{k}\omega},$$

где величины $\epsilon(\omega, \mathbf{k})$ и $\mu(\omega, \mathbf{k})$ зависят от частоты и волнового вектора волны и параметров среды. Они по-прежнему носят названия диэлектрической и магнитной проницаемостей.

Обратим внимание на то, что

при наличии дисперсии волны с разными ω и \mathbf{k} распространяются с разной скоростью:

$$v_{\phi} \equiv v_{\phi}(\omega, \mathbf{k}) = c [\epsilon(\omega, \mathbf{k}) \mu(\omega, \mathbf{k})]^{-1/2}.$$

Если дисперсия отсутствует, т. е. если ϵ и μ являются постоянными, характеризующими среду, то индукции в каждый момент времени t и в каждой точке пространства \mathbf{r} определяются значениями полей в тот же момент времени и в той же точке: $\mathbf{D}(\mathbf{r}, t) = \epsilon_0 \epsilon \mathbf{E}(\mathbf{r}, t)$, $\mathbf{B}(\mathbf{r}, t) = \mu_0 \mu \mathbf{H}(\mathbf{r}, t)$. Чтобы понять, что с этой точки зрения означает дисперсия проницаемостей, сделаем обратное преобразование фурье-соотношений, выражающих индукции через напряженности. Если ввести функции

$$\epsilon(\mathbf{r}, t) = \int \epsilon(\omega, \mathbf{k}) \exp(i\mathbf{k}\mathbf{r} - i\omega t) d^3k d\omega / (2\pi)^4;$$

$$\mu(\mathbf{r}, t) = \int \mu(\omega, \mathbf{k}) \exp(i\mathbf{k}\mathbf{r} - i\omega t) d^3k d\omega / (2\pi)^4,$$

то мы получим

$$\mathbf{D}(\mathbf{r}, t) = \epsilon_0 \int \epsilon(\mathbf{r} - \mathbf{r}', t - t') \mathbf{E}(\mathbf{r}', t') dV' dt',$$

$$\mathbf{B}(\mathbf{r}, t) = \mu_0 \int \mu(\mathbf{r} - \mathbf{r}', t - t') \mathbf{H}(\mathbf{r}', t') dV' dt'.$$

Таким образом,

наличие временной дисперсии означает, что индукции в момент времени t определяются напряженностями не только в этот, но и в другие моменты времени t' . Аналогично, наличие пространственной дисперсии означает, что индукции в пространственной точке \mathbf{r} определяются напряженностями не только в этой, но и в других точках пространства \mathbf{r}' (в таком случае говорят о нелокальной связи индукций с напряженностями).

Существенно, что индукция в момент времени t может определяться значениями напряженности в предыдущие (но не в последующие!) моменты времени. В самом деле, откуда физическому объекту знать, что его ждет в будущем? Поэтому функции $\varepsilon(\mathbf{r}-\mathbf{r}', t-t')$ и $\mu(\mathbf{r}-\mathbf{r}', t-t')$ при $t > t'$ обязательно равны нулю. Это их свойство является следствием общего физического принципа, называемого п р и н ц и п о м п р и ч и н н о с т и.

Функции $\varepsilon(\omega, \mathbf{k})$ и $\mu(\omega, \mathbf{k})$ различны для разных веществ; для того чтобы их найти, нужно знать структуру вещества и поведение его в поле волны. В простейшем случае идеального газа для определения проницаемости нужно знать, как поведут себя отдельные атомы и молекулы во внешнем переменном электромагнитном поле. Это поведение определяется как структурой, так и динамикой атомов и молекул, которые могут быть последовательно описаны и исследованы только в рамках квантовой механики. Таким образом, нахождение проницаемостей $\varepsilon(\omega, \mathbf{k})$ и $\mu(\omega, \mathbf{k})$ представляет собой, вообще говоря, квантово-механическую задачу.

Исключением является задача об определении высокочастотной диэлектрической проницаемости плазмы, т. е. нейтрального в целом газа заряженных частиц, а также задача об определении высокочастотной магнитной проницаемости ферромагнетиков (и других магнитоупорядоченных кристаллов). Эти задачи могут быть решены без использования квантовой механики на основе только классической механики, и мы рассмотрим их в гл. 17 и 19, здесь же остановимся на классической теории дисперсии диэлектрической проницаемости одноатомного газа, основанной на классической модели атома.

Согласно этой модели, электроны в атоме совершают гармонические колебания относительно ядра, т. е. атом можно рассматривать как совокупность гармонических осцилляторов. Рассмотрим один из таких осцилляторов. Если \mathbf{r} — смещение электрона относительно ядра, то изменение со временем этой величины определяется законом движения осциллятора:

$$m_e \ddot{\mathbf{r}} + [m_e \gamma \dot{\mathbf{r}} + m_e \omega_0^2 \mathbf{r}] = \mathbf{f},$$

где \mathbf{f} — внешняя сила, действующая на электрон, m_e — масса электрона, ω_0 — собственная частота его колебаний и γ — коэффициент затухания, который, как будет видно далее, должен быть обязательно введен. (Введение коэффициента затухания соответствует тому, что написанное уравнение описывает возбужденное состояние атома, время жизни которого обратно пропорционально γ .) Наша задача заключается теперь в том, чтобы выяснить, как ведет себя осциллятор в поле плоской монохроматической электромагнитной волны $\mathbf{f} = e\mathbf{E}_0 \exp(i\mathbf{k}\mathbf{r} - i\omega t)$ (скорость атомного электрона мала по сравнению со скоростью света, поэтому мы учли здесь только силу, действующую на него со стороны электрического поля волны). Предполагая, что размеры атома a малы по сравнению с длиной волны λ , т. е. $\lambda \gg a$, можно пренебречь в экспоненте

величиной kr , так как $kr \sim 2\pi(a/\lambda) \ll 1$. Поэтому можно считать электрическое поле зависящим только от времени, так что $\mathbf{i} = e\mathbf{E}_0 \exp(-i\omega t)$ и уравнение движения электрона, связанного с ядром, принимает вид

$$\ddot{\mathbf{r}} + \gamma \dot{\mathbf{r}} + \omega_0^2 \mathbf{r} = \frac{e}{m_e} \mathbf{E}_0 e^{-i\omega t}.$$

Нас должно интересовать только вынужденное решение этого уравнения

$$\mathbf{r} = \frac{\frac{e}{m_e} \mathbf{E}_0 e^{-i\omega t}}{-\omega^2 - i\gamma\omega + \omega_0^2},$$

обращающееся в нуль при $\mathbf{E}_0 = 0$ (так как мы пользуемся комплексным методом, то должна быть взята вещественная часть этого выражения). Имея это решение, можно найти вектор поляризации \mathbf{P} газа в переменном поле волны. Этот вектор определяется формулой $\mathbf{P} = n\mathbf{d}$, где \mathbf{d} — индуцированный полем дипольный момент атома и n — плотность атомов (электронов). Ясно, что найденная величина радиуса смещения \mathbf{r} и определяет $\mathbf{d} = e\mathbf{r}$. Вспоминая, что $\mathbf{P} = \alpha\mathbf{E}$, где α — поляризуемость, найдем поляризацию газа в переменном поле:

$$\alpha = \frac{e^2 n}{\epsilon_0 m_e} (-\omega^2 - i\gamma\omega + \omega_0^2)^{-1}.$$

Мы можем теперь найти диэлектрическую проницаемость газа в переменном поле: она всегда связана с α соотношением $\epsilon = 1 + \alpha$, и поэтому

$$\epsilon = 1 + \frac{e^2 n}{\epsilon_0 m_e} (-\omega^2 - i\gamma\omega + \omega_0^2)^{-1}.$$

Если в состав атомов входят электроны разных оболочек, т. е. электроны, колеблющиеся с разными частотами ω_i , то эта формула должна быть заменена формулой

$$\epsilon = 1 + \sum_i \frac{e^2 n_i}{\epsilon_0 m_e} (-\omega^2 - i\gamma_i\omega + \omega_i^2)^{-1},$$

где n_i — плотность электронов i -оболочки и γ_i — соответствующий коэффициент затухания (в ГС здесь вместо $1/\epsilon_0$ стоит множитель 4π).

Мы видим, что классическая модель атома приводит к зависимости диэлектрической проницаемости газа от частоты волны, иными словами, к временной дисперсии диэлектрической проницаемости. Только в области низких частот, когда $\omega \ll \omega_i$, диэлектрическая проницаемость может считаться не зависящей от частоты, и мы приходим к случаю, рассмотренному в предыдущем параграфе. Что же касается пространственной дисперсии, то мы не получили ее, так как предполагали, что длина волны значительно больше размеров атома. Дело в том, что волновой вектор может входить в выражение для диэлектрической проницаемости только в виде безразмерной комбинации ka , где a — величина, характеризующая атом или среду в целом и имеющая размерность длины. В случае одноатомного газа, рассматриваемого нами, единственной величиной размерности длины, которая может входить в выражение для ϵ , является линейный размер атома a . Рассматривая длинные волны, для которых $ka \ll 1$, мы, естественно, не получаем пространственной дисперсии.

Возвратимся к простейшему выражению для ϵ в случае одной атомной частоты ω_0 и исследуем частотную зависимость диэлектрической проницаемости. Прежде всего обратим внимание на то, что если коэффициент затухания γ равен нулю, то при $\omega \approx \omega_0$ диэлектрическая проницаемость обращается в бесконечность. Именно по этой причине мы ввели в рассмотрение затухание осциллятора. При этом ϵ становится комплексной величиной и, следовательно, комплексной будет и фазовая скорость волны v_Φ .

Мы изложили теорию дисперсии одноатомного газа, рассматривая атом как классический осциллятор. В действительности, однако, атом подчиняется не классической, а квантовой механике и представление об атомном электроны как о классическом осцилляторе является неправильным. Согласно квантовой механике, если атом находится в некотором i -м состоянии с энергией ϵ_i , то поляризуемость его определяется формулой

$$\alpha'_i = \frac{e^2}{\epsilon_0 m_e} \sum_s f_{is} (\omega_{is}^2 - \omega^2)^{-1},$$

где $\omega_{is} = (\epsilon_i - \epsilon_s)/\hbar$ — частота перехода атома из состояния i в состояние s , f_{is} — некоторые константы, характеризующие атом и не зависящие от частоты внешнего поля ω . Их называют *силами осциллятора*. Суммирование производится по всем состояниям s , отличным от i . (Приведенная формула для поляризуемости атома не учитывает затухания, связанного с конечностью времени жизни возбужденного состояния атома.) Зная α'_i , легко найти диэлектрическую проницаемость газа как функцию частоты:

$$\epsilon(\omega) = 1 + n\alpha'_i = 1 + \frac{e^2 n}{\epsilon_0 m_e} \sum_s f_{is} (\omega_{is}^2 - \omega^2)^{-1}.$$

Сила осциллятора f_{is} определяется матричным элементом \mathbf{d}_{is} дипольного момента \mathbf{d} атома и частотой перехода ω_{is} :

$$f_{is} = 2m_e \omega_{is} |d_{is}|^2 / (e^2 \hbar).$$

16.3. Комплексная фазовая скорость

Разъясним теперь комплексность скорости волны v_Φ . Вводя обозначение $N = \sqrt{\epsilon(\omega, \mathbf{k})\mu(\omega, \mathbf{k})}$, имеем $v_\Phi = c/N$. Комплексную величину N (ее называют *комплексным показателем преломления*) можно записать в виде $N = N' + iN''$, где N' и N'' — вещественные величины. Рассмотрим плоскую монохроматическую волну $\mathbf{E} = \mathbf{E}_0 \exp[i\omega(z/v_\Phi - t)]$, распространяющуюся в положительном направлении оси z , и подставим вместо скорости v_Φ ее выражение через показатель преломления:

$$\mathbf{E} = \mathbf{E}_0 e^{i\omega(z/v'_\Phi - t)} e^{-\alpha z},$$

где $v'_\Phi = c/N'$, $\alpha = (\omega/c)N''$.

Смысл фазы волны φ и амплитуды волны имеют, очевидно, величины

$$\varphi(z, t) \equiv \omega \left(\frac{z}{v'_\Phi} - t \right), \quad \mathbf{E}_0(z) \equiv \mathbf{E}_0 e^{-\alpha z}.$$

Таким образом, амплитуда плоской монохроматической волны перестает быть одинаковой во всех точках пространства и начинает убывать вдоль направления распространения волны, т. е. вдоль оси z .

Что касается фазы волны, то она остается постоянной для наблюдателя, движущегося вдоль оси z со скоростью $dz/dt = v'_\phi$; поэтому вещественная величина v'_ϕ имеет смысл фазовой скорости волны, величину же v_ϕ называют *комплексной фазовой скоростью*.

Величина κ , которую можно назвать *коэффициентом затухания*, определяет глубину проникновения волны в рассматриваемой среде. В самом деле, на глубине $\delta \sim 1/\kappa$ амплитуда волны уменьшится в e раз и, следовательно, плотность энергии — в $e^2 \sim 10$ раз. Очевидно, в направлении распространения волны поле может только убывать, иначе для тела достаточно больших размеров оно могло бы стать сколь угодно большим (речь идет о средах, в которых не действуют какие-либо источники энергии). Поэтому обязательно $\kappa > 0$ и, следовательно, $N'' > 0$, т. е.

мнимая часть комплексного показателя преломления обязательно положительна.

Мы рассмотрели, как убывает в пространстве амплитуда монохроматической волны. Поставим теперь другую задачу: пусть в некоторый момент времени $t = 0$ возбуждена волна, амплитуда которой всюду одинакова; как будет вести себя эта амплитуда с течением времени? Записав поле волны в виде $E = E_0 \exp[ik(z - v_\phi t)]$ и выражая комплексную фазовую скорость через показатель преломления, имеем (при $N'' \ll N'$)

$$E = E_0(t) \exp[ik(z - v'_\phi t)], \quad E_0(t) = E_0 \exp(-\kappa v'_\phi t).$$

Очевидно, смысл амплитуды волны имеет теперь величина $E_0(t)$, но она убывает со временем по экспоненциальному закону. Величину $\kappa v'_\phi$ называют *декрементом затухания*; она определяет время $\tau \sim (v'_\phi \kappa)^{-1}$, за которое энергия волны убывает в $e^2 \sim 10$ раз. Неравенство $N'' > 0$ обеспечивает отсутствие нарастания амплитуды волны в случае равновесной (точнее, замкнутой) физической системы и тем самым выполнение для такой системы закона сохранения энергии.

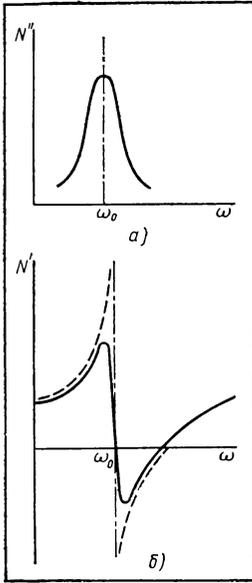
Заметим, однако, что возможны неравновесные физические системы, для которых мнимая часть комплексного показателя преломления не обязательно положительна, а может быть и отрицательной (в каком-то диапазоне частот); такие системы используются на практике как генераторы колебаний. К ним относятся, например, лазеры и система плазма — пучок (см. гл. 17 и 23).

Проиллюстрируем дисперсию, т. е. частотную зависимость, комплексного коэффициента преломления на рассмотренном выше примере классического одноатомного газа. Если $\mu = 1$ и рассматривается только одна собственная частота колебаний электронов, то

$$N = \sqrt{\epsilon} = \sqrt{1 + \frac{e^2 n}{\epsilon_0 m_e (-\omega^2 - i\gamma\omega + \omega_0^2)}}.$$

Считая газ достаточно разреженным, получим отсюда

$$N = 1 + \frac{e^2 n}{2\epsilon_0 m_e (-\omega^2 - i\gamma\omega + \omega_0^2)}$$



16.1. Дисперсия мнимой (а) и вещественной (б) составляющих комплексного коэффициента преломления $N = N' + iN''$ для классического одноатомного газа

и, следовательно,

$$N' = 1 + \frac{e^2 n (-\omega^2 + \omega_0^2)}{2 \epsilon_0 m_e [(-\omega^2 + \omega_0^2)^2 + \gamma^2 \omega^2]},$$

$$N'' = \frac{e^2 n \gamma \omega}{2 \epsilon_0 m_e [(-\omega^2 + \omega_0^2)^2 + \gamma^2 \omega^2]}.$$

Зависимости N'' и N' от частоты представлены на рис. 16.1, а, б. Пунктирная кривая на рис. 16.1, б соответствует предельному случаю $\gamma = 0$. В этом случае $N'' = 0$ и показатель преломления обращается в бесконечность при $\omega = \omega_0$, т. е. когда частота волны совпадает с частотой колебаний электрона. При этом если частота ω стремится к ω_0 со стороны меньших значений, то $N \rightarrow +\infty$, если же со стороны больших значений, то $N \rightarrow -\infty$. Ситуация существенно изменяется при конечном γ : при этом ветви кривой не уходят на $+\infty$ и $-\infty$, а плавно переходят одна в другую. Поэтому если при $\gamma = 0$ показатель преломления всегда возрастал с ростом ω , то при $\gamma \neq 0$ вблизи $\omega = \omega_0$ возникает область, в которой $dN'/d\omega < 0$. Это так называемая область *аномальной дисперсии*; области же частот, в которых $dN'/d\omega > 0$, называют областями *нормальной дисперсии*. Функция $N''(\omega)$ имеет максимум вблизи $\omega = \omega_0$, так что область аномальной дисперсии соответствует большому поглощению волны.

16.4. Скин-эффект

Результаты предыдущего раздела позволяют рассмотреть проникновение переменных полей и токов в проводник — так называемый скин-эффект (skin по-английски означает «кожа»). Начнем с простейшей задачи, а именно исследуем проникновение электромагнитного поля из вакуума в металл, занимающий полупространство $z > 0$.

Заметим прежде всего, что если использовать понятие комплексного коэффициента преломления, то металл с проводимостью σ ничем не отличается от диэлектриков с мнимой частью проницаемости (относительной) $\epsilon''(\omega) = \sigma/(\omega\epsilon_0)$. В самом деле, обратимся к уравнению Максвелла в среде

$$\text{rot } \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{j},$$

и подставим в него $\mathbf{D} = \epsilon_0 \epsilon \mathbf{E}$ и $\mathbf{j} = \sigma \mathbf{E}$. Мы получим тогда уравнение

$$\operatorname{rot} \mathbf{H} = (-i\omega\varepsilon_0\varepsilon + \sigma)\mathbf{E},$$

ничем не отличающееся от уравнения для диэлектрика с проницаемостью

$$\tilde{\varepsilon} = \varepsilon + \frac{i}{\varepsilon_0\omega} \sigma.$$

Проводимость металлов очень велика, порядка 10^8 См·м⁻¹ и выше, поэтому вплоть до колоссальных частот выполняется неравенство $|\varepsilon| \ll \sigma/(\varepsilon_0\omega)$, означающее, что можно пренебречь током смещения по сравнению с током проводимости. Поэтому мнимая часть $\tilde{\varepsilon}$ практически не отличается от $\sigma/(\varepsilon_0\omega)$ и комплексный коэффициент преломления имеет вид

$$N = N' + iN'' = \left(\varepsilon + \frac{i}{\varepsilon_0\omega} \sigma \right)^{1/2} = (1+i) \left(\frac{\sigma}{2\varepsilon_0\omega} \right)^{1/2}.$$

В предыдущем параграфе мы видели, что амплитуды полей убывают в глубь образца по экспоненциальному закону $\exp(-z/\delta)$, где $\delta = c/(\omega N'')$. Подставляя сюда $N'' = [\sigma/(2\varepsilon_0\omega)]^{1/2}$, получим

$$\delta = c [\sigma/(2\varepsilon_0\omega)]^{-1/2}.$$

На расстоянии δ амплитуды полей убывают в e раз, и поэтому эта величина может служить мерой глубины проникновения электромагнитного поля в металл. Ее называют *глубиной скин-слоя*. Мы видим, что глубина скин-слоя убывает с увеличением частоты как $\omega^{1/2}$ и что она тем меньше, чем больше проводимость металла. Если взять медь, для которой при комнатных температурах $\sigma = 5 \cdot 10^8$ См·м⁻¹, то для частоты $\omega = 2\pi \cdot 50$ с⁻¹ глубина скин-слоя составит $\delta = 0,6$ см и $\delta = 0,06$ см для частоты $\omega = 2\pi \cdot 500$ с⁻¹.

Пока мы предполагали, что металл заполняет все полупространство $z > 0$. Ясно, однако, что наши результаты применимы и для ограниченного образца, если только его толщина велика по сравнению с глубиной скин-слоя.

Остановимся в связи с этим на задаче о протекании по проводнику переменного тока. Переменные токи всегда вызываются приложенными к проводникам переменными полями, поэтому рассмотренный нами скин-эффект имеет прямое отношение к хорошо известному на практике явлению вытеснения переменного тока к периферии проводника. Это явление приводит к увеличению сопротивления проводника (но не удельного сопротивления вещества!) по сравнению со случаем постоянного тока.

Оценим сопротивление металлического проводника переменному току. Мы должны, очевидно, считать, что ток течет не по всему поперечному сечению проводника, а только по тонкому скин-слою глубины δ (предполагается, что δ мала по сравнению с радиусом проводника). Это значит, что эффективная площадь поперечного сечения, по которой течет ток, равна $L\delta$, где L —периметр сечения. Эта площадь меньше истинной площади S в $S/(L\delta)$ раз; во столько же раз должно воз-

расти сопротивление R_ω переменному току по сравнению с сопротивлением R постоянному току:

$$R_\omega = R(L\delta/S).$$

Так как $\delta \sim \omega^{-1/2}$, то R_ω пропорционально $\sqrt{\omega}$.

Вытеснение переменного тока на периферию проводника должно приводить к уменьшению (при заданной силе тока) той части энергии магнитного поля тока, которая заключена внутри проводника. Но энергия магнитного поля тока, разделенная на квадрат силы тока, определяет коэффициент самоиндукции проводника. Уменьшение энергии поля должно, следовательно, приводить к уменьшению коэффициента самоиндукции проводника, по которому протекает переменный ток.

При расчете толщины скин-слоя мы считали справедливым закон Ома $\mathbf{j} = \sigma \mathbf{E}$, где σ — удельная электрическая проводимость для постоянного тока. Такое простое выражение для плотности тока справедливо, однако, только в случае однородного поля — поле не должно существенно меняться на расстояниях порядка длины свободного пробега электрона. Но в проводнике переменное поле меняется на расстояниях порядка глубины скин-слоя δ . Поэтому формулой $\mathbf{j} = \sigma \mathbf{E}$ в случае переменного поля можно пользоваться только тогда, когда $l \ll \delta$. Если же $\delta \leq l$, то поле существенно неоднородно на расстояниях порядка l , так что формула $\mathbf{j} = \sigma \mathbf{E}$ со статической проводимостью σ несправедлива. Так как с понижением температуры металла длина свободного пробега электронов растет, то понижение температуры способствует выполнению условия $\delta \leq l$, так же как и увеличение частоты поля.

Постараемся разобраться, как происходит проникновение переменного электромагнитного поля в металл в том случае, когда толщина скин-слоя мала по сравнению с длиной свободного пробега (этот случай называют *аномальным скин-эффектом*). Заметим с этой целью, что при $l \gg \delta$ электрон (если он движется под не очень малым углом к поверхности металла) проводит мало времени в электрическом поле и поэтому практически не взаимодействует с ним. Значит, для электропроводности существенны только те электроны, которые движутся параллельно поверхности металла или под малым углом θ к ней: $\theta \ll \theta_0$, где $\theta_0 \sim \delta/l$. Относительное число таких электронов пропорционально θ_0 , поэтому их плотность в скин-слое по порядку величины составляет $\tilde{n} \sim n\delta/l$, где n — общая плотность электронов проводимости.

Учтем теперь, что проводимость определяется плотностью электронов, «чувствующих» электрическое поле. Очевидно, эффективная проводимость $\tilde{\sigma}$ окажется по порядку величины равной

$$\tilde{\sigma} = \frac{\tilde{n}}{n} \sigma \sim \frac{\delta}{l} \sigma.$$

Подстановка этого выражения в общую формулу для глубины скин-слоя $\delta = c[\sigma\omega/(2\varepsilon_0)]^{-1/2}$ дает

$$\tilde{\delta} = \left(\frac{2\varepsilon_0 c^2 l}{\sigma \omega} \right)^{1/3}.$$

Мы видим, что глубина скин-слоя при аномальном скин-эффекте пропорциональна $\omega^{-1/3}$, а не $\omega^{-1/2}$, как в случае нормального скин-эффекта.

Обратим внимание на то, что в формулу для $\tilde{\delta}$ коэффициент электропроводности σ входит вместе с длиной свободного пробега l электрона в комбинации σ/l . Но $\sigma = e^2 n l / \rho_F$, где ρ_F — граничный импульс Ферми. Поэтому формула для $\tilde{\delta}$ может быть переписана в виде

$$\tilde{\delta} = \left(\frac{2\varepsilon_0 c^2 \rho_F}{e^2 n \omega} \right)^{1/3}.$$

Мы видим, что $\tilde{\delta}$ фактически не зависит от l : в формулу входят лишь такие величины, как плотность электронов проводимости и их граничный импульс (но не длина свободного пробега). Поэтому $\tilde{\delta}$ в отличие от δ не зависит от температуры.

Длина свободного пробега входит, однако, в критерий применимости формулы для $\tilde{\delta}$. Действительно, эта формула справедлива, если $\delta \ll l$, т. е. если $\sigma \omega l^2 / (2\varepsilon_0 c^2) \gg 1$. Подставляя сюда $\sigma = e^2 n l / \rho_F$, получим

$$\omega \gg \omega_0; \quad \omega_0 = \frac{2\varepsilon_0 c^2 \rho_F}{e^2 n l^3}.$$

Этому условию должны удовлетворять частоты для того, чтобы наблюдался аномальный скин-эффект. Полагая $n \sim 10^{23} \text{ см}^{-3}$, $\rho_F \sim 10^{14} \text{ кг} \cdot \text{м} \cdot \text{с}^{-1}$, $l \sim 10^{-15} \text{ м}$, получим $\omega_0 \sim 10^7 \text{ с}^{-1}$.

Легко видеть, что $\tilde{\delta} = l(\omega_0/\omega)^{1/3}$; $\delta = l(\omega_0/\omega)^{1/2}$. Первая формула справедлива при $\omega > \omega_0$, а вторая — при $\omega < \omega_0$. Мы видим, что $\delta > l$, $\tilde{\delta} < l$.

16.5. Фазовая и групповая скорости

Вернемся к понятию фазовой скорости. Как мы уже говорили, оно относится к плоской и строго монохроматической волне $\mathbf{E} = \mathbf{E}_0 \exp(i\mathbf{k}\mathbf{r} - i\omega t)$. Фаза такой волны $\varphi = \mathbf{k}\mathbf{r} - \omega t$ имеет при заданном значении t постоянные значения на плоскостях $\mathbf{k}\mathbf{r} = \text{const}$, перпендикулярных волновому вектору \mathbf{k} , и каждое значение фазы переносится в пространстве с постоянной скоростью

$$\frac{d\mathbf{r}}{dt} = \frac{\omega \mathbf{n}}{k} = \mathbf{v}_\Phi,$$

где $\mathbf{n} = \mathbf{k}/k$ — единичный вектор в направлении распространения волны \mathbf{k} . Это и есть фазовая скорость данной плоской монохроматической волны. Так как для такой волны $\omega = ck/\sqrt{\varepsilon\mu}$, то $v_\Phi = c/\sqrt{\varepsilon\mu}$, а так как ε и μ зависят, вообще говоря, от ω и \mathbf{k} , то фазовая скорость в общем случае является функцией частоты и волнового вектора, причем функцией комплексной (см. § 16.2, 16.3).

Но плоская монохроматическая волна существует во всем пространстве и во все времена и интенсивность ее повсюду и всегда одинакова. Ясно, что такие свойства волны нельзя рассматривать иначе, чем абстракцию. Поэтому возникает вопрос: какое отношение имеет эта волна к реальным волнам? Ответ гласит, что и в материальной среде, как и в вакууме, плоская монохроматическая волна может рассматриваться как элемент, или «кирпич», с помощью которого можно построить реальную волну, существующую конечное время в ограниченной области пространства.

Рассмотрим для простоты среду без поглощения, но обладающую частотной дисперсией диэлектрической проницаемости, так что фазовая скорость $v_\phi = v_\phi(\omega)$ является для нее некоторой вещественной функцией частоты. Пусть напряженность E имеет определенное направление и зависит только от одной пространственной координаты, скажем z , и от времени: $E = E(z, t)$. Согласно теореме Фурье, ее можно представить в виде суперпозиции плоских монохроматических волн:

$$E(z, t) = \int E_k e^{i(kz - \omega t)} dk,$$

где $\omega = kv_\phi$. Так как v_ϕ есть некоторая функция ω , то и $k = \omega/v_\phi$ — некоторая функция ω и, следовательно, $\omega = \omega(k)$ может рассматриваться как некоторая функция k . В написанном интеграле, определяющем $E(z, t)$, пределы, вообще говоря, бесконечны, так что общее выражение для напряженности поля имеет вид

$$E(z, t) = \int_{-\infty}^{\infty} E_k e^{i[kz - \omega(k)t]} dk,$$

где ω следует считать четной функцией k (т. е. $\omega(-k) = \omega(k)$), так как дисперсионные свойства среды не могут зависеть от направления распространения волны. Положив здесь $t = 0$, мы получим напряженность поля в начальный момент времени как функцию z :

$$E(z, 0) \equiv f(z) = \int_{-\infty}^{\infty} E_k e^{ikz} dk.$$

Согласно теореме Фурье, этот интеграл можно обратить и найти

$$E_k = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(z) e^{ikz} dz.$$

Эта формула определяет фурье-компоненту E_k по известному распределению напряженности поля $E(z, 0) \equiv f(z)$ в начальный момент времени. Зная же фурье-компоненты E_k , мы можем согласно написанному выше разложению Фурье найти поле во все моменты времени.

Таким образом, действительно, плоская монохроматическая волна представляет собой тот структурный элемент, который позволяет единым образом представить поле в виде суперпозиции этих элементов.

В интеграле Фурье компоненты с $k > 0$ соответствуют волнам, распространяющимся в положительном направлении оси z , а слагаемые с $k < 0$ — волнам, распространяющимся в отрицательном направлении этой оси.

Рассмотрим прежде всего простейший случай, когда отсутствует дисперсия, так что $\omega = v_\phi |k|$, где v_ϕ не зависит от ω и k . Переписав интеграл Фурье в виде

$$E(z, t) = \int_0^{\infty} E_k e^{i(kz - kv_\phi t)} dk + \int_{-\infty}^0 E_k e^{i(kz - |k|v_\phi t)} dk,$$

легко заключить, что первое слагаемое представляет собой некоторую функцию от $(z - v_\phi t)$, а второе — функцию от $(z + v_\phi t)$:

$$E(z, t) = f_1(z - v_\phi t) + f_2(z + v_\phi t).$$

Первое слагаемое описывает волну, бегущую вдоль положительной полуоси z , а второе — волну, бегущую вдоль отрицательной полуоси z . Волны эти не меняют, очевидно, формы в процессе своего распространения. Если, например, в момент времени $t=0$ задан электромагнитный импульс $f(z)$ и он распространяется вдоль положительной оси z , то значение поля в точке z в более поздние моменты времени определится формулой $E(z, t) = f(z - v_\phi t)$, т. е. импульс как бы сносится вдоль оси z с фазовой скоростью v_ϕ .

Но такая ситуация имеет место только в простейшем и идеализированном случае среды, не обладающей дисперсией проницаемости. Фактически же дисперсия есть всегда, и поэтому разным компонентам Фурье отвечает разная фазовая скорость, так что нельзя говорить об общей скорости движения суперпозиции плоских монохроматических волн. Подчеркнем, что

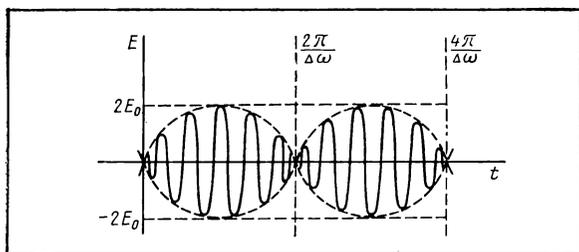
существование дисперсии не только экспериментальный факт, но и следствие общих физических принципов.

В самом деле, какова бы ни была структура вещества, его структурные элементы (например, атомы) не могут успевать реагировать на колебания очень высокой частоты, так что в пределе $\omega \rightarrow \infty$ диэлектрические восприимчивости должны обращаться в нуль и, следовательно, проницаемости — в единицу. Иными словами, для частот, превосходящих все характерные частоты среды, любая среда должна вести себя как вакуум.

Возвращаясь к вопросу о распространении суперпозиции плоских монохроматических волн, рассмотрим сперва суперпозицию двух таких волн с близкими частотами и волновыми векторами:

$$E(z, t) = E_0 e^{i(kz - \omega t)} + E_0 e^{i(k'z - \omega' t)},$$

где $k = (k_0 - \Delta k/2)$; $\omega = (\omega_0 - \Delta\omega/2)$; $k' = (k_0 + \Delta k/2)$; $\omega' = (\omega_0 + \Delta\omega/2)$, причем $\Delta k \ll k_0$ и $\Delta\omega \ll \omega_0$ (для простоты амплитуды волн считаются одинаковыми). Это выражение можно преобразовать к виду



16.2.

Биения при сложении двух волн равной амплитуды

$$E(z, t) = A(z, t) e^{i(k_0 z - \omega_0 t)}; A(z, t) = 2 E_0 \cos^{1/2}(\Delta k z - \Delta \omega t)$$

Второй множитель в выражении для E соответствует плоской монохроматической волне со средними значениями частоты ω_0 и волнового вектора k_0 (применяя термин радиотехники, ее можно назвать *несущей волной*). Функцию же $A(z, t)$ можно рассматривать как амплитуду этой волны; она медленно меняется в пространстве и времени от нуля до удвоенной амплитуды каждой из волн (в общем случае, когда амплитуды обеих волн не равны, — от разности до суммы амплитуд). Частота изменения результирующей амплитуды равна $\Delta\omega/2$; она значительно меньше частоты ω_0 несущей волны. Таким образом, в результате суперпозиции двух волн близкой частоты возникают *биения* с медленно меняющейся амплитудой (рис. 16.2).

Амплитуда биений остается неизменной, если $\Delta k z - \Delta \omega t = \text{const}$, или $z - (d\omega/dk)t = \text{const}$, поскольку $\Delta\omega$ и Δk малы. Иными словами, амплитуда биений переносится в пространстве со скоростью

$$v_g = d\omega/dk.$$

Эту скорость называют *групповой*. Она совпадает с фазовой скоростью $v_\phi = \omega/k$ только при линейной зависимости ω от k .

Рассмотрим теперь более общий случай, когда в суперпозиции участвуют не две, а непрерывная совокупность плоских монохроматических волн, частоты и волновые векторы которых меняются в узких интервалах $\omega_0 - \Delta\omega/2 < \omega < \omega_0 + \Delta\omega/2$ и $k_0 - \Delta k/2 < k < k_0 + \Delta k/2$, тогда

$$E(z, t) = \int_{k_0 - \Delta k/2}^{k_0 + \Delta k/2} E_k \exp \{i(kz - \omega t)\} dk.$$

Такая суперпозиция носит название *волнового пакета* (или *цуга*) волн (мы с ней уже сталкивались в гл. 15).

Учитывая, что $\Delta k \ll k$, можно приближенно заменить частоту ω на $\omega_0 + (d\omega/dk)_0(k - k_0)$, где $(d\omega/dk) = v_g$ — значение групповой скорости при $k = k_0$. Поэтому $E(z, t) = A(z, t) \exp \{i(k_0 z - \omega_0 t)\}$, где

$$A(z, t) = \int_{k_0 - \Delta k/2}^{k_0 + \Delta k/2} E_k \exp \{i(k - k_0)(z - v_g t)\} dk.$$

Как и в рассмотренном выше случае, экспоненциальный множитель в выражении для E отвечает плоской монохроматической волне со средними значениями частоты и волнового вектора, величина же $A(z, t)$ определяет амплитуду результирующей волны — волнового пакета. Эта величина медленно меняется в пространстве (на характерных расстояниях $\sim 1/\Delta k$) и во времени (с характерными временами $1/\Delta\omega$). Она представляет собой функцию одного переменного $z - v_g t$ (а не двух независимых переменных z и t), т. е. $A(z, t) = f(z - v_g t)$. Поэтому можно сказать, что амплитуда переносится в пространстве с постоянной скоростью, равной по значению групповой скорости в точке $k = k_0$. Иными словами, волновой пакет движется как единое целое, без изменения формы.

Вывод о неизменности формы волнового пакета является, однако, не точным, а приближенным. Дело в том, что при рассмотрении фурье-представления пакета мы заменили частоту ω на $\omega_0 + v_g(k - k_0)$, иными словами, учли лишь первые два слагаемых в разложении частоты в ряд Тейлора по степеням $k - k_0$:

$$\omega = \omega_0 + v_g(k - k_0) + \frac{1}{2}v_g'(k - k_0)^2 + \dots,$$

где $v_g' = dv_g/dk$. Чтобы оценить роль отброшенного следующего (третьего) слагаемого, заметим, что оно умножается в экспоненте на t ; поэтому, для того чтобы можно было им пренебречь, должно выполняться условие $v_g'(\Delta k)^2 t \ll 1$. За время t пакет проходит расстояние $L = v_g t$, так что это условие можно переписать в виде

$$L \ll L_0, \quad L_0 = [v_g / \{v_g'(\Delta k)^2\}]^{1/2}$$

Только на расстояниях, меньших L_0 , волновой пакет ведет себя как единое целое, движущееся с групповой скоростью. Перемещаясь, волновой пакет расплывается; это расплывание становится существенным, когда пакет перемещается на расстояние $\sim L_0$. На больших расстояниях понятие групповой скорости теряет смысл.

Рассмотрим в качестве иллюстрации среду, для которой частота ω связана с волновым вектором соотношением $\omega = \omega_0(1 + \frac{1}{2}a^2 k^2)$, где ω_0 и a — некоторые константы (такая дисперсионная зависимость характерна для плазменных колебаний и спиновых волн, см. гл. 17, 19). Пусть в начальный момент времени задан импульс в форме модулированной синусоидальной волны, огибающая которой есть $\exp[-z^2/(2d^2)]$:

$$E(z, 0) \equiv (z) = E_0 \cos k_0 z \exp[-z^2/(2d^2)].$$

Спрашивается, в течение какого времени этот импульс будет двигаться как единое целое, без изменения формы? Для ответа на этот вопрос разложим $f(z)$ в интеграл Фурье $f(z) = \int E_k \exp(ikz) dk$ и рассмотрим волновой пакет:

$$E(z, t) = \int E_k \exp(ikz - i\omega t) dk.$$

Тогда время существования неразмытого пакета определяется величиной $t_0 \sim (\Delta k)^{-2} (dv_g/dk)^{-1}$ (строго говоря, оно должно быть велико

по сравнению с t). В рассматриваемом случае разброс волновых векторов около значений $k = k_0$ по порядку величин равен $\Delta k \sim d^{-1}$, а групповая скорость $v_g = d\omega/dk = \omega_0 a^2 k$. Поэтому $t_0 \sim d^2/(\omega_0 a^2)$, т. е. пакет не расплывается, если $t \ll d^2/(a^2 \omega_0)$.

Аккуратный расчет показывает, что волновой пакет имеет при $t \neq 0$, так же как и при $t = 0$, форму модулированной синусоидальной волны с огибающей, определяющейся гауссовой кривой $\exp[-z^2/(2d_t^2)]$, где d_t — эффективная ширина кривой, но уже не постоянная, а зависящая от времени:

$$d_t^2 = d^2 (1 + \omega_0 a^2 t / d^2) e^2;$$

Она растет со временем, и соответственно пакет расплывается. Если $\omega_0 a^2 t / d^2 \ll 1$, то $d_t \simeq d$ и волновой пакет совпадает по форме с начальным импульсом.

Но дисперсия диэлектрической проницаемости — не единственная причина изменения формы волнового пакета в реальной среде. До сих пор мы считали среду непоглощающей, т. е. идеально прозрачной, поглощение же энергии поля тоже приводит к изменению формы пакета. Поэтому в условиях большого поглощения понятие групповой скорости вообще теряет смысл.

Под волновым пакетом мы понимали суперпозицию волн с небольшим разбросом волновых векторов (по сравнению со средним значением волнового вектора). Выясним теперь, какую область в пространстве занимает пакет. Рассмотрим для этого простейший пример — цуг конечной длины Δz . В этом случае

$$E(z, 0) \equiv f_1(z) = \begin{cases} E_0 \exp(i k_0 z) & (|z| \leq \Delta z/2), \\ 0 & (|z| > \Delta z/2) \end{cases}$$

при этом фурье-компонента

$$E_k = \frac{1}{2\pi} E_0 \int_{-\Delta z/2}^{\Delta z/2} \exp[i(k_0 - k)z] dz = \frac{E_0}{\pi(k - k_0)} \sin \frac{(k - k_0) \Delta z}{2}.$$

Обозначим $(k - k_0)\Delta z/2 = \xi$, тогда E_k пропорционально $(\sin \xi)/\xi$. График этой функции представлен на рис. 16.3. Она не мала (т. е. порядка единицы) в интервале ξ от $-\pi/2$ до $\pi/2$. Ширина этого интервала составляет π , поэтому ширина Δk интервала волновых векторов, где не мало E_k , составляет $\Delta k \sim 2\pi/\Delta z$. Таким образом,

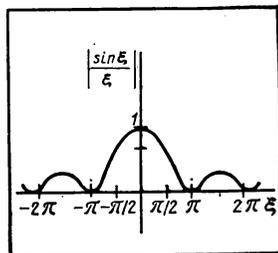
$$\Delta k \Delta z \sim 2\pi.$$

Такое же соотношение справедливо для любых волновых пакетов. Мы видим, что чем они уже, тем больше в них разброс волновых векторов. В общем случае, когда имеется трехмерный волновой пакет, размеры которого вдоль осей x, y, z равны $\Delta x, \Delta y, \Delta z$, для интервалов $\Delta k_x, \Delta k_y, \Delta k_z$, в которых не мала фурье-компонента поля, мы можем сразу написать

$$\Delta k_x \Delta x \sim 2\pi, \quad \Delta k_y \Delta y \sim 2\pi, \quad \Delta k_z \Delta z \sim 2\pi.$$

Возникает вопрос: нет ли аналогичного соотношения для частоты? Если обозначить Δt интервал времени, которое пакет проводит в данной точке пространства, то $\Delta t \sim \Delta z/v_g$, поэтому $v_g \Delta t \Delta k \sim 2\pi$. Но $v_g \Delta k \sim \Delta\omega$, где $\Delta\omega$ — разброс по частотам в пакете. Таким образом,

$$\Delta\omega \Delta t \sim 2\pi.$$



16.3.

График функции $\frac{\sin \xi}{\xi}$

Это соотношение связывает интервал времени прохождения пакета через данную точку пространства с частотным разбросом в пакете. Но оно имеет и другой, более общий и глубокий смысл, а именно: оно связывает время жизни каждого состояния любой физической системы с соответствующим этому состоянию разбросом по частоте.

Пусть, например, имеется импульс $f(t)$ (электромагнитный или какой-либо иной), существующий конечное время τ . В простейшем случае импульс экспоненциально затухает за время τ :

$$f(t) = Ae^{-t/\tau} \quad (t \geq 0);$$

следовательно, если $t < 0$, то $f(t) = 0$. Тогда для его фурье-компоненты имеем

$$f_\omega = \frac{1}{2\pi} \int_0^\infty f(t) e^{i\omega t} dt = \frac{\tau}{2\pi} (1 - i\omega\tau)^{-1}.$$

По модулю f_ω не мало в интервале частот $-\Delta\omega/2 \leq \omega \leq \Delta\omega/2$, где $\Delta\omega \sim 2\pi/\tau$. Иными словами, разброс частот в импульсе обратно пропорционален времени его жизни, причем по порядку величины $\Delta\omega\tau \sim 2\pi$.

На примере волнового пакета видно, сколь важно при использовании соотношения $\Delta\omega\Delta t \sim 2\pi$ уточнить, что понимается под Δt . В самом деле, частотный разброс $\Delta\omega \sim v_g \Delta k$ характеризует обратное время прохождения пакета через фиксированную точку пространства. Под временем же жизни пакета обычно понимается время значительно большее — от создания пакета до его расплывания, т. е. время, когда амплитуда поля не мала в системе отсчета, движущейся с пакетом. Легко видеть, что переходу в систему отсчета, движущуюся со скоростью \mathbf{v} , отвечает замена $\omega \rightarrow (\omega - \mathbf{k}\mathbf{v})$. В самом деле, если $f(\mathbf{r}, t)$ — какая-либо зависящая от координат и времени величина, то в движущейся системе эта же величина есть $f(\mathbf{r} - \mathbf{v}t, t)$. Но сдвигке $\mathbf{r} \rightarrow \mathbf{r} - \mathbf{v}t$ в обратном преобразовании Фурье эквивалентна замена $\omega \rightarrow \omega - \mathbf{k}\mathbf{v}$, так как

$$\begin{aligned} & \int f(\mathbf{r} - \mathbf{v}t, t) \exp(-i\mathbf{k}\mathbf{r} + i\omega t) d^3r dt = \\ & = \int f(\mathbf{r}, t) \exp\{-i\mathbf{k}\mathbf{r} + i(\omega - \mathbf{k}\mathbf{v})t\} d^3r dt. \end{aligned}$$

В системе отсчета, движущейся с групповой скоростью, разброс частот в пакете уже не равен $\Delta\omega \sim v_g \Delta k$: слагаемое, содержащее $(\Delta\omega - v_g \Delta k)$, в первом по Δk приближении исчезает, так что $\Delta\omega = = 1/2 v_g' (\Delta k)^2$. Вспоминая, что время жизни пакета есть $L_0/v_g \sim \sim v_g^{-1} (\Delta k)^{-2}$, мы видим, что соотношение $\Delta\omega \Delta t \sim 2\pi$ в движущейся системе отсчета приобретает более глубокий физический смысл, а именно: $\Delta\omega$ — это разброс по частотам (значительно меньший, чем в лабораторной системе), а Δt — это время жизни пакета (значительно большее, чем $(v_g \Delta k)^{-1}$).

Свяжем групповую скорость с показателем преломления N . Дифференцируя соотношение $\omega = ck/N$, получим

$$v_g \equiv d\omega/dk = \frac{c}{N} - \frac{ck}{N^2} \frac{dN}{d\omega} \frac{d\omega}{dk},$$

откуда

$$v_g = c \left[N \left(1 + \frac{ck}{N^2} \frac{dN}{d\omega} \right) \right]^{-1} = c \left[\frac{d}{d\omega} (N\omega) \right]^{-1}.$$

В области нормальной дисперсии (см. предыдущий параграф), по определению, $dN/d\omega > 0$. Далее, в проанализированных в предыдущем параграфе примерах в области нормальной дисперсии $N > 1$ (последнее неравенство имеет место в области нормальной дисперсии для большинства сред, хотя и не является обязательным). Поэтому в области нормальной дисперсии имеют место неравенства $v_g < v_\phi < < c$.

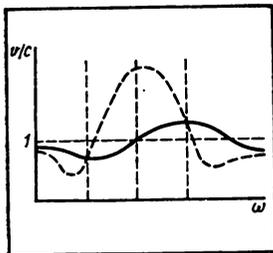
В области аномальной дисперсии $dN/d\omega < 0$, причем $|dN/d\omega|$ может быть очень велико. При этом величина v_g может быть и больше c (и даже стать бесконечной). Групповая скорость может быть и отрицательной — медленно меняющаяся амплитуда может переноситься в сторону, противоположную той, в которую переносится фаза волны. Во всех этих случаях пользоваться понятием групповой скорости нужно с большой осторожностью и при решении вопроса о движении волнового пакета следует обратиться к исходному представлению Фурье и исследовать его поведение.

Зависимость v_g и v_ϕ от частоты в области аномальной дисперсии представлена на рис. 16.4.

Итак, как фазовая, так и групповая скорости могут превышать скорость света в пустоте c . Но

каковы бы ни были значения v_ϕ и v_g и какова бы ни была форма начального импульса поля, передний край фронта волны всегда движется со скоростью c .

Этот результат является универсальным и не зависит не только от формы импульса или сигнала, но и от свойства среды — среда может быть как прозрачной, так и (в какой-то области частот) непрозрачной, может быть диэлектриком или проводником.



16.4.

Зависимости фазовой (сплошная кривая) и групповой (пунктирная кривая) скоростей от частоты в области аномальной дисперсии

Дело в том, что осцилляторы, из которых построена среда и которые обуславливают дисперсию проницаемостей, во время прохождения фронта волны еще не успеют сдвинуться с места и поэтому их движение не скажется на скорости фронта. Фронт любого сигнала проходит через среду, как через пустое пространство, и действие среды оказывается лишь позади фронта. Это приводит к согласию с требованием теории относительности о том, что скорость сигнала не может превосходить c (см. гл. 7). Вывод о том, что скорость фронта волны не может быть больше, чем c , может быть доказан строго математически: он является следствием того, что для любой среды при $\omega \rightarrow \infty$, как отмечалось выше, $\epsilon(\omega) \rightarrow 1$.

16.6. Энергия поля в диспергирующих средах

Выше мы показали (см. гл. 2 и 5), что для изменения электрической и магнитной индукций \mathbf{D} и \mathbf{B} на $d\mathbf{D}$ и $d\mathbf{B}$ в единичном объеме нужно затратить работу

$$dA = \frac{1}{2} (\mathbf{E}d\mathbf{D} + \mathbf{H}d\mathbf{B})$$

В случае статических полей электрическая индукция \mathbf{D} зависит только от электрического поля \mathbf{E} (и через диэлектрическую проницаемость может зависеть от температуры), а магнитная индукция \mathbf{B} зависит только от \mathbf{H} (и через магнитную проницаемость может зависеть от температуры; в особом типе кристаллов, называемых сегнетомагнетиками, \mathbf{D} зависит также и от \mathbf{H} , а \mathbf{B} — также и от \mathbf{E}). Поэтому при постоянной температуре dA является полным дифференциалом термодинамической функции F , называемой *свободной энергией*:

$$dA = dF, \quad F = \int (\mathbf{E}d\mathbf{D} + \mathbf{H}d\mathbf{B}),$$

здесь интегрирование совершается от нулевых до данных значений напряженностей \mathbf{E} и \mathbf{H} . В простейшем случае линейных связей $\mathbf{D} = \epsilon_0 \epsilon \mathbf{E}$, $\mathbf{B} = \mu_0 \mu \mathbf{H}$ свободная энергия F имеет вид

$$F = \frac{1}{2} (\mathbf{E}\mathbf{D} + \mathbf{H}\mathbf{B}) = \frac{1}{2} (\epsilon_0 \epsilon E^2 + \mu_0 \mu H^2)$$

(первое из этих равенств справедливо и при более общих линейных связях, а именно когда среда анизотропна и $D_i = \epsilon_0 \epsilon_{ik} E_k$, $B_i = \mu_0 \mu_{ik} H_k$, где ϵ_{ik} и μ_{ik} — тензоры проницаемостей среды).

Величина F представляет собой полную плотность свободной энергии поля в веществе, т. е. в F входит как энергия самого поля, так и свободная энергия вещества, связанная с полем.

Учитывая общее термодинамическое соотношение $F = W - TS$, где W и S — плотности энергии и энтропии, можно, зная F , определить и $W = F - T(\partial F / \partial T)$. Если ϵ и μ не зависят от температуры, то $F = W$. О величине W можно сказать то же самое, что и о F , а именно: W представляет собой сумму плотностей собственно электромаг-

нитой энергии, т. е. $(\epsilon_0 E^2 + \mu_0 H^2)/2$, и энергии вещества, связанной с рассматриваемым полем.

Ситуация осложняется, когда мы переходим от статических к переменным полям. Дело в том, что в случае переменных полей, как мы видели в § 16.2, векторы \mathbf{D} и \mathbf{B} , взятые в некоторый момент времени, зависят, как правило, от значений \mathbf{E} и \mathbf{H} не только в этот момент, но и от их значений в более ранние моменты времени. Это свойство среды (*временная дисперсия*) приводит к тому, что величина $dA = \mathbf{E}d\mathbf{D} + \mathbf{H}d\mathbf{B}$, сохраняя свой прежний смысл работы при изменении индукций, как правило, не представляет собой полного дифференциала какой-либо функции полей, которую можно было бы интерпретировать как энергию поля и связанную с полем энергию (или свободную энергию) вещества. Такая ситуация возможна только в специальном случае, когда и для переменного поля (в рассматриваемой области частот) имеют место те же соотношения $\mathbf{D} = \epsilon_0 \epsilon \mathbf{E}$ и $\mathbf{B} = \mu_0 \mu \mathbf{H}$ с постоянными ϵ и μ , как и для статического поля.

Но, согласно первому началу термодинамики, для любой физической системы работа, совершаемая над системой, равна сумме изменения внутренней энергии W и количества теплоты, выделяемой в системе. Поэтому и для поля в веществе можно написать соотношение

$$dA = \mathbf{E}d\mathbf{D} + \mathbf{H}d\mathbf{B} = dW + dQ,$$

где dQ — теплота, выделяемая в единичном объеме вещества при изменении векторов индукций на $d\mathbf{D}$ и $d\mathbf{B}$ (подчеркнем, что, хотя мы используем для выделяемой теплоты символ dQ , эта величина не является дифференциалом какой-либо функции). Ясно, что знание только dA не позволяет определить обе величины dW и dQ порознь, нужно знать еще динамику самого вещества, так как полная энергия W содержит как энергию поля, т. е. $(\epsilon_0 E^2 + \mu_0 H^2)/2$, так и энергию вещества в этом поле, а последнюю можно найти только если известны законы движения частиц вещества в поле.

Рассмотрим в качестве примера газ осцилляторов. Тогда под W нужно понимать величину

$$W = 1/2 (\epsilon_0 E^2 + \mu_0 H^2) + K + U,$$

где K и U — плотности кинетической и потенциальной энергий системы осцилляторов в электромагнитном поле:

$$K = n \frac{m \dot{\mathbf{r}}^2}{2}, \quad U = n \frac{m \omega_0^2 \mathbf{r}^2}{2}.$$

Здесь m и ω_0 — масса и собственная частота колебаний осциллятора, \mathbf{r} — его радиус-вектор (относительно положения равновесия) и n — плотность осцилляторов.

Если e — заряд осциллятора и γ — коэффициент затухания, то

$$\ddot{\mathbf{r}} + \gamma \dot{\mathbf{r}} + \omega_0^2 \mathbf{r} = \frac{e}{m} \mathbf{E}.$$

Определив из этого уравнения \mathbf{r} , можно вычислить плотность энергии W . Кроме того, можно найти количество теплоты Q , выделяющейся в единичном объеме газа осцилляторов за 1 с:

$$Q = nm\gamma\bar{r}^2.$$

Если поле является монохроматическим, т. е. $\mathbf{E} = E_0 \exp(-i\omega t)$, то

$$\mathbf{r} = -\frac{e}{m} E_0 \exp(-i\omega t) (\omega^2 - \omega_0^2 + i\gamma\omega)^{-1}.$$

Положив $\omega_0 = 0$, мы придем к случаю плазмы, в котором вычисления особенно просты. Именно: в этом случае для плотности энергии получим

$$W = \frac{\epsilon_0 E_0^2}{2} \left\{ \left[1 - \frac{\omega_e^2 (\omega^2 - \gamma^2)}{(\omega^2 + \gamma^2)^2} \right] \cos^2 \omega t + \frac{\gamma \omega \omega_e^2}{(\omega^2 + \gamma^2)^2} \sin 2\omega t + \frac{\omega^2 \omega_e^2}{(\omega^2 + \gamma^2)^2} \right\},$$

где $\omega_e^2 = e^2 n / (\epsilon_0 m)$ (магнитное слагаемое $\mu_0 H^2 / 2$ опущено). В единичном объеме за 1 с выделяется теплота

$$Q = \omega_e^2 \gamma \left[\frac{1}{\omega^2 + \gamma^2} - \frac{\omega^2 - \gamma^2}{(\omega^2 + \gamma^2)^2} \cos 2\omega t + \frac{2\gamma\omega}{(\omega^2 + \gamma^2)^2} \sin 2\omega t \right] \frac{\epsilon_0 E_0^2}{2}.$$

Величины W и Q являются, естественно, функциями времени; для их средних значений по периоду поля получим

$$\langle W \rangle = \left(1 + \frac{\omega_e^2}{\omega^2 + \gamma^2} \right) \frac{\epsilon_0 |E_0|^2}{4}, \quad \langle Q \rangle = \frac{\omega_e^2 \gamma}{\omega^2 + \gamma^2} \frac{\epsilon_0 |E_0|^2}{2}.$$

Запишем также формулы для $\langle W \rangle$ и $\langle Q \rangle$ и для общего случая газа осцилляторов, когда $\omega_0 \neq 0$:

$$\langle W \rangle = \left[1 + \frac{\omega_e^2 (\omega^2 + \omega_0^2)}{(\omega^2 - \omega_0^2) + \omega^2 \gamma^2} \right] \frac{\epsilon_0 |E_0|^2}{4},$$

$$\langle Q \rangle = \frac{\omega_e^2 \gamma \omega^2}{(\omega^2 - \omega_0^2)^2 + \omega^2 \gamma^2} \frac{\epsilon_0 |E_0|^2}{2}.$$

Обратимся теперь к выражению для диэлектрической проницаемости газа осцилляторов, полученному в § 16.2, и запишем его в форме

$$\epsilon = 1 - \frac{\omega_e^2 (\omega^2 - \omega_0^2 - i\gamma\omega)}{(\omega^2 - \omega_0^2)^2 + \gamma^2 \omega^2}.$$

Мы видим, что $\langle Q \rangle$ выражается через мнимую часть ϵ'' диэлектрической проницаемости:

$$\langle Q \rangle = \epsilon_0 \omega \epsilon''(\omega) |E_0|^2 / 2.$$

Эта важная формула оказывается справедливой не только для газа осцилляторов (или плазмы), но и для любой среды —

потери энергии переменного поля на нагревание среды всегда связаны таким же соотношением с мнимой частью диэлектрической проницаемости среды

(если, как, например, в случае ферромагнетика, существенна и магнитная проницаемость, то нужно еще добавить слагаемое $\mu_0 \mu''(\omega) |H_0|^2/2$).

Что касается средней плотности энергии поля $\langle W \rangle$, то она, вообще говоря, через диэлектрическую проницаемость среды не выражается. Только в том случае, когда потери энергии на нагревание среды равны нулю ($\langle Q \rangle = 0$), средняя плотность энергии $\langle W \rangle$ может быть выражена через $\epsilon(\omega)$.

В случае газа осцилляторов, полагая $\gamma = 0$ и сравнивая формулы для $\langle W_e \rangle$ и $\epsilon(\omega)$, легко убедиться в справедливости соотношения

$$\langle W_e \rangle = \frac{\epsilon_0}{4} \frac{d}{d\omega} [\omega \epsilon(\omega)] |E_0|^2.$$

Такое же соотношение имеет место и для других сред, если только $\langle Q \rangle = 0$. В этом случае среда является прозрачной, поэтому можно сказать, что средняя плотность энергии электрического поля пропорциональна $d(\omega \epsilon)/d\omega$ (и, естественно, квадрату напряженности поля) для прозрачной среды, или, более точно, в области прозрачности среды, т. е. для интервала частот, где $\langle Q \rangle = 0$. К написанному выражению для W_e мы должны добавить среднюю энергию магнитного поля, которая при $\mu = 1$ равна $\langle W_m \rangle = |H_0|^2 \mu_0/4$, а так как $\sqrt{\mu_0 \mu} H_0 = \sqrt{\epsilon_0 \epsilon} E_0$, то $\langle W_m \rangle = \epsilon_0 \epsilon |E_0|^2/4$. Поэтому средняя плотность энергии всего электромагнитного поля в диспергирующей среде (при $\mu = 1$) определяется формулой

$$\langle W \rangle = \langle W_e \rangle + \langle W_m \rangle = \left\{ \frac{d}{d\omega} [\omega \epsilon(\omega)] + \epsilon(\omega) \right\} \frac{\epsilon_0 |E_0|^2}{4}.$$

Мы считали поле монохроматическим, но приведенная формула для $\langle W \rangle$ справедлива и для пакета монохроматических волн, частоты которых лежат в узком интервале около некоторой частоты ω . Чтобы убедиться в этом, напомним, что скорость волнового пакета как целого определяется групповой скоростью $v_g = c / \left[\frac{d}{d\omega} (\omega N) \right]$, где $N = \sqrt{\epsilon(\omega)}$ — показатель преломления (предполагается, что $\mu = 1$). Если мы умножим v_g на среднюю плотность электромагнитной энергии $\langle W \rangle$, получим, очевидно, среднее значение $\langle S \rangle$ плотности потока энергии $S = [EH]$. Так как $\sqrt{\epsilon_0 \epsilon} E = \sqrt{\mu_0 \mu} H$, то

$$\langle S \rangle = c \sqrt{\epsilon_0 \epsilon(\omega)} |E_0|^2.$$

Итак, $\langle S \rangle = v_g \langle W \rangle$, откуда $\langle W \rangle = \langle S \rangle / v_g$. Подставляя сюда выражение для $\langle S \rangle$ и учитывая, что

$$\frac{1}{2} \left[\frac{d}{d\omega} (\omega N^2) + N^2 \right] = N \frac{d}{d\omega} (\omega N),$$

мы придем к найденной выше формуле для $\langle W \rangle$.

Этот вывод формулы для $\langle W \rangle$ еще раз подчеркивает важность предположения о прозрачности среды или, точнее, об очень малом поглощении энергии в среде; в противном случае теряет смысл, как мы видим в предыдущем параграфе, понятие групповой скорости, а вместе с ним и определение $\langle W \rangle$.

16.7. Отражение и преломление волн

В однородной и безграничной среде могут распространяться, как мы видели, плоские монохроматические волны. Если среда неоднородна, то монохроматические волны в ней, вообще говоря, уже не являются плоскими. Но существует один очень важный случай, когда и в неоднородной среде могут распространяться плоские монохроматические волны, — это случай двух однородных полубесконечных сред, соприкасающихся вдоль плоской границы. В этом параграфе мы рассмотрим такой случай.

Плоские волны удовлетворяют уравнениям Максвелла внутри каждого из полупространств, заполненных однородной средой. Поэтому вопрос сводится к тому, как удовлетворить граничным условиям на поверхности раздела сред:

$$\mathbf{E}_t^{(1)} = \mathbf{E}_t^{(2)}, \quad \mathbf{H}_t^{(1)} = \mathbf{H}_t^{(2)}, \quad D_n^{(1)} = D_n^{(2)}, \quad B_n^{(1)} = B_n^{(2)}$$

(индексы (1), (2) обозначают обе среды, а t , n — тангенциальные и нормальные компоненты полей). Этим условиям, как мы сейчас убедимся, можно удовлетворить с помощью трех плоских монохроматических волн (с одной и той же частотой), из которых две волны распространяются в одной из сред и одна — во второй среде.

Пусть в среде 1 из бесконечности по направлению к границе раздела сред (плоскость $z = 0$, рис. 16.5) распространяется плоская монохроматическая волна, для которой

$$\mathbf{E} = \mathbf{E}_0 \exp(ikr - i\omega t); \quad \mathbf{B} = \sqrt{\mu\epsilon} k^{-1} [\mathbf{kE}] \sqrt{\mu_0\epsilon_0}.$$

Тогда от границы в среде 2 распространяется волна

$$\mathbf{E}' = \mathbf{E}'_0 \exp(ik'r - i\omega t),$$

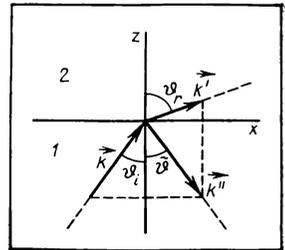
$$\mathbf{B}' = \sqrt{\mu'\epsilon'} (k')^{-1} [\mathbf{k}'\mathbf{E}'] \sqrt{\mu_0\epsilon_0}$$

и, кроме того, в среде 1 возникнет отраженная волна

$$\mathbf{E}'' = \mathbf{E}''_0 \exp(ik''r - i\omega t),$$

$$\mathbf{B}'' = \sqrt{\mu\epsilon} (k'')^{-1} [\mathbf{k}''\mathbf{E}''] \sqrt{\mu_0\epsilon_0},$$

распространяющаяся от границы раздела. Здесь \mathbf{k} , \mathbf{k}' , \mathbf{k}'' — волновые векторы падающей, прошедшей (или, как говорят, преломленной) и отраженной волн и ϵ , μ и ϵ' , μ' — диэлектрические и магнитные прони-



16.5. К выводу закона Снеллиуса (для определенности представлен случай $N > N'$)

щаемости обеих сред (не путать с вещественными частями проницаемостей, для которых в остальных разделах используются те же обозначения).

Покажем, что эти три волны, вместе взятые, удовлетворяют необходимым граничным условиям. При этом, например, под $\mathbf{E}^{(1)}$ следует понимать сумму напряженностей падающей и отраженной волн, а под $\mathbf{E}^{(2)}$ — напряженность преломленной волны:

$$\mathbf{E}^{(1)} = \mathbf{E} + \mathbf{E}''; \quad \mathbf{H}^{(1)} = \mathbf{H} + \mathbf{H}''; \quad \mathbf{D}^{(1)} = \mathbf{D} + \mathbf{D}''; \quad \mathbf{B}^{(1)} = \mathbf{B} + \mathbf{B}'';$$

$$\mathbf{E}^{(2)} = \mathbf{E}'; \quad \mathbf{H}^{(2)} = \mathbf{H}'; \quad \mathbf{D}^{(2)} = \mathbf{D}'; \quad \mathbf{B}^{(2)} = \mathbf{B}' \quad (z = 0).$$

Прежде всего ясно, что для выполнения граничных условий на поверхности раздела должны быть одинаковы экспоненциальные множители $\exp(ik_x x - i\omega t)$, $\exp(ik'_x x - i\omega t)$ и $\exp(ik''_x x - i\omega t)$, иначе не смогут выполняться равенства тангенциальных составляющих полей и нормальных составляющих индукций (волновой вектор падающей волны \mathbf{k} лежит в плоскости xz , так что $k_y = 0$ и, следовательно, $k'_y = k''_y = 0$; рис. 16.5). Иными словами, $k'_x = k''_x = k_x$. Эти соотношения означают, что *волновые векторы падающей, преломленной и отраженной волн лежат в одной плоскости, в которой лежит также нормаль \mathbf{n} к границе раздела*. Эту плоскость мы будем называть *плоскостью падения*.

Замечая, что $k_x = k \sin \theta_i$, $k'_x = k' \sin \theta_r$, $k''_x = k'' \sin \tilde{\theta}$, где θ_i , θ_r , $\tilde{\theta}$ — углы падения, преломления и отражения (рис. 16.5), и учитывая, что $k = k'' = \omega/v_\Phi$, $k' = \omega/v'_\Phi$, где $v_\Phi = c/\sqrt{\epsilon\mu}$ и $v'_\Phi = c/\sqrt{\epsilon'\mu'}$ — фазовые скорости волн в обеих средах, имеем

$$\sin \theta_i / v_\Phi = \sin \theta_r / v'_\Phi, \quad \tilde{\theta} = \theta_i.$$

Угол отражения равен углу падения, а отношение синусов углов преломления и падения равно отношению фазовых скоростей волн в соответствующих средах.

Последнее соотношение, называемое *законом Снеллиуса*, можно представить в виде

$$\sin \theta_i / \sin \theta_r = N' / N,$$

где $N = c/v_\Phi$ и $N' = c/v'_\Phi$ — показатели преломления обеих сред.

Мы видим, что если $N' < N$, то $\theta_r > \theta_i$.

Среду с большим показателем преломления называют *оптически более плотной*. Таким образом, волновой вектор расположен ближе к нормали \mathbf{n} в той из сред, которая является оптически более плотной. Направление волнового вектора совпадает с направлением луча, поэтому можно сказать, что луч, попадая в более плотную оптически среду, отклоняется в сторону нормали к границе раздела сред, а попадая в менее плотную — в сторону от нормали.

Если $N > N'$, то при увеличении угла падения θ_i мы сможем дойти до угла преломления $\theta_r = \pi/2$ (рис. 16.6). Соответствующий угол падения, который мы обозначим θ_c , определяется из условия

$$\sin \theta_c = N' / N.$$

Его называют углом полного внутреннего отражения. Название связано с тем, что при $\theta > \theta_c$ синус угла преломления формально становится больше единицы, из чего следует, что при $\theta > \theta_c$ преломленная волна отсутствует, а имеется только отраженная волна.

Такой вывод, однако, не совсем верен: при $\theta_i > \theta_c$ существует преломленная волна, но она сильно затухает, распространяясь в глубь второй среды. Действительно, рассмотрим пространственную зависимость поля волны во второй среде: $E' \sim \exp(ik'_x x + ik'_z z)$.

Так как $k'_x = k_x = k \sin \theta_i$, то эта величина вещественна. Что же касается $k'_z = \sqrt{(kN'/N)^2 - k_x'^2}$, то эта величина будет чисто мнимой, так как $N'/N = \sin \theta_c$ и, следовательно, $k'_z = ik\sqrt{\sin^2 \theta_i - \sin^2 \theta_c}$. Таким образом, все компоненты полей во второй среде ведут себя как $\exp\{-kz\sqrt{\sin^2 \theta_i - \sin^2 \theta_c} + ikx \sin \theta_i\}$, т. е. убывают по мере увеличения z . Можно сказать, что при полном внутреннем отражении волна проникает во вторую среду на глубину

$$l \sim \lambda (\sin^2 \theta_i - \sin^2 \theta_c)^{-\frac{1}{2}},$$

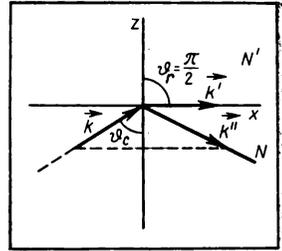
где $\lambda = 2\pi/k$ — длина волны в первой среде, причем глубина проникновения l сильно зависит от соотношения между θ_i и θ_c .

Вдоль оси x волна распространяется с волновым вектором $k \sin \theta_i$; не затухая. Поэтому при $\theta_i > \theta_c$ преломленная волна фактически проявляется как поверхностная волна. Заметим, что

при полном внутреннем отражении (в случае, если обе среды прозрачны) энергия поля не убывает со временем. Таким образом, полное внутреннее отражение представляет собой не затухание, а непропускание волны.

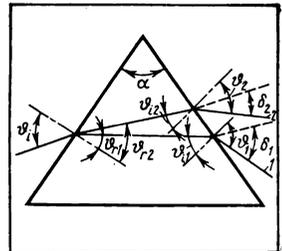
Показатель преломления зависит от частоты, или, что то же самое, от длины волны. Поэтому и отношение синусов углов падения и преломления зависит от длины волны. Отсюда следует, что если падающий луч представляет собой суперпозицию различных монохроматических волн, то они преломятся по-разному, в результате чего из одного падающего луча возникнет ряд (непрерывный или дискретный) пространственно разделенных преломленных лучей, каждый из которых будет уже монохроматическим.

Основываясь на этой идее, Ньютон (в 1666 г.) разложил пучок естественного света на его отдельные составляющие. Ньютон



16.6.

К определению угла полного внутреннего отражения θ_c



16.7.

Схема опыта по разложению белого света в призме

пользовался для этой цели стеклянной призмой, в которой луч света испытывал двойное преломление (рис. 16.7). Ясно, что при входе луча в призму $\sin\theta_{r1} = \sin\theta_i/N_1$, $\sin\theta_{r2} = \sin\theta_i/N_2$, где N_1 , N_2 — показатели преломления стекла для двух разных цветов 1 и 2, θ_{r1} , θ_{r2} — соответствующие углы преломления (θ_i — угол падения первоначального луча). Если α — угол при вершине призмы, то нормали к ее граням образуют угол $\pi - \alpha$, так что $\theta_{i1} = \alpha - \theta_{r1}$, $\theta_{i2} = \alpha - \theta_{r2}$; поэтому

$$\sin\theta_1 = \sin\theta_{i1}N_1 = \sin(\alpha - \theta_{r1})N_1,$$

$$\sin\theta_2 = \sin\theta_{i2}N_2 = \sin(\alpha - \theta_{r2})N_2$$

(θ_{i1} , θ_{i2} — углы падения, θ_1 , θ_2 — преломление при выходе луча из призмы). Легко убедиться, что угол δ между направлениями падающего и дважды преломленного луча (угол отклонения луча) тем больше, чем больше показатель преломления вещества призмы. Для стекла N растет с ростом частоты, поэтому фиолетовый луч отклоняется к основанию призмы сильнее, чем красный.

Угол δ минимален, если $\theta = \theta_i$ (θ — любой из углов θ_1 , θ_2 ; δ — соответствующий угол отклонения луча при двойном преломлении). Если обозначить δ_{\min} значение δ в минимуме, то, как нетрудно показать,

$$\sin \frac{\delta_{\min} + \alpha}{2} \Big/ \sin \frac{\alpha}{2} = N.$$

Эта формула может служить для экспериментального определения показателя преломления.

Все наши выводы основывались пока только на законе Снеллиуса, вытекающем из равенства экспонент типа $\exp(ik_x x - i\omega t)$ для различных волн на границе раздела двух сред. Но мы должны полностью использовать граничные условия, приравняв нормальные составляющие индукций и тангенциальные составляющие напряженностей полей по обе стороны границы. В результате получаются соотношения между амплитудами волн, с помощью которых амплитуды преломленной и отраженной волн можно выразить через амплитуду падающей волны. Эти соотношения существенно зависят от ориентации полей падающей волны относительно плоскости падения. При линейных связях между полями и индукциями, т. е. при не зависящих от полей проницаемостях достаточно рассмотреть две ориентации, или, иначе говоря, две поляризации: когда напряженность электрического поля падающей волны перпендикулярна плоскости (k , n) и когда она параллельна этой плоскости. Для первого случая мы получим

$$E'' = \frac{\sin(\theta_r - \theta_i)}{\sin(\theta_r + \theta_i)} E, \quad E' = \frac{2\sin\theta_r \cos\theta_i}{\sin(\theta_r + \theta_i)} E.$$

Во втором случае удобнее производить вычисления для напряженности магнитного поля (направленной при этом по нормали к плоскости падения).

Легко убедиться, что

$$H'' = \frac{\operatorname{tg}(\theta_i - \theta_r)}{\operatorname{tg}(\theta_i + \theta_r)} H, \quad H' = \frac{\sin 2\theta_i}{\sin(\theta_i + \theta_r) \cos(\theta_i - \theta_r)} N$$

(напомним, что E, H — напряженности поля падающей, E', H' — преломленной и E'', H'' — отраженной волн).

Эти соотношения называют формулами Френеля. С их помощью можно определить коэффициент отражения R , под которым понимают отношение проекций вектора Пойнтинга на нормаль к границе для отраженной и падающей волн. Для первой из рассмотренных поляризаций коэффициент отражения определяется формулой

$$R_{\perp} = \frac{\sin^2(\theta_r - \theta_i)}{\sin^2(\theta_r + \theta_i)},$$

а для второй — формулой

$$R_{\parallel} = \frac{\operatorname{tg}^2(\theta_r - \theta_i)}{\operatorname{tg}^2(\theta_r + \theta_i)}.$$

Из последней формулы видно, что если $\theta_r + \theta_i = \pi/2$, то $R_{\parallel} = 0$. Это условие означает, что $\cos\theta_r = \sin\theta_i$, а так как $\sin\theta_i/\sin\theta_r = N'/N$, то коэффициент R_{\parallel} обращается в нуль при $\theta_i = \theta_0$, где

$$\operatorname{tg} \theta_0 = N'/N$$

При $\theta_i = \theta_0$ в отраженной волне полностью отсутствует одна из поляризаций; иными словами, отраженный свет полностью поляризован (напряженность электрического поля в нем нормальна к плоскости падения). Поэтому угол θ_0 называют *углом полной поляризации* (другой термин — *угол Брюстера*, в честь ученого, обнаружившего в 1815 г. это явление). Заметим, что при $\theta_i = \theta_0$ отраженный и преломленный лучи перпендикулярны друг другу.

ОСНОВНЫЕ ФОРМУЛЫ

Связь между амплитудами напряженностей полей в плоской волне

$$\sqrt{\varepsilon_0 \varepsilon} E_0 = \sqrt{\mu_0 \mu} H_0,$$

$$E_0 \perp H_0$$

Комплексный показатель преломления и комплексная фазовая скорость

$$N = \sqrt{\varepsilon(\omega, \mathbf{k}) \mu(\omega, \mathbf{k})},$$

$$v_{\phi} = c/N$$

Глубина скин-слоя при нормальном скин-эффекте

$$\delta = c \left(\frac{2\varepsilon_0}{\sigma \omega} \right)^{1/2}$$

Глубина скин-слоя при аномальном скин-эффекте

$$\tilde{\delta} = \left(\frac{2\varepsilon_0 c^2 p_F}{e^2 n \omega} \right)^{1/2}$$

Групповая скорость электромагнитной волны

$$v_g \equiv \frac{d\omega}{dk} = c \left[\frac{d}{d\omega} (N\omega) \right]^{-1}$$

Соотношения неопределенностей в оптике	$\Delta k_x \Delta x \sim 2\pi, \Delta k_y \Delta y \sim 2\pi,$ $\Delta k_z \Delta z \sim 2\pi, \Delta \omega \Delta t \sim 2\pi$
Средняя плотность электромагнитной энергии в диспергирующей среде	$W = \frac{\epsilon_0}{4} \left[\frac{d}{d\omega} (\omega \epsilon(\omega)) + \epsilon(\omega) \right] E_0 ^2$
Закон Снеллиуса	$\frac{\sin \theta_i}{\sin \theta_r} = \frac{N_r}{N_i}$
Синус угла полного внутреннего отражения	$\sin \theta_c = N_r/N_i$
Угол полной поляризации (угол Брюстера)	$\theta_0 = \text{arctg}(N_r/N_i)$



Глава 17. КОЛЕБАНИЯ ПЛАЗМЫ

17.1. Ленгмюровские колебания

Наряду с твердым, жидким и газообразным состояниями вещество может находиться еще в другом своеобразном состоянии — в виде *ионизированной плазмы*, состоящей из двух заряженных газов: газа электронов и газа ионов, между которыми действуют электрические силы. В равновесии плазма в среднем электрически нейтральна — заряды электронов и ионов взаимно компенсируются; если же, переместив часть электронов, создать в плазме возмущение плотности заряда, то в ней возбуждаются особые высокочастотные электростатические волны — так называемые *ленгмюровские колебания*.

В ленгмюровских колебаниях принимают участие главным образом электроны; ионы можно считать покоящимися, и они образуют однородный фон положительного заряда. Нескомпенсированный электронный заряд ρ вызывает электрическое поле \mathbf{E} , причем согласно уравнению Пуассона

$$\text{div } \mathbf{E} = \rho/\epsilon_0.$$

Под действием этого поля электроны приобретают ускорение

$$\dot{\mathbf{v}} = (e/m_e)\mathbf{E},$$

где e и m_e — заряд и масса электрона. Если бы мы хотели учесть столкновение между частицами, то могли бы ввести в это уравнение силу трения, пропорциональную скорости \mathbf{v} . Мы не будем делать этого: нас интересует случай бесстолкновительной плазмы, когда время свободного пробега частиц велико по сравнению с периодом колебаний.

Будем считать возмущение малым; тогда можно записать плотность тока в виде $\mathbf{j} = en_0\mathbf{v}$, где n_0 — равновесная плотность электронов. Поэтому $\partial\mathbf{j}/\partial t = e^2n_0\mathbf{E}/m_e$, так что

$$\operatorname{div} \frac{\partial\mathbf{j}}{\partial t} = \frac{e^2n_0}{m_e} \operatorname{div} \mathbf{E} = \frac{e^2n_0}{\epsilon_0 m_e} \rho.$$

С другой стороны, закон сохранения заряда приводит к уравнению непрерывности

$$\frac{\partial\rho}{\partial t} + \operatorname{div} \mathbf{j} = 0,$$

поэтому для $\partial\mathbf{j}/\partial t$ можно записать и другое выражение:

$$\operatorname{div} \frac{\partial\mathbf{j}}{\partial t} = -\frac{\partial^2\rho}{\partial t^2}.$$

Сравнив эти два выражения, получим уравнение для малых колебаний:

$$\frac{\partial^2\rho}{\partial t^2} + \frac{e^2n_0}{\epsilon_0 m_e} \rho = 0.$$

Это уравнение описывает, очевидно, колебания осциллятора с частотой ω_e , равной

$$\omega_e = \sqrt{\frac{e^2n_0}{\epsilon_0 m_e}}.$$

Эту частоту называют *ленгмюровской* (или *плазменной*).

Частота ω_e велика даже в плазме невысокой плотности. Например, при плотности электронов $n_0 = 10^{20} \text{ м}^{-3}$ имеем $\omega_e = 5,6 \cdot 10^{11} \text{ с}^{-1}$; из электромагнитных волн в вакууме такую частоту имеет волна длиной 3 мм.

17.2. Диэлектрическая проницаемость электронной плазмы

Мы пока не учитывали теплового движения электронов плазмы; при этом ленгмюровские колебания оказались незатухающими и лишенными дисперсии — частота их не зависит от волнового вектора. Если учесть тепловое движение электронов, то возникнет дисперсия этих колебаний и еще одно очень интересное явление — затухание колебаний даже в отсутствие столкновений между частицами.

В не очень плотной плазме равновесное распределение электронов по скоростям является максвелловским:

$$f(\mathbf{v}, \mathbf{r}, t) = f_M(v) = A \exp\left(-\frac{m_e v^2}{2T}\right), \quad A = \frac{n_0}{4\pi} \left(\frac{m_e}{2\pi T}\right)^{3/2}$$

(функция распределения предполагается нормированной согласно условию $\iint f d^3v = n_0$; температура предполагается выраженной в

энергетических единицах). Если плазма очень плотная, то распределение является фермиевским, т. е. таким же, как и распределение электронов в металле (см. гл. 9). Однако такую плотную, как говорят, *вырожденную плазму*, мы рассматривать не будем.

Представим себе теперь, что в некоторый момент времени распределение электронов претерпевает изменение, так что их функция распределения перестает быть максвелловской и начинает зависеть помимо скорости еще от времени t и координат электрона \mathbf{r} . Спрашивается: каким образом можно найти неравновесную функцию распределения $f(\mathbf{v}, \mathbf{r}, t)$? Для этого служит кинетическое уравнение, с которым мы уже сталкивались в гл. 9 при изучении электронов в металле. Повторяя приведенные там рассуждения, можно сказать, что в результате столкновений число электронов в момент времени $t + dt$ в элементе объема фазового пространства $d^3v d^3r$ станет равным $[f(\mathbf{v}, \mathbf{r}, t) - (\partial f / \partial t)^{\text{coll}} dt] d^3v d^3r$, где $(\partial f / \partial t)^{\text{coll}}$ — интеграл столкновений, а так как частицы не рождаются и не исчезают, то эта величина должна составить $f(\mathbf{v} + \dot{\mathbf{v}} dt, \mathbf{r} + \mathbf{v} dt, t + dt) d^3v d^3r$. Учитывая, что

$$f(\mathbf{v} + \dot{\mathbf{v}} dt, \mathbf{r} + \mathbf{v} dt, t + dt) = f(\mathbf{v}, \mathbf{r}, t) + \frac{\partial f}{\partial v} \dot{\mathbf{v}} dt + \\ + \frac{\partial f}{\partial \mathbf{r}} \mathbf{v} dt + \frac{\partial f}{\partial t} dt,$$

где $\dot{\mathbf{v}} = \mathbf{F}$ — сила, действующая на частицу, и приравнявая оба выражения, получим

$$\frac{\partial f}{\partial t} + \mathbf{v} \frac{\partial f}{\partial \mathbf{r}} + \frac{\mathbf{F}}{m_e} \frac{\partial f}{\partial \mathbf{v}} = (\frac{\partial f}{\partial t})^{\text{coll}}.$$

Это и есть кинетическое уравнение для определения функции распределения электронов. Входящий сюда интеграл столкновений существенно зависит от свойств рассматриваемой системы, и структура его в случае плазмы иная, чем в случае металла. Мы не станем здесь выяснять эту структуру, потому что далее будем интересоваться только случаем так называемой бесстолкновительной плазмы, когда столкновениями можно вообще пренебречь. В этом случае кинетическое уравнение имеет вид

$$\frac{\partial f}{\partial t} + \mathbf{v} \frac{\partial f}{\partial \mathbf{r}} + \frac{\mathbf{F}}{m_e} \frac{\partial f}{\partial \mathbf{v}} = 0.$$

Этим уравнением — его называют уравнением Власова — можно пользоваться в случае высокочастотных возмущений, когда $\omega \tau \gg 1$, где ω — частота возмущения функции распределения электронов и τ — среднее время между столкновениями. В самом деле, по порядку величины $\partial f / \partial t \sim \omega \delta f$, $(\partial f / \partial t)^{\text{coll}} \sim \delta f / \tau$, где δf — возмущение функции распределения, поэтому неравенство $\omega \tau \gg 1$ позволяет пренебречь интегралом столкновений по сравнению со слагаемым $\partial f / \partial t$.

Воспользуемся уравнением Власова для определения высокочастотной электрической восприимчивости плазмы. Введем для этого

слабое возмущающее плазму электрическое поле, которое будем считать потенциальным (но не статическим!). Потенциал предположим имеющим вид плоской монохроматической волны $\varphi = \varphi_0 \exp\{i\mathbf{k}\mathbf{r} - i\omega t\}$, где φ_0 — комплексная амплитуда потенциала, \mathbf{k} и ω — волновой вектор и частота волны. Напряженность электрического поля \mathbf{E} связана с φ соотношением $\mathbf{E} = -\text{grad}\varphi$, поэтому

$$\mathbf{E} = \mathbf{E}_0 \exp\{i\mathbf{k}\mathbf{r} - i\omega t\}, \quad \mathbf{E}_0 = -i\mathbf{k}\varphi_0$$

Поле это направлено вдоль \mathbf{k} , т. е. является продольным.

Возмущение функции распределения $\delta f = f - f_M$ будем искать также в виде плоской монохроматической волны. Пренебрегая квадратичными по амплитуде возмущения слагаемыми и полагая $\mathbf{F} = e\mathbf{E}$, получим из уравнения Власова

$$(-i\omega + i\mathbf{k}\mathbf{v})\delta f - i\frac{e}{m}\varphi_0\mathbf{k}\frac{\partial f_M}{\partial \mathbf{v}} = 0.$$

Так как $\partial f_M / \partial \mathbf{v} = -m\mathbf{v}f_M/T$, то

$$\delta f = \frac{e\varphi_0}{T}f_M \frac{\mathbf{k}\mathbf{v}}{\omega - \mathbf{k}\mathbf{v}}.$$

С δf связано возмущение плотности заряда $\delta\rho$:

$$\delta\rho = e \int \delta f d^3v = \frac{e\varphi_0}{T} \int \frac{\mathbf{k}\mathbf{v}}{\omega - \mathbf{k}\mathbf{v}} f_M(v) d^3v.$$

Найдем теперь вектор электрической поляризации \mathbf{P} . Согласно гл. 6, $\text{div}\mathbf{P} = \delta\rho$, а так как все величины пропорциональны $\exp\{i\mathbf{k}\mathbf{r} - i\omega t\}$, то $\text{div}\mathbf{P} = i\mathbf{k}\mathbf{P}$. Вектор \mathbf{P} (как и векторы \mathbf{E} и \mathbf{D}) направлен вдоль \mathbf{k} , т. е. $\mathbf{P} = \mathbf{k}P/k$. Поэтому

$$P = \frac{\delta\rho}{ik} = -\frac{e^2E}{Tk^2} \int \frac{\mathbf{k}\mathbf{v}}{\omega - \mathbf{k}\mathbf{v}} f_M(v) d^3v.$$

Вспоминая, что $P = \varepsilon_0\alpha E$, где α — диэлектрическая восприимчивость среды, а диэлектрическая проницаемость $\varepsilon = 1 + \alpha$, получим

$$\varepsilon(\omega, \mathbf{k}) = 1 + \alpha(\omega, \mathbf{k}) = 1 - \frac{e^2}{\varepsilon_0 Tk^2} \int \frac{\mathbf{k}\mathbf{v}}{\omega - \mathbf{k}\mathbf{v}} f_M(v) d^3v.$$

Зная диэлектрическую проницаемость, мы можем определить частоту собственных колебаний плазмы. В самом деле, в отсутствие сторонних зарядов согласно уравнению Пуассона, $\mathbf{D} = 0$. Но $\mathbf{D} = \varepsilon_0\varepsilon\mathbf{E}$; поэтому

для того, чтобы в отсутствие сторонних зарядов в среде было ненулевое электростатическое поле, требуется, чтобы обратилась в нуль диэлектрическая проницаемость среды:

$$\varepsilon(\omega, \mathbf{k}) = 0$$

Это соотношение позволяет, в принципе, определить ω как функцию \mathbf{k} . Величина ω и представляет собой собственную частоту колебаний плазмы.

Рассмотрим сперва «холодную» плазму, в которой средняя тепловая скорость электронов достаточно мала. В этом случае подынтегральное выражение в формуле для α можно разложить в ряд по степеням kv :

$$\frac{kv}{\omega - kv} = \frac{kv}{\omega} \left[1 + \frac{kv}{\omega} + \left(\frac{kv}{\omega} \right)^2 + \dots \right].$$

При интегрировании по скоростям нечетные слагаемые пропадают; вклад второго и четвертого слагаемых легко вычислить, в результате мы получим следующее выражение для ϵ :

$$\epsilon(\omega, \mathbf{k}) = 1 - \frac{e^2 n_0}{\epsilon_0 m_e \omega^2} \left(1 + \frac{3Tk^2}{m_e \omega^2} \right).$$

Но величина $\sqrt{e^2 n_0 / (\epsilon_0 m_e)} = \omega_e$ есть как раз введенная в предыдущем параграфе ленгмюровская частота; поэтому если положить $T = 0$, то из уравнения $\epsilon = 0$ мы получим $\omega = \omega_e$. Во втором слагаемом в скобках можно вместо ω^2 подставить ω_e^2 ; замечая, что $[T\epsilon_0 / (e^2 n_0)]^{1/2} \equiv r_D$ есть введенный в гл. 10 радиус Дебая, найдем

$$\omega = \omega_e \sqrt{1 + 3r_D^2 k^2} \approx \omega_e \left(1 + \frac{3}{2} r_D^2 k^2 \right).$$

Эта формула определяет закон дисперсии плазменных колебаний. Пользоваться ею можно, если $kr_D \ll 1$, а так как $k = 2\pi/\lambda$, где λ — длина волны, то должно выполняться неравенство $\lambda \gg r_D$.

17.3. Ионный звук

Равновесное максвелловское распределение устанавливается в газе в результате столкновений между частицами, поэтому бесстолкновительная плазма, в которой время свободного пробега частиц очень велико, может долго находиться в неравновесном состоянии. Особенно долго может сохраняться двухтемпературное состояние, когда распределения частиц максвелловские, но температура электронов T_e не равна температуре ионов T_i . При этом в эксперименте обычно $T_e \gg T_i$ — электроны горячие, а ионы холодные. Это связано с тем, что при многих методах нагрева плазмы первоначально нагреваются электроны, а затем энергия постепенно передается от электронов к ионам. Если написать законы сохранения энергии и импульса для столкновений между частицами плазмы, то можно убедиться, что обмен энергиями между частицами одного сорта происходит в m_i/m_e раз быстрее, чем обмен энергиями между электронами и ионами (m_e и m_i — массы электрона и иона). Поэтому быстрее всего успевает установиться максвелловское распределение в каждом из газов (электронном и ионном) порознь, а затем температуры газов T_e и T_i медленно приближаются друг к другу до тех пор, пока плазма не перейдет в состояние полного равновесия с единой температурой электронов и ионов.

В бесстолкновительной плазме с горячими электронами и холодными ионами существуют наряду с высокочастотными ленгмюровски-

ми колебаниями еще и низкочастотные колебания, в которых принимают участие не только электроны, но и ионы. Фазовая скорость ω/k низкочастотных колебаний значительно меньше тепловой скорости электронов $\langle v_e \rangle = \sqrt{T_e/m_e}$, но значительно больше тепловой скорости ионов $\langle v_i \rangle = \sqrt{T_i/m_i}$, $\langle v_i \rangle \ll \omega/k \ll \langle v_e \rangle$.

Так как колебания происходят с низкой частотой, то электроны успевают «подстраиваться» в каждый момент времени к существующему в этот момент полю, т. е. можно считать, что они находятся в состоянии равновесия. Поэтому их плотность $n_e(\mathbf{r}, t)$ определяется распределением Больцмана: $n_e(\mathbf{r}, t) = n_0 \exp[-e\varphi(\mathbf{r}, t)/T_e]$, где φ — потенциал поля. В слабом поле

$$n_e(\mathbf{r}, t) = n_0 [1 - e\varphi(\mathbf{r}, t)/T_e].$$

Величина n_0 имеет, очевидно, смысл невозмущенной плотности электронов (совпадающей с невозмущенной плотностью ионов; заряд иона для простоты считается единичным). Отсюда следует, что возмущение электронной плотности

$$\delta n_e = -(en_0/T_e)\varphi.$$

Определим теперь возмущение ионной плотности δn_i . Ионная плотность удовлетворяет, очевидно, уравнению непрерывности

$$\frac{\partial n_i}{\partial t} + \operatorname{div} n_i \mathbf{u}_i = 0,$$

где \mathbf{u}_i — гидродинамическая скорость ионов, удовлетворяющая уравнению $d\mathbf{u}_i/dt = \frac{e}{m_i} \operatorname{grad}\varphi$. Если колебания малы и все величины изменяются по закону $\exp[i(\mathbf{k}\mathbf{r} - \omega t)]$, то $\mathbf{u}_i = -ek\varphi/(m_i\omega)$ и поэтому для δn_i получим

$$\delta n_i = n_0 k\mathbf{u}_i/\omega = -ek^2 n_0 \varphi/(m_i\omega^2).$$

Обратимся теперь к уравнению Пуассона $\Delta\varphi = -\rho/\epsilon_0$ и учтем, что невозмущенная плотность заряда равна нулю, поэтому $\Delta\varphi = e(\delta n_e - \delta n_i)/\epsilon_0$. Подставляя сюда найденные выражения для δn_e и δn_i и замечая, что $\Delta\varphi = -k^2\varphi$, имеем

$$-k^2\varphi = \frac{e^2 n_0 \varphi}{\epsilon_0 T_e} \left(1 - \frac{T_e}{m_i} \frac{k^2}{\omega^2}\right),$$

откуда

$$\omega \equiv \omega_s(k) = \frac{kv_s}{\sqrt{1 + k^2 r_D^2}}, \quad v_s = \sqrt{\frac{T_e}{m_i}}.$$

Мы получили колебания с частотой $\omega_s(k)$. Их фазовая скорость ω_s/k , как и утверждалось, мала по сравнению с $\langle v_e \rangle$, так как $m_e \ll m_i$. Условие же $\omega_s/k \gg \langle v_i \rangle$ выполняется в случае сильно неравновесной плазмы, когда $T_e \gg T_i(1 + k^2 r_D^2)$. При нарушении этого неравенства фазовая скорость колебаний имеет порядок теп-

ловой скорости ионов и колебания сильно затухают из-за резонансного поглощения ионами (см. § 17.4).

В области длинных волн ($kr_D \ll 1$) эти колебания характеризуются линейным законом дисперсии $\omega_s(k) = kv_s$ и поэтому их называют *ионным звуком*, а величину v_s — *скоростью ионного звука*. В коротковолновой области ($kr_D \gg 1$) частота ионного звука близка к ионной плазменной частоте $\omega_i = \sqrt{e^2 n_0 / (\epsilon_0 m_i)}$.

Обратим, однако, внимание на существенное различие между ионным и обычным звуками. Частота обычного звука мала по сравнению с частотой столкновений; при этом колебания происходят адиабатно и их фазовая скорость равна $\gamma \sqrt{T/m_i}$, где γ — показатель адиабаты. Наоборот, частота ионного звука (хотя и малая по сравнению с ленгмюровской, что позволяет назвать их низкочастотными) велика по сравнению с частотой столкновений; поэтому эти колебания происходят изотермически.

17.4. Резонансное взаимодействие волн и частиц и затухание Ландау

В плазме, как и во всяком газе, происходят столкновения между частицами. Эти столкновения приводят, естественно, к затуханию колебаний частиц, а следовательно, и к затуханию электромагнитных волн, распространяющихся в плазме, в частности ленгмюровских и ионно-звуковых.

Но столкновения — не единственный механизм затухания плазменных колебаний. В разреженной плазме он играет ничтожную роль, а определяющим становится другой механизм, которого нет в обычном газе, состоящем из нейтральных атомов. Этот механизм, действующий в бесстолкновительной плазме, заключается во взаимодействии частиц с полем волны. Взаимодействие это происходит особенно эффективно в том случае, когда фазовая скорость волны $v_\phi = \omega(k)/k$ совпадает (или близка) с проекцией скорости частицы v_k на направление волнового вектора ($v_k = v_\phi$), т. е. когда

$$\omega(k) = v_k k = kv.$$

Частицы, для которых выполняется это условие, называют *резонансными*, а само взаимодействие частиц с волной при этом называют *резонансным*.

Рассмотрим подробнее резонансное взаимодействие частиц с ленгмюровской волной $\phi = \phi_0 \cos(kx - \omega t)$, где ϕ — потенциал поля волны и ϕ_0 — его амплитуда (волна распространяется вдоль оси x). Вследствие этого взаимодействия, как мы сейчас убедимся, поле затухает, т. е. ϕ_0 изменяется со временем по закону $\exp(-\gamma t)$, где γ — коэффициент затухания, но затухание очень мало и мы сперва пренебрежем им.

Перейдем в систему отсчета, движущуюся с фазовой скоростью волны v_ϕ . Поле в этой системе имеет вид $\phi(x') = \phi_0 \cos kx'$, где $x' = x - v_\phi t$, и представляет собой совокупность статических ям и гор-

бов. Для электронов, движущихся в таком поле, имеет место закон сохранения энергии:

$$mv'^2/2 + e\varphi(x') = W = \text{const},$$

где $v' = (v_k - v_\phi)$ — проекция скорости частицы на ось x в этой системе.

Из этого соотношения видно, как происходит обмен энергией между частицами и полем. Ясно, что у частиц с кинетической энергией $mv'^2/2$, значительно превышающей потенциальную энергию $e\varphi_0$, обмена энергией в среднем не происходит. Напротив, частицы с малой энергией (меньшей или порядка $e\varphi_0$) захватываются в потенциальную яму и сильно взаимодействуют с полем. Точнее, сильно взаимодействуют с волной частицы, скорости которых v_k отличаются от v_ϕ на величину, меньшую или порядка $\Delta v_k \equiv \sqrt{e\varphi_0/m}$. Если при этом $v' > 0$, т. е. частица обгоняет волну, то она отдает энергию волне; если же $v' < 0$, т. е. частица отстает от волны, то она получает энергию от волны.

Отсюда можно заключить, что работа поля над частицами определяется разностью чисел частиц, обгоняющих волну и отстающих от нее. Скорости частиц близки к резонансной скорости v_ϕ , а так как величина Δv_k мала, то эта разность пропорциональна производной от функции распределения электронов по v_k при $v_k = v_\phi$. Этой же производной пропорционален и декремент затухания волны:

$$\gamma \sim - \left. \frac{\partial f}{\partial v_k} \right|_{v_k = v_\phi} \cdot$$

Эта важная формула показывает, что результирующий эффект резонансного взаимодействия существенно зависит от знака производной от функции распределения при $v_k = v_\phi$.

Если функция распределения $f(v_k)$ при $v_k = v_\phi$ убывает, то число быстрых частиц меньше числа медленных и колебания затухают (рис. 17.1). Если же распределение частиц таково, что при $v_k = v_\phi$ функция распределения возрастает, то число частиц, обгоняющих волну, превышает число частиц, отстающих от нее, и колебания плазмы не затухают, а нарастают.

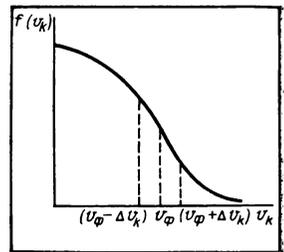
Первый случай имеет место, например, для равновесной плазмы с максвелловским распределением электронов, а второй осуществляется, например, при прохождении пучка заряженных частиц через плазму (см. следующий параграф).

Затухание, возникающее в первом случае, называют *затуханием Ландау*. Декремент затухания при максвелловском распределении электронов определяется формулой

$$\gamma_L(k) = \sqrt{\frac{\pi}{8}} \frac{\omega_e}{(kr_D)^3} \exp\left(-\frac{3}{2} + \frac{1}{2} k^2 r_D^2\right).$$

Формула эта справедлива, если $kr_D \ll 1$; при этом $\gamma_L \ll \omega_e$.

В случае ленгмюровских колебаний в



17.1. Распределение частиц по скоростям в случае затухающих колебаний

плазме с максвелловским распределением электронов при $v_\phi \gg v_e \sim \sqrt{T_e/m_e}$ резонансные электроны принадлежат к «хвосту» максвелловского распределения, где их число экспоненциально мало, поэтому экспоненциально мал и декремент затухания. При увеличении волнового вектора фазовая скорость ленгмюровских колебаний убывает, число резонансных электронов, эффективно взаимодействующих с волной, возрастает и поэтому увеличивается затухание Ландау. При $kr_D \sim 1$ фазовая скорость ленгмюровских колебаний сравнивается с тепловой скоростью электронов и число резонансных электронов оказывается настолько большим, что по порядку величины декремент затухания совпадает с частотой.

Подчеркнем, что рассматриваемая картина поглощения колебаний резонансными электронами относится к случаю достаточно слабых полей. Именно: необходимо, чтобы за время

$$\Delta t = (kv_k)^{-1} = [m_e / (ek^2 \varphi_0)]^{1/2},$$

за которое захваченная частица пролетает потенциальную яму, могло произойти затухание амплитуды поля, т. е. необходимо, чтобы выполнялось неравенство

$$\gamma \Delta t = \gamma \left(\frac{m_e}{ek^2 \varphi_0} \right)^{1/2} \gg 1.$$

Действительно, в противном случае резонансная частица успевает совершить несколько колебаний в потенциальной яме, прежде чем изменится амплитуда поля. Если при одном столкновении со «стенкой» ямы такая частица приобрела энергию от поля, то в последующем столкновении она уже отдаст энергию и т. д. В результате обмен энергией между резонансными частицами и полем сильно замедляется и, следовательно, затухание поля существенно уменьшается по сравнению с рассмотренным нами затуханием в слабом поле.

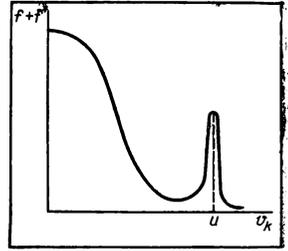
Рассмотрев затухание ленгмюровских колебаний, остановимся еще кратко на затухании ионного звука. Ограничимся при этом случаем длинных волн $kr_D \ll 1$. В этом случае декремент определяется формулой $\gamma_s \sim (\partial f / \partial v_k) / v_k = v_s$; подставляя сюда в качестве f максвелловскую функцию распределения, получим $\gamma_s \sim (m_e/m_i)^{1/2} kv_s$. Точная формула имеет вид

$$\gamma_s = \left(\frac{\pi m_e}{8 m_i} \right)^{1/2} kv_s.$$

Мы видим, что в случае ионного звука декремент затухания не содержит малого экспоненциального множителя; в отличие от случая ленгмюровских колебаний резонансные электроны уже не принадлежат к «хвосту» максвелловского распределения. Тем не менее условие слабого затухания $\gamma \ll \omega$ выполняется: $\gamma_s / (kv_s) \sim (m_e/m_i)^{1/2} \ll 1$.

17.5. Пучковая неустойчивость

Рассмотрим теперь второй случай, когда имеется интервал скоростей, в котором функция распределения электронов имеет положительную производную и в котором, следовательно, величина γ отрицательна. При $\gamma < 0$ амплитуда колебаний нарастает и мы можем говорить об их возбуждении. Величину $-\gamma = |\gamma|$ называют *инкрементом нарастания* колебаний. Такая ситуация возникает, как говорилось выше, при прохождении пучка электронов (или других заряженных частиц) через плазму. В простейшем случае электроны плазмы и электроны пучка имеют максвелловские распределения f и f' с разными температурами T и T' :



17.2. Пример немасвелловского распределения электронов

$$f = n_0 \left(\frac{m_e}{2\pi T} \right)^{3/2} \exp \left(-\frac{m_e v^2}{2T} \right), \quad f' = n'_0 \left(\frac{m_e}{2\pi T'} \right)^{3/2} \exp \left\{ \frac{-m_e (v - u)^2}{2T'} \right\},$$

где n_0 и n'_0 — плотности частиц в плазме и в пучке и u — направленная скорость электронов пучка. Суммарная функция распределения электронов $f + f'$ для этого случая представлена на рис. 17.2. Ясно, что в этом случае возможно возбуждение колебаний. Возбуждаться могут ленгмюровские колебания, а при $T_e \gg T_i$ также и ионно-звуковые колебания.

Будем считать, что плотность пучка мала ($n'_0 \ll n_0$), и предположим сперва, что электроны пучка имеют большой разброс по скоростям (горячий пучок). Ленгмюровские колебания возбуждаются, если $u > \omega_e / k_z$ (ось z — направлена вдоль u). Инкремент нарастания максимален, если $u - v'_e \gg v_e$ (здесь v_e, v'_e — средние тепловые скорости электронов плазмы и пучка); при этом

$$|\gamma| = \sqrt{\frac{\pi}{8}} \left(\frac{\omega'_e}{k v'_e} \right)^2 \omega_e,$$

где $\omega'_e = \sqrt{e^2 n'_0 / (\epsilon_0 m)}$. В этом случае нарастают практически все колебания, волновой вектор которых направлен в ту же сторону, что и u ($\cos \theta > 0$, где θ — угол между k и u), и удовлетворяет неравенству $kr_D \ll 1$. Более точно, нарастающими являются колебания, волновые векторы которых лежат внутри конуса $\theta < \theta_0$, где θ_0 близко к $\pi/2$. При уменьшении направленной скорости пучка угол раствора конуса θ_0 уменьшается, и при некотором критическом значении u (равном нескольким v_e) возбуждение колебаний прекращается.

Для возбуждения ионно-звуковых колебаний необходимо, чтобы скорость пучка превышала их фазовую скорость.

В рассмотренном случае горячего пучка каждая резонансная частица возбуждает колебания независимо, поэтому инкремент оказы-

вадается пропорциональным плотности частиц пучка. Иная ситуация возникает в случае холодного моноэнергетического пучка электронов, т. е. пучка с очень малым разбросом частиц по скоростям. В этом случае инкремент нарастания ленгмюровских колебаний оказывается пропорциональным $(n'_0)^{1/3}$, а не n'_0 (плотность пучка по-прежнему считается малой). Если при этом $k = \omega_e/u$, то

$$|\gamma| = \frac{\sqrt{3}}{2^{4/3}} \left(\frac{n'_0}{n_0} \right)^{1/3} \omega_e.$$

Формула эта справедлива, если $(kv'_e / \omega'_e)^2 \ll 1$.

Обратим внимание на то, что нарастание колебаний означает, по существу, неустойчивость системы плазма — пучок. Такую неустойчивость называют *пучковой*. Ясно, что экспоненциальное нарастание колебаний по закону $\exp(|\gamma|t)$ не может продолжаться бесконечно долго, так как амплитуда колебаний при $t \rightarrow \infty$ становится бесконечно большой. Это значит, что при достаточно больших значениях t нашими формулами нельзя пользоваться, так как они получены в предположении малости возмущения (малости поля и малости отклонений плотности частиц от равновесной), или, как говорят, в линейной теории. Поэтому, чтобы выяснить поведение системы плазма — пучок при $t \rightarrow \infty$, нужна существенно нелинейная теория, т. е. теория, учитывающая нелинейные процессы.

Простейшим из нелинейных процессов является обратное воздействие колебаний плазмы на неосциллирующую часть функций распределения частиц, которая уже не будет равновесной, максвелловской. Амплитуда колебаний предполагается еще малой, и поэтому взаимодействие волн друг с другом не учитывается. Такая теория, в которой учитывается воздействие колебаний на неосциллирующую часть функции распределения, но предполагается справедливым принцип суперпозиции колебаний, носит название *к в а з и л и н е й н о й т е о р и и*.

Так как амплитуда колебаний считается достаточной малой, то в квазилинейной теории фактически рассматривается влияние колебаний плазмы только на распределение резонансных частиц, т. е. частиц, проекция скорости которых ω на направление распространения близка к фазовой скорости волны $\omega \sim \omega/k = v_\phi$; именно эти частицы ответственны как за затухание, так и за раскачку колебаний, поскольку они сильно взаимодействуют с плазменными колебаниями. Нерезонансные частицы в среднем не обмениваются энергией с волнами, и поэтому можно считать, что их распределение не испытывает влияния со стороны колебаний плазмы. Неосциллирующая часть функции распределения \bar{f} определяется как среднее от функции распределения f по фазам колебаний: $\bar{f} = \langle f \rangle$. Эта функция не совпадает с равновесной функцией и зависит от времени, но изменяется со временем очень медленно. Проинтегрировав \bar{f} по проекциям скорости, перпендикулярным \mathbf{k} , мы получим распределение частиц $\bar{f}(\omega, t)$ по ω в момент времени t . Производная от этой функции по ω определяет

инкремент нарастания колебаний в момент времени t :

$$\gamma(t) \sim -(\partial \bar{f} / \partial \omega)_{\omega = v_{\phi}}.$$

Согласно квазилинейной теории, эта производная все время медленно уменьшается, т. е. функция \bar{f} при $\omega = v_{\phi}$ сглаживается. В результате на графике этой функции возникает плоский участок с $\partial \bar{f} / \partial \omega = 0$, или, как говорят, плато, обмен энергией между частицами и волнами прекращается и инкремент нарастания колебаний обращается в нуль.

Явление пучковой неустойчивости, имеет место также в двухтемпературной плазме с током. Ионы, имеющие большую массу, в этом случае можно считать покоящимися. Мы ограничимся рассмотрением длинноволновых ($kr_D \ll 1$) ионно-звуковых колебаний. Для таких колебаний, как мы видели в предыдущем параграфе, v_{ϕ} не зависит от волнового вектора и совпадает со скоростью ионного звука v_s . Если \mathbf{u} — средняя направленная скорость электронов, то $\partial f / \partial \omega > 0$ в области скоростей $0 < \omega < u$; поэтому если $u > v_s$, то ионно-звуковые колебания нарастают. Инкремент нарастания имеет вид

$$|\gamma| = \left(\frac{\pi m_e}{8 m_i}\right)^{1/2} k(u \cos \theta - v_s),$$

где θ — угол между \mathbf{k} и \mathbf{u} . Мы видим, что нарастают колебания, волновые векторы которых лежат внутри конуса с углом раствора θ , где $\cos \theta = v_s / u$.

17.6. Колебания плазмы в магнитном поле

Перейдем к рассмотрению колебаний плазмы в магнитном поле, которое будем считать постоянным и однородным. Плазма, как и в предыдущих параграфах, предполагается бесстолкновительной, т. е. частоты ее колебаний считаются значительно большими, чем частота соударений. Кроме того, плазма предполагается холодной, т. е. не учитывается тепловое движение частиц. В этом случае скорости частиц плазмы \mathbf{v}_{α} ($\alpha = e$ для электронов и $\alpha = i$ для ионов) как функции координат и времени удовлетворяют уравнениям

$$\frac{d_{\alpha} \mathbf{v}_{\alpha}}{dt} = \frac{e_{\alpha}}{m_{\alpha}} (\mathbf{E} + [\mathbf{v}_{\alpha} (\mathbf{B}_0 + \mathbf{B})]),$$

где e_{α} и m_{α} — заряд и масса частицы сорта α , \mathbf{B}_0 — индукция постоянного внешнего магнитного поля, \mathbf{E} и \mathbf{B} — напряженность электрического поля и магнитная индукция, возникающие в результате колебаний, и d_{α} / dt — полная производная по времени, учитывающая движение частиц сорта α (она определяется формулой $d_{\alpha} / dt = \partial / \partial t + \mathbf{v}_{\alpha} \nabla$; см. § 17.7). Кроме \mathbf{v}_{α} , нужно знать еще плотность частиц n_{α} , которая удовлетворяет уравнению непрерывности

$$\frac{\partial n_\alpha}{\partial t} + \operatorname{div}(n_\alpha \mathbf{v}_\alpha) = 0.$$

Плотность n_α можно представить в виде $n_\alpha = n_0 + n'_\alpha$, где n_0 — равновесная плотность частиц и n'_α — переменная часть плотности, связанная с колебаниями (мы считаем для простоты, что $e_i = -e_e$; в этом случае, очевидно, равновесная плотность одинакова для электронов и ионов).

Мы рассмотрим только малые (линейные) колебания и пренебрежем поэтому квадратичными величинами $[\mathbf{v}_\alpha \mathbf{B}]$, $(\mathbf{v}_\alpha \nabla) \mathbf{v}_\alpha$, $\mathbf{v}_\alpha n'_\alpha$, после чего линеаризованные уравнения движения примут вид

$$\frac{\partial \mathbf{v}_\alpha}{\partial t} = \frac{e_\alpha}{m_\alpha} \left(\mathbf{E} + [\mathbf{v}_\alpha \mathbf{B}_0] \right), \quad \frac{\partial n'_\alpha}{\partial t} + n_0 \operatorname{div} \mathbf{v}_\alpha = 0.$$

Решение этих уравнений будем искать в виде плоских волн: $\exp\{i(\mathbf{k}\mathbf{r} - \omega t)\}$. При этом уравнения переписуются в виде

$$i\omega \mathbf{v}_\alpha + \frac{e_\alpha}{m_\alpha} \mathbf{E} + \frac{e_\alpha}{m_\alpha} [\mathbf{v}_\alpha \mathbf{B}_0] = 0, \quad n'_\alpha = n_0 \frac{\mathbf{k}\mathbf{v}_\alpha}{\omega}.$$

Из первого из этих уравнений следует, что

$$v_{\alpha x} = \frac{e_\alpha (i\omega E_x - \omega_{B\alpha} E_y)}{m_\alpha (\omega^2 - \omega_{B\alpha}^2)}, \quad v_{\alpha y} = \frac{e_\alpha (i\omega E_y + \omega_{B\alpha} E_x)}{m_\alpha (\omega^2 - \omega_{B\alpha}^2)}, \quad v_{\alpha z} = \frac{ie_\alpha E_z}{m_\alpha \omega},$$

где $\omega_{B\alpha} = e_\alpha B_0/m_\alpha$ — циклотронная частота (гирочастота) частиц сорта α (ось z выбрана в направлении \mathbf{B}_0).

Вычислив \mathbf{v}_α и n'_α , можно определить плотности тока и заряда $\mathbf{j} = e_e n_0 (\mathbf{v}_e - \mathbf{v}_i)$, $\rho = e_e (n'_e - n'_i)$ и подставить полученные выражения в уравнения Максвелла:

$$\operatorname{rot} \mathbf{E} = -\partial \mathbf{B} / \partial t, \quad \operatorname{div} \mathbf{B} = 0, \\ \mu_0^{-1} \operatorname{rot} \mathbf{B} = \varepsilon_0 \partial \mathbf{E} / \partial t + \mathbf{j}, \quad \varepsilon_0 \operatorname{div} \mathbf{E} = \rho.$$

Для плоских монохроматических волн эти уравнения примут вид

$$[\mathbf{kE}] = \omega \mathbf{B}, \quad \mu_0^{-1} [\mathbf{kB}]_i = -\varepsilon_0 \omega \varepsilon_{ij} E_j.$$

Здесь подразумевается суммирование по индексу $j = 1, 2, 3$ и ε_{ij} означает матрицу

$$\varepsilon_{ij} = \begin{pmatrix} \varepsilon_1 & i\varepsilon_2 & 0 \\ -i\varepsilon_2 & \varepsilon_1 & 0 \\ 0 & 0 & \varepsilon_3 \end{pmatrix}$$

с компонентами

$$\varepsilon_1 = 1 - \sum_\alpha \frac{\omega_{p\alpha}^2}{\omega^2 - \omega_{B\alpha}^2}, \quad \varepsilon_2 = - \sum_\alpha \frac{\omega_{p\alpha}^2 \omega_{B\alpha}}{\omega (\omega^2 - \omega_{B\alpha}^2)},$$

$$\epsilon_3 = 1 - \sum_{\alpha} \frac{\omega_{p\alpha}^2}{\omega^2},$$

где $\omega_{p\alpha}^2 = e^2 n_{\alpha} / (\epsilon_0 m_{\alpha})$. Матрица ϵ_{ij} представляет собой *тензор диэлектрической проницаемости* плазмы в магнитном поле (плазму в этом случае называют также *магнитоактивной плазмой*). При $\mathbf{V}_0 = 0$ мы получим $\epsilon_{ij} = \delta_{ij} \epsilon_3$ и придем (пренебрегая $1/m_i$ по сравнению с $1/m_e$) к формулам § 17.2.

Мы должны теперь исключить \mathbf{V} из уравнений Максвелла и записать условие разрешимости полученной однородной системы уравнений для компонент \mathbf{E} . Мы получим тогда

$$\Delta \equiv \det \Lambda_{ij} = 0, \quad \Lambda_{ij} = N^2 \left(\frac{k_i k_j}{k^2} - \delta_{ij} \right) + \epsilon_{ij},$$

где $N = ck/\omega$ — показатель преломления. Подставив ϵ_{ij} , найдем

$$AN^4 + BN^2 + C = 0;$$

$$A = \epsilon_1 \sin^2 \theta + \epsilon_3 \cos^2 \theta; \quad B = -\epsilon_1 \epsilon_3 (1 + \cos^2 \theta) - (\epsilon_1^2 - \epsilon_2^2) \sin^2 \theta;$$

$$C = \epsilon_3 (\epsilon_1^2 - \epsilon_2^2),$$

где θ — угол между \mathbf{k} и \mathbf{V}_0 . Полученное соотношение представляет собой дисперсионное уравнение, связывающее частоты возможных колебаний магнитоактивной плазмы с волновым вектором. Задав \mathbf{k} , получим пять частот $\omega_{\nu}(k)$, $\nu = 1, 2, 3, 4, 5$, т. е. в холодной магнитоактивной плазме (без учета теплового движения частиц) существует пять ветвей высокочастотных колебаний.

В явном виде зависимости $\omega_{\nu}(k)$ удается получить в некоторых предельных случаях. Мы остановимся здесь только на продольных волнах, когда поле \mathbf{E} параллельно \mathbf{k} . В этом случае дисперсионное уравнение приобретает вид $A = 0$.

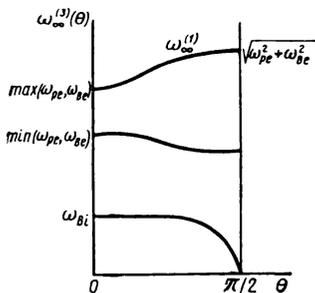
При $\mathbf{V}_0 = 0$ продольными колебаниями являются ленгмюровские, или плазменные, колебания. Поэтому и при $\mathbf{V}_0 \neq 0$ о колебаниях с частотами, определяемыми из условия $A = 0$, говорят как о продольных плазменных колебаниях в магнитном поле, а сами частоты называют *плазменными* или *гибридными резонансами* (так как эти частоты являются комбинациями плазменной и циклотронной частот).

Дисперсионное уравнение $A = 0$ для этих резонансов имеет вид

$$1 - \frac{\omega_{pe}^2}{\omega^2} \cos^2 \theta - \frac{\omega_{pe}^2}{\omega^2 - \omega_{Be}^2} \sin^2 \theta - \frac{\omega_{pi}^2}{\omega^2} \cos^2 \theta - \frac{\omega_{pi}^2}{\omega^2 - \omega_{Bi}^2} \sin^2 \theta = 0.$$

Это уравнение третьей степени относительно ω^2 определяет три резонансные частоты $\omega_{\infty}^{(j)}(\theta)$ ($j = 1, 2, 3$) (индекс ∞ означает, что эти частоты являются собственными при $k \rightarrow \infty$).

Если пренебречь вкладом ионов (из-за малости отношения m_e/m_i), то частоты $\omega_{\infty}^{(1,2)}$ определяются формулами



17.3. Зависимость частот продольных колебаний плазмы в магнитном поле от угла θ

$$\omega_{\infty}^{(1,2)}(\theta) = \left[\frac{1}{2} (\omega_{pe}^2 + \omega_{Be}^2) \pm \frac{1}{2} \sqrt{(\omega_{pe}^2 + \omega_{Be}^2)^2 - 4\omega_{pe}^2 \omega_{Be}^2 \cos^2 \theta} \right]^{1/2}$$

(знаки «+» и «-» соответствуют $j = 1$ и $j = 2$). Третья частота $\omega_{\infty}^{(3)}(\theta)$ зависит от вклада ионов. Если $\omega = \omega_{Bi}$, то

$$\omega_{\infty}^{(3)}(\theta) = \omega_{Bi} \left(1 - \frac{m_e}{2m_i} \operatorname{tg}^2 \theta \right).$$

На рис. 17.3 изображена зависимость гибридных резонансов от угла θ ($\max(a, b)$ и $\min(a, b)$ означают большую и меньшую из величин a, b).

17.7. Магнитогидродинамическое течение

Рассмотрим теперь медленные движения плазмы. Если характерные времена таких движений (например, периоды колебаний) велики по сравнению с временем свободного пробега частиц, то плазма ничем не отличается от обычной жидкости или газа и, подобно им, описывается гидродинамическими уравнениями. Специфика плазмы (впрочем, роднящая ее с такими жидкостями, как ртуть) проявляется во внешнем магнитном поле и связана с тем, что плазма — хороший проводник электричества. Поэтому далее вместо термина *плазма* для общности мы будем пользоваться термином *проводящая жидкость*.

Напомним, что если p — давление жидкости, то на ее единичный объем со стороны окружающей жидкости действует сила $-\nabla p$. При наличии внешнего магнитного поля к этой силе нужно еще прибавить магнитную силу $[jB]$. Масса единичного объема жидкости есть, очевидно, ее плотность ρ , поэтому

$$\rho \frac{dv}{dt} = -\nabla p + [jB]$$

(обратим внимание на то, что в предшествующей части этой главы под ρ понимается плотность заряда и только в последних двух параграфах — массовая плотность). Здесь \mathbf{v} — скорость определенной частицы жидкости. Но течение жидкости удобнее описывать, не рассматривая ее отдельных частиц, а вводя поле скоростей $\mathbf{v} = \mathbf{v}(\mathbf{r}, t)$. При этом $\mathbf{v}(\mathbf{r}, t)$ — скорость не какой-либо определенной частицы жидкости, а любой ее частицы, проходящей пространственную точку \mathbf{r} в момент времени t . В таком методе описания движения (его называют методом Эйлера в отличие от следящего за отдельной частицей описания, называемого методом Лагранжа) ускорение $d\mathbf{v}/dt$ равно, очевидно, $\partial\mathbf{v}/\partial t + (\mathbf{v}\nabla)\mathbf{v}$. Действительно, рассмотрим сперва стационарное поле скоростей, когда $\partial\mathbf{v}/\partial t = 0$. Тогда

если в точке \mathbf{r} скорость равна \mathbf{v} и мы следим за определенной частицей, то следует учесть, что за время dt эта частица очутится в точке $\mathbf{r} + d\mathbf{r} = \mathbf{r} + \mathbf{v}dt$, в которой поле скоростей составляет $\mathbf{v}(\mathbf{r} + d\mathbf{r}) = \mathbf{v}(\mathbf{r}) + (d\mathbf{r}\nabla)\mathbf{v}$; поэтому изменение скорости определенной частицы равно $(d\mathbf{r}\nabla)\mathbf{v}$; разделив эту величину на dt , получим $(\mathbf{v}\nabla)\mathbf{v}$. Если поле не стационарно, то к этому выражению следует прибавить $\partial\mathbf{v}/\partial t$. Таким образом,

$$\frac{\partial\mathbf{v}}{\partial t} + (\mathbf{v}\nabla)\mathbf{v} = -\frac{1}{\rho}\nabla p + \frac{1}{\rho}[\mathbf{j}\mathbf{B}].$$

Входящие в гидродинамическое уравнение движения величины \mathbf{v} и ρ должны, очевидно, удовлетворять уравнению непрерывности

$$\frac{\partial\rho}{\partial t} + \text{div}(\rho\mathbf{v}) = 0,$$

выражающему закон сохранения массы жидкости.

Обратимся теперь к величинам \mathbf{B} , \mathbf{H} , \mathbf{E} , определяющим электромагнитное поле в жидкости. Они должны удовлетворять уравнениям Максвелла

$$\text{rot}\mathbf{E} = -\partial\mathbf{B}/\partial t, \quad \text{div}\mathbf{B} = 0, \quad \text{rot}\mathbf{H} = \mathbf{j}.$$

Под \mathbf{j} здесь можно, пренебрегая током смещения, понимать ток проводимости

$$\mathbf{j} = \sigma\mathbf{E}^* = \sigma(\mathbf{E} + [\mathbf{v}\mathbf{B}]),$$

где σ — удельная электрическая проводимость и \mathbf{E}^* — напряженность электрического поля в системе отсчета, связанной с движущимся элементом жидкости. Предполагая проводимость среды очень большой, можно считать, что поле $\mathbf{E}^* = \mathbf{j}/\sigma$ равно нулю, поэтому

$$\mathbf{E} = -[\mathbf{v}\mathbf{B}].$$

Первое уравнение Максвелла принимает теперь вид

$$\partial\mathbf{B}/\partial t = \text{rot}[\mathbf{v}\mathbf{B}].$$

Подставив далее в гидродинамическое уравнение движения $\mathbf{j} = \text{rot}\mathbf{H}$, получим

$$\frac{\partial\mathbf{v}}{\partial t} + (\mathbf{v}\Delta)\mathbf{v} = -\frac{1}{\rho}\nabla p - \frac{1}{\rho}[\mathbf{B}\text{rot}\mathbf{H}].$$

Это уравнение вместе с записанным выше первым уравнением Максвелла и определяет скорость течения и магнитное поле в среде.

Из первого уравнения следует, что если взять произвольную жидкую поверхность, т. е. поверхность, образованную частицами среды, то магнитный поток через нее не меняется при движении частиц. Действительно, обозначив такую поверхность Σ , получим

$$\frac{\partial}{\partial t} \int_{\Sigma} \mathbf{B}ds = \int_{\Sigma} \text{rot}[\mathbf{v}\mathbf{B}] ds = \int_L [\mathbf{v}\mathbf{B}] dl,$$

где L — контур, охватывающий Σ (последнее равенство написано на основании формулы Стокса). Но $d\mathbf{l} \parallel \mathbf{v}$, поэтому последнее выражение равно нулю и

$$\Phi \equiv \int_{\Sigma} \mathbf{B} ds = \text{const.}$$

Рассмотрим далее две жидкие поверхности, на которых в некоторый момент времени $B_n = 0$ (\mathbf{n} — направление нормали к поверхности). Это равенство сохранится тогда и во все последующие моменты времени. Линия пересечения этих поверхностей является, очевидно, магнитной силовой линией, а так как линия пересечения жидких поверхностей представляет собой жидкую линию, то мы приходим к выводу, что магнитные силовые линии скреплены с жидкими линиями. На этом основании говорят о *вмороженности магнитных силовых линий*.

Если взять элемент достаточно тонкой жидкой силовой трубки, то с течением времени его длина и площадь сечения Δs будут меняться, но магнитный поток останется неизменным, т. е.

$$B\Delta s = B_0\Delta s_0$$

(индексом 0 обозначены начальные значения). С другой стороны, масса жидкости в элементе трубки также не будет меняться, т. е.

$$\rho\Delta s\Delta l = \rho_0\Delta s_0\Delta l_0$$

(Δl , Δl_0 , ρ , ρ_0 — длины элемента трубки и значения плотности жидкости в рассматриваемый и начальный моменты времени). Разделив первое из написанных равенств на второе, получим

$$\frac{B}{\rho\Delta l} = \frac{B_0}{\rho_0\Delta l_0} = \text{const.}$$

Таким образом, величина B/ρ изменяется пропорционально длине элемента магнитной силовой линии. Если жидкость несжимаема, то магнитная индукция изменяется пропорционально «растяжению» силовых линий.

17.8. Магнитогидродинамические волны

Как известно, возмущение плотности в каком-либо месте жидкости (или газа) не локализуется в этом месте, а распространяется в виде волны — продольной звуковой волны, квадрат скорости v_s которой равен адиабатной производной давления по плотности: $v_s^2 = (dp/d\rho)_s$.

В проводящей жидкости (или газе), находящейся в магнитном поле, происходит модификация звуковых волн — появляются две ветви магнитозвуковых волн с разными скоростями, зависящими от магнитного поля. Кроме того, возникает еще третья ветвь колебаний — поперечная, не связанная с изменением плотности среды. Ее называют *альве-*

новской (в честь открывшего ее Альвена); скорость ее определяется магнитным полем (и плотностью среды). Все три указанные волны называют *магнитогидродинамическими*.

Чтобы понять происхождение этих волн и найти их скорости, обратимся к уравнениям, описывающим магнитогидродинамическое течение, и положим в них $\rho = \rho_0 + \rho'$, $p = p_0 + p'$, $\mathbf{V} = \mathbf{V}_0 + \mu_0 \mathbf{h}$, где ρ_0 , p_0 , \mathbf{V}_0 — постоянные и не зависящие от координат значения величин, соответствующие равновесному состоянию среды, и ρ' , p' , \mathbf{h} — малые добавки к ним, описывающие возмущение. Скорость возникающего течения \mathbf{v} также мала, поэтому, линеаризуя уравнения, имеем

$$\rho_0 \frac{d\mathbf{v}}{dt} = -\nabla p' + [\text{rot } \mathbf{h} \mathbf{B}_0],$$

$$\rho_0 \text{div } \mathbf{v} + \frac{\partial \rho'}{\partial t} = 0, \quad \frac{\partial \mathbf{h}}{\partial t} = \text{rot } [\mathbf{v} \mathbf{B}_0].$$

Для получения замкнутой системы уравнений нам остается найти p' . Давление жидкости является функцией плотности и температуры. В звуковой волне происходят быстрые изменения плотности (в том смысле, что теплообмен между различными элементами жидкости не успевает произойти); иными словами, возмущение происходит адиабатно; поэтому $p' = (dp/d\rho)_s \rho'$, где $(dp/d\rho)_s = v_s^2$ (индекс s означает, что производная берется при постоянной энтропии; заметим, что в этом параграфе v_s — скорость обычного, а не ионного звука). Учитывая эту связь, мы имеем семь линейных уравнений для определения семи неизвестных: ρ' , \mathbf{v} , \mathbf{h} .

Заметим, что эти же уравнения можно использовать для описания не только плазмы со столкновениями, но и двухтемпературной бесстолкновительной плазмы. В этом случае вместо обычной скорости звука в уравнения войдет (как в разделе о ионном звуке) скорость изотермического (не адиабатного) ионного звука.

Будем искать решения наших уравнений в виде плоских монохроматических волн. Тогда

$$\omega \rho_0 \mathbf{v} = v_s^2 \mathbf{k} \rho' + [\mathbf{B}_0 [\mathbf{k} \mathbf{h}]];$$

$$-\rho_0 \mathbf{k} \mathbf{v} + \omega \rho' = 0; \quad -\omega \mathbf{h} = -\mu_0^{-1} [\mathbf{k} [\mathbf{v} \mathbf{B}_0]].$$

Используя формулу $[\mathbf{a} [\mathbf{bc}]] = \mathbf{b}(\mathbf{ac}) - \mathbf{c}(\mathbf{ab})$, перепишем эти уравнения в виде

$$\omega \rho_0 \mathbf{v} = v_s^2 \mathbf{k} \rho' + \frac{1}{\omega \mu_0} \{ \mathbf{k} B_0^2 (\mathbf{k} \mathbf{v}) - \mathbf{k} (\mathbf{B}_0 \mathbf{v}) (\mathbf{k} \mathbf{B}_0) +$$

$$+ \mathbf{v} (\mathbf{k} \mathbf{B}_0)^2 - \mathbf{B}_0 (\mathbf{k} \mathbf{v}) (\mathbf{k} \mathbf{B}_0) \},$$

$$\rho_0 \mathbf{k} \mathbf{v} = \omega \rho', \quad \mu_0 \omega \mathbf{h} = -\mathbf{v} (\mathbf{k} \mathbf{B}_0) + \mathbf{B}_0 (\mathbf{k} \mathbf{v}).$$

Согласно последнему уравнению, $\mathbf{k} \mathbf{h} = 0$, т. е. колебание магнитного поля является поперечным.

Умножая первое уравнение скалярно сперва на \mathbf{B}_0 , а затем на \mathbf{k} и используя второе уравнение, получим

$$\omega^2 (\mathbf{B}_0 \mathbf{v}) = v_s^2 (\mathbf{kB}_0) (\mathbf{k}\mathbf{v}),$$

$$\omega \rho_0 (\mathbf{k}\mathbf{v}) = v_s^2 k^2 \rho_0 \frac{(\mathbf{k}\mathbf{v})}{\omega} + \frac{(\mathbf{k}\mathbf{v})}{\mu_0 \omega} \left\{ k^2 B_0^2 - \frac{k^2 v_s^2}{\omega^2} (\mathbf{kB}_0)^2 \right\}.$$

Из второго уравнения следует, что если $B_0 = 0$, то $\omega = kv_s$, т. е. возникает, как и следовало ожидать, лишь обычная звуковая волна.

При $\mathbf{B}_0 \neq 0$ следует различать две возможности: $\rho' \neq 0$ и $\rho' = 0$. В первом случае, когда возникает волна плотности, $\mathbf{k}\mathbf{v} \neq 0$, поэтому второе из написанных соотношений можно сократить на $\mathbf{k}\mathbf{v}$. В результате мы получим

$$\omega^2 = v_s^2 k^2 + \frac{1}{\mu_0 \rho_0} \left\{ B_0^2 k^2 - v_s^2 \frac{k^2}{\omega^2} (\mathbf{kB}_0)^2 \right\},$$

откуда

$$v^4 - v_s^2 v^2 - v_A^2 (v^2 - v_s^2 \cos^2 \theta) = 0,$$

где $v = \omega/k$, θ — угол между \mathbf{k} и \mathbf{B}_0 и $v_A^2 = B_0^2/(\mu_0 \rho_0)$. Смысл величины v_A мы объясним ниже.

Полученное квадратное уравнение для v^2 имеет два корня:

$$v_{\pm}^2 = \frac{1}{2} \left\{ v_s^2 + v_A^2 \pm \sqrt{(v_s^2 + v_A^2)^2 - 4v_s^2 v_A^2 \cos^2 \theta} \right\}.$$

Волну с фазовой скоростью v_+ называют *быстрой магнитозвуковой*, а со скоростью v_- — *медленной магнитозвуковой*. При $\mathbf{B}_0 = 0$ быстрая магнитозвуковая волна переходит в обычную звуковую, а медленная исчезает.

Рассмотрим теперь случай $\rho' = 0$. Согласно уравнению непрерывности, при этом $\mathbf{k}\mathbf{v} = 0$; из соотношения $\omega^2 (\mathbf{B}_0 \mathbf{v}) = v_s^2 (\mathbf{kB}_0) (\mathbf{k}\mathbf{v})$ следует, что $(\mathbf{B}_0 \mathbf{v}) = 0$. Иными словами, при $\rho' = 0$ скорость жидкости перпендикулярна как волновому вектору \mathbf{k} , так и вектору \mathbf{B}_0 . Уравнения для \mathbf{v} , \mathbf{h} принимают вид

$$\omega \rho_0 \mathbf{v} = \frac{1}{\mu_0 \omega} (\mathbf{kB}_0)^2 \mathbf{v}; \quad \mu_0 \omega \mathbf{h} = -\mathbf{v} (\mathbf{kB}_0).$$

Полученные уравнения имеют нетривиальные решения, если

$$\omega = kv_A \cos \theta, \quad v_A = B_0 / \sqrt{\mu_0 \rho_0}.$$

Эту ветвь магнитогидродинамических колебаний называют *альвеновской волной*, а величину v_A — *альвеновской скоростью*. Заметим, что в альвеновской волне $\mathbf{h} = -\mathbf{v} / \sqrt{\mu_0 \rho_0}$, поэтому $1/2 \mu_0 \mathbf{h}^2 = 1/2 \rho_0 v^2$; иными словами, *плотность магнитной энергии совпадает с плотностью кинетической энергии жидкости*.

Уравнение Власова $\frac{\partial f}{\partial t} + \mathbf{v} \frac{\partial f}{\partial \mathbf{r}} + \frac{\mathbf{F}}{m_e} \frac{\partial f}{\partial \mathbf{v}} = 0$

Диэлектрическая проницаемость плазмы (высокочастотная) $\epsilon(\omega, \mathbf{k}) = 1 - \frac{\omega_e^2}{\omega^2} \left(1 + \frac{3T k^2}{m_e \omega^2} \right),$

$$\omega_e = \sqrt{e^2 n_0 / (\epsilon_0 m_e)}$$

Закон дисперсии плазменных колебаний $\omega = \omega_e \left(1 + \frac{3}{2} r_D^2 k^2 \right),$

$$r_D = \sqrt{\frac{\epsilon_0 T}{e^2 n_0}}$$

Декремент затухания Ландау $\gamma_L = \sqrt{\frac{\pi}{8}} \frac{\omega_e}{(k r_D)^3} \exp \left\{ -\frac{3}{2} + \frac{1}{2 k^2 r_D^2} \right\}$

Частота ионного звука в плазме $\omega_s(k) = \frac{k v_s}{\sqrt{1 + k^2 r_D^2}}, \quad v_s = \sqrt{\frac{T_e}{m_i}}$

Гибридные резонансы $\omega^{(1,2)}(\theta) = \left[\frac{1}{2} \left(\omega_{pe}^2 + \omega_{Be}^2 \right) \pm \frac{1}{2} \sqrt{\left(\omega_{pe}^2 + \omega_{Be}^2 \right)^2 - 4 \omega_{pe}^2 \omega_{Be}^2 \cos^2 \theta} \right]^{1/2}$

Частота альвеновской волны $\omega_A = k v_A \cos \theta, \quad v_A = \frac{B_0}{\sqrt{\mu_0 \rho_0}}$

Скорости быстрой и медленной магнитозвуковых волн $v_{\pm}^2 = \frac{1}{2} \left\{ v_s^2 + v_A^2 \pm \sqrt{\left(v_s^2 + v_A^2 \right)^2 - 4 v_s^2 v_A^2 \cos^2 \theta} \right\}$



Глава 18. КРИСТАЛЛООПТИКА

18.1. Тензор диэлектрической проницаемости

Рассматривая электромагнитные волны в материальных средах, мы пока считали, что среда изотропна, т. е. что ее свойства одинаковы по всем направлениям. Такими средами являются жидкости, газы, плазма (в отсутствие внешних полей), но не кристаллические твердые тела —

кристаллы ведут себя в разных направлениях по-разному, или, как говорят, обладают анизотропией.

Переходя к изучению электромагнитных волн в анизотропных средах, мы по-прежнему будем считать линейной связь между индукциями и напряженностями полей. Но в анизотропном случае индукции уже не обязательно параллельны соответствующим полям. Это значит, что

диэлектрическая и магнитная проницаемости анизотропных сред являются не скалярными, а тензорными величинами.

При этом они, как и в изотропном случае, зависят, вообще говоря, и от частоты и от волнового вектора (временная и пространственная дисперсии). Мы приходим к выводу, что свойства анизотропной среды по отношению к электромагнитным волнам характеризуются двумя тензорами (т. е. восемнадцатью обычными, скалярными) функциями $\epsilon_{ij} \equiv \epsilon_{ij}(\omega, \mathbf{k})$, $\mu_{ij} \equiv \mu_{ij}(\omega, \mathbf{k})$ ($i, j = 1, 2, 3$ или $i, j = x, y, z$), с помощью которых индукции выражаются через поля

$$D_i = \epsilon_0 \epsilon_{ij} E_j, \quad B_i = \mu_0 \mu_{ij} H_j$$

(напомним, что при такой записи по повторяющимся индексам подразумевается суммирование; тензор с двумя индексами называют *тензором второго ранга*).

В этой главе мы рассмотрим вещества, анизотропные по отношению к электрическим, но не магнитным свойствам, именно к таким средам принадлежит большинство кристаллов. Иными словами, мы будем считать, что $\mathbf{B} = \mu_0 \mu \mathbf{H}$, где μ — обычная (скалярная) функция ω и \mathbf{k} , т. е. $\mu_{ij} = \mu \delta_{ij}$.

Анизотропия электрических (как и магнитных) свойств возможна далеко не у всех кристаллов. В самом деле, если кристалл обладает кубической симметрией, то для него нельзя построить тензор второго ранга, не пропорциональный единичному тензору δ_{ij} . Поэтому кубический кристалл по электромагнитным свойствам ничем не отличается от жидкости или газа, т. е. для него $\epsilon_{ij} = \epsilon \delta_{ij}$.

Плотность потока энергии поля в анизотропной среде определяется той же формулой, что и в изотропной среде: $\mathbf{S} = [\mathbf{E}\mathbf{H}]$. Для плотности энергии W и ее изменения за 1 с, т. е. для выделяемой теплоты Q , в случае анизотропной среды справедливы формулы

$$W = \frac{\epsilon_0}{4} \frac{d}{d\omega} (\omega \epsilon_{ij}) E_i^* E_j, \quad Q = \frac{i \epsilon_0 \omega}{2} (\epsilon_{ij}^* - \epsilon_{ji}) E_i E_j^*$$

аналогичные формулам для изотропных сред (см. § 16.6; мы опускаем для простоты магнитную энергию). Формула для W справедлива для прозрачной или почти прозрачной среды.

Как мы видели, для прозрачной изотропной среды проницаемости вещественны. В анизотропных средах условие прозрачности имеет более сложный вид $\epsilon_{ji} = \epsilon_{ij}^*$ и не означает уже вещественности всех компонент тензора проницаемости. В самом деле, запишем ϵ_{ij} в виде $\epsilon_{ij} = \epsilon'_{ij} + i \epsilon''_{ij}$, где ϵ' , ϵ'' — два вещественных тензора. Тогда для

прозрачной среды

$$\varepsilon'_{jt} = \varepsilon'_{ij}, \quad \varepsilon''_{jt} = -\varepsilon''_{ij}.$$

Вектор $\varepsilon'_{ij}E_j$ может быть представлен в виде векторного произведения $\varepsilon''_{ij}E_j = [\mathbf{Eg}]_i$, где \mathbf{g} — некоторый (аксиальный) вектор $g_i = = 1/2 \varepsilon_{ijl}\varepsilon''_{jl}(\varepsilon_{ijl}$ — единичный антисимметричный тензор). Этот вектор называют *вектором гирации*, а среды, для которых он отличен от нуля, — *гиротропными*. В гиروتропной среде, очевидно,

$$\mathbf{D} = \mathbf{D}' + i\varepsilon_0[\mathbf{Eg}], \quad D'_i = \varepsilon_0\varepsilon'_{ij}E_j.$$

Большинство кристаллов в отсутствие магнитного поля не обладает гиروتропией. В этой главе мы рассмотрим только такие кристаллы, посвятив гиروتропным средам следующую главу. Поэтому, ограничиваясь прозрачными кристаллами, будем считать тензор вещественным и симметричным.

Для любого симметричного тензора второго ранга могут быть выбраны три взаимно перпендикулярных направления таким образом, что если направить вдоль них координатные оси x, y, z , то тензор примет диагональную форму

$$\varepsilon_{ij} = \begin{pmatrix} \varepsilon_x & & \\ & \varepsilon_y & \\ & & \varepsilon_z \end{pmatrix}$$

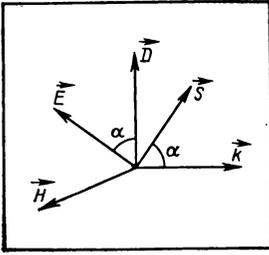
(невывисанные элементы — нули). В случае тензора ε_{ij} такие направления называют *главными диэлектрическими осями*, а величины $\varepsilon_x, \varepsilon_y, \varepsilon_z$ — *главными диэлектрическими проницаемостями* кристалла. Для кубического кристалла все три главные проницаемости совпадают, а выбор осей произволен. Для одноосного кристалла совпадают две главные проницаемости, например $\varepsilon_x = \varepsilon_y$, и выбор соответствующих осей ограничен лишь тем, чтобы они лежали в базисной плоскости (плоскость xy). В двухосном кристалле все три главные проницаемости различны и отсутствует произвол в выборе главных осей. Очевидно,

главные оси, и только они, обладают тем замечательным свойством, что если поле направлено вдоль одной из них, то индукция, как в изотропном случае, параллельна полю.

Заметим, что при изменении частоты могут меняться не только главные диэлектрические проницаемости, но и сами направления главных осей; в этом случае говорят о *дисперсии осей*. Дисперсия осей возможна в двухосных кристаллах определенных систем.

18.2. Фазовая скорость и лучевой вектор

Рассмотрим плоскую монохроматическую волну в прозрачной анизотропной среде. Если ω — частота и \mathbf{k} — волновой вектор, то уравнения Максвелла можно, очевидно, записать в форме



18.1.
Взаимное расположение векторов E, D, H, k, S

$$\mathbf{H} = \frac{1}{\mu_0 \omega} [kE], \quad \mathbf{D} = -\frac{1}{\omega} [kH], \quad kD = kH = 0.$$

Согласно этим уравнениям, три вектора k, D и H взаимно перпендикулярны. Далее, вектору H перпендикулярны E и вектор Пойнтинга $S = [EH]$. Но раз четыре вектора k, D, E, S перпендикулярны H , то они лежат в одной плоскости. В этой плоскости $E \perp S$, а $D \perp k$, поэтому угол между E и D равен углу между S и k (угол α на рис. 18.1, иллюстрирующем взаимное расположение пяти рассматриваемых векторов). Представим вектор k в виде

$$k = (\omega/c)N.$$

Это соотношение напоминает скалярное соотношение $k = (\omega/c)N$ для изотропного случая, поэтому о величине N принято говорить как о *векторном показателе преломления* (хотя теперь ни его модуль, ни компоненты не имеют в отличие от изотропных сред такого простого отношения к закону преломления). Фазовая скорость волны v_ϕ , т. е. скорость перемещения поверхности, на которой постоянна фаза $\phi = \mathbf{kr} - \omega t$, направлена вдоль k и равна по модулю ω/k . Используя вектор N , ее можно представить в виде $v_\phi = cN/N^2$.

Найдем теперь связь между частотой волны и волновым вектором, или, что то же самое, определим вектор N . Подставим с этой целью выражение для H в уравнение для D . Раскрыв двойное векторное произведение, получим

$$\epsilon_0^{-1} D = N^2 E - N(NE).$$

С другой стороны, $\epsilon_0^{-1} D_i = \epsilon_{ij} E_j$. Приравняв друг другу оба выражения для D , получим систему уравнений

$$(N^2 \delta_{ij} - N_i N_j - \epsilon_{ij}) E_j = 0.$$

Эта система имеет нетривиальные решения, если равен нулю ее определитель. Таким образом, уравнение для определения N имеет вид

$$\det(N^2 \delta_{ij} - N_i N_j - \epsilon_{ij}) = 0.$$

Этот определитель просто раскрыть в системе координат, связанной с главными диэлектрическими осями:

$$N^2 (\epsilon_x N_x^2 + \epsilon_y N_y^2 + \epsilon_z N_z^2) - [N_x^2 \epsilon_x (\epsilon_y + \epsilon_z) + N_y^2 \epsilon_y (\epsilon_x + \epsilon_z) + N_z^2 \epsilon_z (\epsilon_x + \epsilon_y)] + \epsilon_x \epsilon_y \epsilon_z = 0.$$

Полученное соотношение, называемое *уравнением Френеля*, является одним из основных уравнений кристаллооптики. Оно определяет ω как неявную функцию k . Решение этой задачи, однако, очень слож-

но ввиду дисперсии главных диэлектрических проницаемостей, а в случае двухосных кристаллов определенных систем — еще и ввиду возможной дисперсии осей. Более простой является задача: найти волновое число при заданной частоте и заданном направлении распространения волны. В такой постановке уравнение Френеля — квадратное уравнение относительно N^2 , а следовательно, и относительно $k^2 = (\frac{\omega}{c} N)^2$. Оно имеет два, вообще говоря, различных корня, которые, как мы сейчас убедимся, соответствуют двум различным плоским поляризациям волны. Действительно, рассмотрим вектор \mathbf{D} , который перпендикулярен \mathbf{N} . Так как направление \mathbf{N} задано, то нас интересуют две перпендикулярные \mathbf{N} проекции \mathbf{D} , которые мы обозначим D_x и D_y (ось z выбрана вдоль \mathbf{N}). Из уравнения $\epsilon_0^{-1} \mathbf{D} = N^2 \mathbf{E} - \mathbf{N}(\mathbf{E}\mathbf{N})$ следует, что они связаны с \mathbf{E} соотношением $\epsilon_0^{-1} \mathbf{D}_i = N^2 E_i$ ($i = x, y$), а так как $\epsilon_0 E_i = \eta_{ij} D_j$, где $\eta_{ij} = (\epsilon^{-1})_{ij}$ — тензор, обратный ϵ_{ij} (т. е. $\eta_{ik} \epsilon_{kj} = \delta_{ij}$), то

$$D_i = N^2 \eta_{ij} D_j.$$

Для того чтобы это уравнение имело нетривиальные решения, должно выполняться условие

$$\det(\delta_{ij}/N^2 - \eta_{ij}) = 0 \quad (i, j = x, y).$$

Это соотношение, представляющее собой квадратное относительно N^2 уравнение, имеет, вообще говоря, два различных корня $N_{1,2}^2$. Им соответствуют два различных вектора электрической индукции $\mathbf{D}^{(1)}$ и $\mathbf{D}^{(2)}$, причем эти векторы, как известно из задачи о нахождении собственных значений и собственных векторов матрицы, взаимно ортогональны: $\mathbf{D}^{(1)} \mathbf{D}^{(2)} = 0$. Мы приходим к важному выводу:

в анизотропной среде плоская монохроматическая волна может быть поляризованной только линейно.

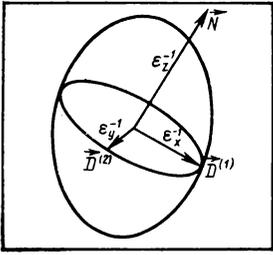
Напомним, что

в изотропной среде плоская монохроматическая волна в общем случае поляризована эллиптически.

Уравнение для вектора \mathbf{D} допускает простую геометрическую интерпретацию. Направим координатные оси вдоль главных диэлектрических осей; в такой системе координат, очевидно, тензор η диагонален и имеет вид

$$\eta = \begin{pmatrix} 1/\epsilon_x & & \\ & 1/\epsilon_y & \\ & & 1/\epsilon_z \end{pmatrix}$$

(остальные его компоненты равны нулю). Построим далее эллипсоид с осями, направленными вдоль главных диэлектрических осей и равными $1/\epsilon_x$, $1/\epsilon_y$, $1/\epsilon_z$ (рис. 18.2). Пересечем этот эллипсоид плоскостью, проходящей через его центр и перпендикулярной направлению \mathbf{N} . Мы получим тогда эллипс, длины и направления главных осей кото-



18.2.

Геометрическое построение для определения направлений поляризации

рого определяют значения N и направление поляризации, т. е. направления векторов $\mathbf{D}^{(1)}$ и $\mathbf{D}^{(2)}$.

Введем наряду с вектором \mathbf{N} еще один важный вектор, характеризующий волну в анизотропной среде, а именно вектор \mathbf{s} , направление которого совпадает с направлением вектора Пойнтинга, или вектора групповой скорости, а длина определяется из соотношения

$$\mathbf{N}\mathbf{s} = N\mathbf{s} \cos \alpha = 1$$

(α — угол между \mathbf{N} и \mathbf{s}). Этот вектор называют *лучевым вектором*. Так как вектор Пойнтинга перпендикулярен \mathbf{E} и \mathbf{H} , то

$$s\mathbf{E} = s\mathbf{H} = 0.$$

Непосредственным вычислением можно показать, что

$$\mathbf{H} = c[\mathbf{sD}], \quad \mathbf{E} = -c\mu_0[\mathbf{sH}].$$

Четыре эти соотношения формально напоминают четыре уравнения Максвелла, если записать последние в форме

$$\mathbf{ND} = \mathbf{NH} = 0, \quad c\mu_0\mathbf{H} = [\mathbf{NE}], \quad c\mathbf{D} = -[\mathbf{NH}].$$

Если учесть аналогию между уравнениями $D_i = \epsilon_0 \epsilon_{ij} E_j$ и $\epsilon_0 E_i = \eta_{ij} D_j$, то мы придем к полезному выводу, сильно облегчающему многие вычисления и геометрические построения, а именно: если справедливо какое-либо уравнение, то справедливо и соотношение, получаемое из него с помощью замен:

$$\epsilon_0 \mathbf{E} \rightarrow \mathbf{D}, \quad \mathbf{D} \rightarrow \epsilon_0 \mathbf{E}, \quad \mathbf{N} \rightarrow \mathbf{s}, \quad \mathbf{s} \rightarrow \mathbf{N}, \quad \epsilon_{ij} \rightarrow \eta_{ij}, \quad \eta_{ij} \rightarrow \epsilon_{ij}.$$

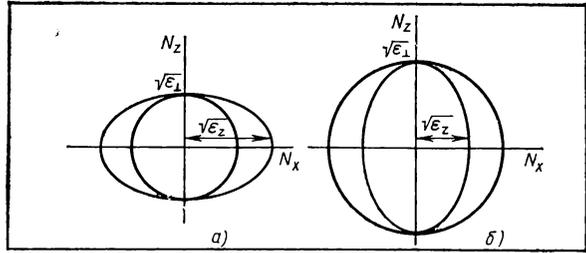
Это правило, в частности, делает простым геометрическое построение, позволяющее получить направление поляризации вектора \mathbf{E} . Для этого нужно построить эллипсоид, главные оси которого направлены по главным диэлектрическим осям, и равны соответственно ϵ_x , ϵ_y , ϵ_z (*эллипсоид Френеля*), и пересечь его плоскостью, перпендикулярной лучевому вектору \mathbf{s} . Мы получим в сечении эллипс; тогда две независимые допустимые поляризации вектора \mathbf{E} определяются направлениями главных осей этого эллипса.

18.3. Одноосные и двухосные кристаллы

Как мы говорили выше, в кубических кристаллах все три главные диэлектрические проницаемости совпадают; такие кристаллы по своим оптическим свойствам ничем не отличаются от изотропных сред. Если симметрия кристалла такова, что совпадают две из трех главных проницаемостей, то кристалл называют *одноосным*. Одна из главных диэлектрических осей при этом сов-

18.3.

Поверхность волновых векторов (a — положительные, b — отрицательные одноосные кристаллы)



падает с осью симметрии кристалла (ось z); о ней говорят как об *оптической оси* кристалла. Две другие оси — это два произвольных взаимно ортогональных направления в базисной плоскости кристалла (x, y). Если $\epsilon_z > \epsilon_\perp$, где $\epsilon_\perp = \epsilon_x = \epsilon_y$, то одноосный кристалл называют *положительным*, в случае же $\epsilon_z < \epsilon_\perp$ — *отрицательным*.

Подставив в уравнение Френеля $\epsilon_x = \epsilon_y \equiv \epsilon_\perp$. Оно распадется тогда на два независимых уравнения:

$$N^2 = \epsilon_\perp; \quad \frac{N_z^2}{\epsilon_\perp} + \frac{N_x^2 + N_y^2}{\epsilon_z} = 1.$$

Используя сформулированную в предыдущем параграфе аналогию между \mathbf{N} и \mathbf{s} , мы можем сразу записать уравнения для лучевого вектора: $s^2 = 1/\epsilon_\perp$, $\epsilon_\perp s_z^2 + \epsilon_z (s_x^2 + s_y^2) = 1$.

Построим поверхности волновых векторов и лучевые поверхности (т. е. поверхности, на которых могут лежать при заданном значении частоты концы векторов \mathbf{N} и \mathbf{s} соответственно). Каждая из этих поверхностей распадется на две: сферу и эллипсоид вращения (рис. 18.3, a, b , на которых изображены продольные сечения поверхности волновых векторов для положительных и отрицательных кристаллов). Мы видим, что сфера и эллипсоид касаются друг друга в двух точках (полюсах), лежащих на оптической оси.

Из уравнений для N и для s следует, что одна из волн, соответствующих данной частоте, распространяется независимо от направления \mathbf{k} с волновым числом $k = (\omega/c)\sqrt{\epsilon_\perp}$, причем в ней $\mathbf{s} \parallel \mathbf{N}$. Иными словами, эта волна ничем не отличается от линейно поляризованной волны в изотропной среде, потому ее называют *обыкновенной волной*.

Что касается второй волны (так называемой *необыкновенной*), то для нее длина волнового вектора зависит от угла θ между ним и оптической осью:

$$\frac{1}{N^2} = \frac{\sin^2 \theta}{\epsilon_z} + \frac{\cos^2 \theta}{\epsilon_\perp}.$$

Лучевой вектор необыкновенной волны лежит в плоскости так называемого *главного сечения*, принадлежащего данному \mathbf{N} , т. е. в плоскости, проходящей через оптическую ось и вектор \mathbf{N} . Он образует с оптической осью угол θ' , связанный с углом θ простым соотношением

$$\operatorname{tg} \theta' = \frac{\varepsilon_2}{\varepsilon_1} \operatorname{tg} \theta.$$

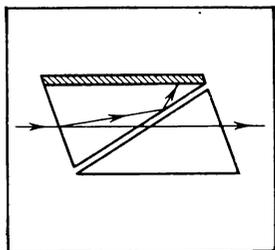
Таким образом, направления \mathbf{s} и \mathbf{N} в необыкновенной волне совпадают только в тех случаях, когда волна распространяется вдоль оси ($\theta = \theta' = 0$) или перпендикулярно ей. Наблюдаемое направление луча определяется вектором \mathbf{s} (а не \mathbf{N}).

Выясним теперь вопрос о поляризованностях обыкновенной и необыкновенной волн. Заметим, для этого, что во всякой волне четыре вектора \mathbf{E} , \mathbf{D} , \mathbf{s} и \mathbf{N} лежат в одной плоскости. В необыкновенной волне направления \mathbf{s} и \mathbf{N} не совпадают, но оба они лежат в плоскости главного сечения, поэтому в этой же плоскости лежат векторы \mathbf{E} и \mathbf{D} . Далее, векторы \mathbf{E} (так же, как и векторы \mathbf{D}) для двух независимо поляризованных волн перпендикулярны друг другу. Следовательно, в обыкновенной волне \mathbf{E} и \mathbf{D} лежат в плоскости, перпендикулярной главному сечению.

Все особенности, отличающие волны в одноосном кристалле от волн в изотропной среде, исчезают, если волна распространяется вдоль оптической оси. При этом нет различия между обыкновенной и необыкновенной волнами, так что плоская монохроматическая волна, поляризована не обязательно линейно, а, вообще говоря, эллиптически.

При преломлении волны (падающей на кристалл не вдоль оси z) в одноосном кристалле возникает две волны (обыкновенная и необыкновенная), волновые векторы которых не равны. Если учесть, что на поверхности кристалла сохраняются частота и тангенциальная составляющая \mathbf{k} , то мы приходим к выводу, что единый луч расщепляется при преломлении на два луча, распространяющиеся под углом друг к другу. Это явление называют *двойным лучепреломлением*. (Обратим внимание на то, что так как наблюдаемое направление луча есть \mathbf{s} , а не \mathbf{N} , то в общем случае это направление лежит вне плоскости падения.)

На явлении двойного лучепреломления базируется целый ряд оптических приборов. Основная их идея заключается в том, что это явление позволяет пространственно разделить два луча с различными (и притом известными) линейными поляризациями. Наиболее известный из таких приборов — *призма Николя* (Николь, 1828 г.). Обычно ее делают из кристалла исландского шпата в виде ромбоэдра, который разрезают по диагонали и склеивают канадским бальзамом (ход лучей в призме Николя показан на рис. 18.4). Углы ромбоэдра подбирают так, чтобы на границе с канадским бальзамом обыкновенный луч испытывал полное внутреннее отражение, в то время как необыкновенный луч свободно проходил через призму. Это возможно потому, что показатель преломления канадского бальзама ($N = 1,53$) меньше показателя преломления для обыкновенной волны в исландском шпате ($N = 1,66$), но



18.4.
Ход лучей в призме Николя

больше показателя [преломления для необыкновенной волны в нем (геометрия призмы такова, что для необыкновенной волны $N = 1,49$).

Призму Николя (или другое устройство), пропускающую свет только заданной поляризации, называют *поляризатором*. Любое такое же устройство может служить в качестве *анализатора* линейно поляризованного света. Если при вращении призмы-анализатора вокруг продольной (по отношению к направлению светового луча) оси в каком-то положении призмы она не пропускает света, то это значит, что анализируемый свет полностью поляризован и вектор \mathbf{D} в нем перпендикулярен главному сечению призмы-анализатора. С помощью подобного рода устройств была установлена поперечность световых волн по отношению к электрическим векторам (Юнг, 1817, Френель, 1819).

Остановимся еще кратко на *двухосных кристаллах*. Для них $\epsilon_x \neq \epsilon_y \neq \epsilon_z$, поэтому при изучении их оптических свойств приходится исследовать уравнение Френеля и другие соотношения § 18.1 в их общем виде. Возникающие при этом формулы громоздки, и мы не будем их приводить. Заранее ясно, однако, что и в двухосных кристаллах имеет место двойное лучепреломление, причем в отличие от кристалла одноосного обе преломленные волны являются необыкновенными.

Из уравнения Френеля вытекает, что при определенных условиях световой луч в кристаллической среде может расщепиться на множество лучей, распространяющихся по поверхности некоторого конуса. Это явление называют *конической рефракцией*, точнее — *внутренней конической рефракцией* (оно было теоретически предсказано в 1832 г. Гамильтоном и в том же году обнаружено экспериментально). Чтобы понять, почему оно происходит, заметим, что тензор диэлектрической проницаемости ϵ_{ik} и обратный ему тензор $(\epsilon^{-1})_{ik}$ представимы в виде

$$\begin{aligned}\epsilon_{ik} &= a\delta_{ik} + b(g_i^{(1)}g_k^{(2)} + g_k^{(1)}g_i^{(2)}); \\ (\epsilon^{-1})_{ik} &= c\delta_{ik} + d(h_i^{(1)}h_k^{(2)} + h_k^{(1)}h_i^{(2)}),\end{aligned}$$

где a, b, c, d — скаляры и $g_i^{(1)}, g_i^{(2)}, h_i^{(1)}, h_i^{(2)}$ — единичные векторы, которые могут зависеть как от ω , так и от \mathbf{k} . Векторы $g_i^{(1)}$ и $g_i^{(2)}$ называют *бирадиальными* или *оптическими осями первого рода*, а векторы $h_i^{(1)}$ и $h_i^{(2)}$ — *бинормальными* или *оптическими осями второго рода*.

Обратимся теперь к уравнению Френеля

$$|\omega^2\epsilon_{ij} + k_i k_j - k^2\delta_{ij}| = 0.$$

Используя приведенные выражения для ϵ_{ij} и $(\epsilon^{-1})_{ij}$, получим после раскрытия определителя квадратное относительно ω^2 уравнение, решение которого имеет вид

$$\omega^2 = ck^2 + d([\mathbf{kh}^{(1)}][\mathbf{kh}^{(2)}] \pm |[\mathbf{kh}^{(1)}]||[\mathbf{kh}^{(2)}]|)$$

(знаки « \pm » соответствуют различным поляризациям волны).

Будем для простоты предполагать, что c , d и $\mathbf{h}^{(1)}$, $\mathbf{h}^{(2)}$ — константы, не зависящие от ω и \mathbf{k} , и вычислим при этих предположениях групповую скорость $\mathbf{v}_g = d\omega/d\mathbf{k}$:

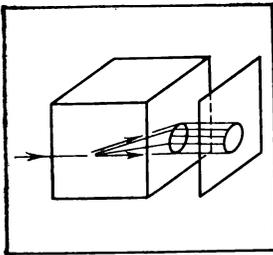
$$\mathbf{v}_g = c\mathbf{v} - \frac{1}{2} \left\{ [\mathbf{h}^{(1)} [\mathbf{h}^{(2)} \mathbf{v}]] + [\mathbf{h}^{(2)} [\mathbf{h}^{(1)} \mathbf{v}]] \pm \left(|[\mathbf{h}^{(1)} \mathbf{v}]| \frac{[\mathbf{h}^{(2)} [\mathbf{h}^{(1)} \mathbf{v}]]}{|[\mathbf{h}^{(2)} \mathbf{v}]|} + |[\mathbf{h}^{(2)} \mathbf{v}]| \frac{[\mathbf{h}^{(1)} [\mathbf{h}^{(2)} \mathbf{v}]]}{|[\mathbf{h}^{(1)} \mathbf{v}]|} \right) \right\},$$

где $\mathbf{v} = \mathbf{k}/\omega$. Мы видим, что если волновой вектор \mathbf{k} направлен вдоль какой-либо из бинормалей, то знаменатель и числитель одной из дробей, входящих в \mathbf{v}_g , обращаются в нуль. Чтобы раскрыть эту неопределенность, предположим, что вектор \mathbf{k} направлен не строго по бинормали, скажем $\mathbf{h}^{(1)}$, а имеет бесконечно малую составляющую $\delta\mathbf{k}_\perp$, перпендикулярно $\mathbf{h}^{(1)}$. Тогда можно показать, что вектор \mathbf{v}_g определяется формулой

$$\mathbf{v}_g = c\mathbf{v} - \frac{d}{2} \left| \sin \varphi \right| \mathbf{v} \left(\frac{\mathbf{h}_\perp^{(2)}}{|\mathbf{h}_\perp^{(2)}|} \mp \frac{\delta\mathbf{v}_\perp}{|\delta\mathbf{v}_\perp|} \right),$$

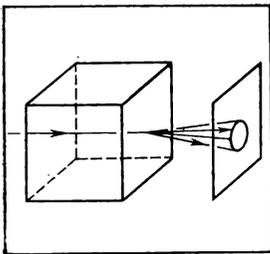
где φ — угол между бинормальями, а $\mathbf{h}_\perp^{(2)}$ — составляющая $\mathbf{h}^{(2)}$, перпендикулярная $\mathbf{h}^{(1)}$.

Мы видим, что бесконечно малая составляющая $\delta\mathbf{k}_\perp$ оказывает конечное воздействие на групповую скорость \mathbf{v}_g . Именно поэтому луч, входящий в кристалл по бинормали $\mathbf{h}^{(1)}$, расщепляется и распространяется по образующим конуса, размеры и положение которого определяются приведенной формулой для \mathbf{v}_g . Это и есть явление *внутренней конической рефракции* (рис. 18.5).



18.5. Внутренняя коническая рефракция

Заметим, что существует еще явление *внешней рефракции*, заключающееся в том, что луч, идущий в кристалле по одной из бинормалей, выходя из кристалла, расщепляется и распространяется по образующим конуса (рис. 18.6).



18.6. Внешняя коническая рефракция

18.4. Искусственная анизотропия

Если деформировать первоначально изотропное тело, то в нем возникнет выделенное направление и тело приобретет свойства одноосного кристалла, в том числе и свойство двойного лучепреломления. Это явление, открытое Зеебеком (1813) и Брюстером (1815), сейчас широко используется для наблюдения за напряжениями в изделиях из обычного и органического стекла, в частности за возникаю-

щими при недостаточно медленном охлаждении стекла остаточными напряжениями. Обычно при этом регистрируется разность хода обыкновенного и необыкновенного лучей. Как всегда, она максимальна для света, падающего перпендикулярно оптической оси, — выделенного направления в образце. Разность хода, естественно, пропорциональна напряжению и длине пути света в теле; коэффициент пропорциональности зависит от вещества, частоты света и температуры и может быть как положительным, так и отрицательным.

Разумеется, если деформировать одноосный кристалл, прикладывая нагрузку в направлении, отличном от направления оптической оси, то он приобретет свойства двухосного кристалла.

Другой способ создания анизотропии в первоначально изотропной среде, применяемый в случае жидких взвесей (суспензии, коллоидальные растворы) с анизотропными по форме частицами, состоит в том, что если такую взвесь заставить течь, то ориентации взвешенных частиц упорядочатся. Степень их упорядочения и, следовательно, возникающая анизотропия пропорциональны градиентам скорости. Этот эффект, обнаруженный *Максвеллом* (1873), позволяет при исследовании потока жидкостей, текущих мимо препятствий, получать информацию о направлении и значении градиента скорости.

Эффективная анизотропия возникнет и в первоначально изотропном теле, например в жидкости, если поместить его в постоянное однородное электрическое поле $E^{(0)}$ (*электрический эффект Керра*, 1875). Эффект Керра в первоначально оптически изотропном теле квадратичен по напряженности поля $E^{(0)}$. Для его описания можно исходить из следующего выражения для тензора диэлектрической проницаемости:

$$\epsilon_{ij} = (\epsilon + AE^{(0)2}) \delta_{ij} + BE_i^{(0)} E_j^{(0)},$$

где ϵ — диэлектрическая проницаемость в отсутствие поля $E^{(0)}$ и A, B — постоянные, зависящие от вещества, температуры и частоты падающего света.

Оптическая ось получающейся одноосной среды (ось z) совпадает, очевидно, с направлением напряженности электрического поля $E^{(0)}$, поэтому двойное лучепреломление в такой среде максимально для света, падающего перпендикулярно напряженности постоянного электрического поля. В этом случае, как мы видели в предыдущем параграфе, показатель преломления обыкновенной волны $N_o = \sqrt{\epsilon_{\perp}}$, а необыкновенной $N_e = \sqrt{\epsilon_z}$. Поэтому на пути l разность хода лучей

$$\begin{aligned} \delta &= l \left[\sqrt{\epsilon + (A + B)(E^{(0)})^2} - \sqrt{\epsilon + A(E^{(0)})^2} \right] \approx \\ &\approx Bl (E^{(0)})^2 / (2\sqrt{\epsilon}). \end{aligned}$$

Вводя вместо разности хода сдвиг фазы $\varphi = k\delta$ (k — волновой вектор), представим его в виде

$$\varphi = 2\pi Kl (E^{(0)})^2.$$

Коэффициент K , связанный с B соотношением $K = B/(4\pi\sqrt{\epsilon_0})$, называют *постоянной Керра*.

Постоянная Керра имеет разный порядок величины и разный знак для разных веществ. Для нитробензола, например, $K = 2 \cdot 10^{-5}$; для воды $K = 5 \cdot 10^{-7}$ (в единицах СГС).

На эффекте Керра основано устройство ряда приборов, с помощью которых осуществляются электрооптическая модуляция и отклонение света.

В случае постоянного однородного магнитного поля H^0 имеет место явление, полностью аналогичное эффекту Керра, — эффект Коттона — Мутона. Все соотношения, описывающие этот эффект, получаются из соответствующих соотношений для эффекта Керра с помощью замены $E^{(0)} \rightarrow H^0$. Постоянная, характеризующая эффект Коттона — Мутона, очень мала, например для нитробензола $K = 2,5 \cdot 10^{-12}$ (в единицах СГСМ). Это значит, что при $H^0 = 3,4$ кЭ и толщине слоя нитробензола 1 см разность хода между обыкновенным и необыкновенным лучами составит $2,8 \cdot 10^{-2}$ длины волны.

ОСНОВНЫЕ ФОРМУЛЫ

Связь между индукциями и напряжениями полей в анизотропной среде

$$D_i = \epsilon_0 \epsilon_{ij} E_j, \quad B_i = \mu_0 \mu_{ij} H_j$$

Вектор гирации

$$g_i = \epsilon_{ijl} \epsilon'_{jl}, \quad \mathbf{D} = \mathbf{D}' + i\epsilon_0 [\mathbf{E}g], \quad \mathbf{D}' = \epsilon_0 \epsilon'_{ij} E_j$$

Плоская монохроматическая волна в анизотропной среде

$$\mathbf{H} = \frac{1}{\mu_0 \omega} [k\mathbf{E}], \quad \mathbf{D} = -\frac{1}{\omega} [k\mathbf{H}]$$

Вектор Пойнтинга и векторный показатель преломления

$$\mathbf{S} = [\mathbf{E}\mathbf{H}], \quad \mathbf{N} = \frac{c}{\omega} \mathbf{k}$$

Фазовая скорость

$$v_\phi = cN/N^2$$

Лучевой вектор

$$s \parallel \mathbf{S}, \quad s\mathbf{N} = 1,$$

$$\mathbf{H} = c[s\mathbf{D}], \quad \mathbf{E} = -\mu_0 c [s\mathbf{H}]$$

Уравнение Френеля

$$N^2 (\epsilon_x N_x^2 + \epsilon_y N_y^2 + \epsilon_z N_z^2) - [N_x^2 \epsilon_x (\epsilon_y + \epsilon_z) + N_y^2 \epsilon_y (\epsilon_z + \epsilon_x) + N_z^2 \epsilon_z (\epsilon_x + \epsilon_y)] + \epsilon_x \epsilon_y \epsilon_z = 0$$

Уравнения Френеля для одноосного кристалла

$$N^2 = \epsilon_\perp, \quad \frac{N_z^2}{\epsilon_\perp} + \frac{N_x^2 + N_y^2}{\epsilon_z} = 1$$

Волновое число для обыкновенной волны

$$k = \frac{\omega}{c} \sqrt{\epsilon_{\perp}}$$

Уравнение необыкновенной волны

$$\frac{\omega^2}{c^2 k^2} = \frac{\sin^2 \theta}{\epsilon_z} + \frac{\cos^2 \theta}{\epsilon_{\perp}}$$

Эффект Керра

$$\epsilon_{ij} = (\epsilon + A E_0^2) \delta_{ij} + B E_{0i} E_{0j}$$



Глава 19. КОЛЕБАНИЯ ГИРОТРОПНЫХ СРЕД

19.1. Вектор гирации

В § 18.1 было показано, что в случае прозрачной (непоглощающей) среды связь между электрической индукцией и напряженностью поля может быть записана в виде

$$\mathbf{D} = \mathbf{D}' + i\epsilon_0 [\mathbf{E}\mathbf{g}],$$

где $\epsilon_0^{-1} \mathbf{D}' = \epsilon'_{ij} E_j$, $\epsilon'_{ij} = \epsilon'_{ji}$ и \mathbf{g} — аксиальный вектор, называемый *вектором гирации*. До сих пор мы рассматривали среды с $\mathbf{g} = 0$; именно к таким средам относится большинство жидкостей, газов и кристаллов в отсутствие внешнего магнитного поля. Переходя теперь к изучению гиротропных сред, выясним прежде всего физический смысл вектора гирации.

Будем для простоты считать, что кроме гиротропии среда не обладает никакой электрической анизотропией, т. е.

$$\epsilon_0^{-1} \mathbf{D} = \epsilon \mathbf{E} + i [\mathbf{E}\mathbf{g}],$$

где ϵ — скалярная диэлектрическая проницаемость; таковы жидкости, газы, плазма, кубические кристаллы.

Как мы видели в § 18.2, в случае плоской монохроматической волны в изотропной среде с $\mathbf{g} = 0$ векторы \mathbf{D} и \mathbf{E} связаны между собой соотношением

$$\epsilon_0^{-1} \mathbf{D} = \epsilon \mathbf{N}^2 - \mathbf{N} (\mathbf{N}\mathbf{E}), \quad \mathbf{N} = kc/\omega.$$

Рассмотрим случай малой гиротропии, когда $|\mathbf{g}| \ll \epsilon$, и поэтому по-прежнему будем пользоваться этим соотношением вместе с соотношением

$$\epsilon_0 \mathbf{E} = \frac{1}{\epsilon} \mathbf{D} + \frac{i}{\epsilon^2} [\mathbf{g}\mathbf{D}],$$

справедливым, как легко видеть, при $|\mathbf{g}| \ll \epsilon$.

Условие совместности этих соотношений имеет (в том же приближении, т. е. при $\mathbf{g} \ll \epsilon$), следующий вид:

$$N = N_{\pm}, \quad N_{\pm} = N_0 \pm \frac{1}{2N_0} g \cos \theta,$$

где $N_0 = \sqrt{\epsilon}$, а θ — угол между векторами \mathbf{g} и \mathbf{N} . Каждому из двух значений N отвечает своя поляризация волны, а именно

$$D_x^{(\pm)} = \pm iD_y^{(\pm)}$$

(ось z направлена вдоль \mathbf{N} ; знак тот же, что в правой части выражения для N). Мы видим, что в гиротропной среде плоские монохроматические волны могут обладать только круговой поляризацией. В том случае, когда вектор \mathbf{D} вращается (если смотреть вдоль направления волнового вектора) направо ($\mathbf{D}^{(+)}$), говорят о *правой*, а когда этот вектор вращается налево — о *левой поляризации* ($\mathbf{D}^{(-)}$) волны.

Пусть теперь линейно поляризованная плоская волна падает в нормальном направлении на плоскую пластинку толщины l , изготовленную из гиротропного вещества. Выберем ось z в направлении падения волны, а ось x — в направлении вектора \mathbf{E} в падающей волне. Линейно поляризованную волну можно представить в виде суперпозиции двух волн с двумя различными круговыми поляризациями; в пластинке эти две волны должны распространяться уже с различными волновыми векторами $k_{\pm} = \frac{\omega}{c} N_{\pm}$. Обозначив D_0 значение D_x при входе света в пластинку (при $z = 0$), имеем

$$D_x = \frac{D_0}{2} (e^{ik_+z} + e^{ik_-z}), \quad D_y = \frac{D_0}{2} (-e^{ik_+z} + e^{ik_-z}).$$

Введем среднее значение волнового числа $k = 1/2(k_+ + k_-)$ и величину $\kappa = 1/2(k_+ - k_-)$. Тогда

$$D_x = D_0 e^{ikz} \cos \kappa z, \quad D_y = D_0 e^{ikz} \sin \kappa z.$$

При выходе из пластинки имеем

$$D_y/D_x = \operatorname{tg} \kappa l.$$

Мы видим, что при выходе из пластинки две волны с круговыми поляризациями складываются снова в линейно поляризованную волну, однако плоскость поляризации оказывается повернутой по отношению к первоначальной на угол κl . С этим свойством среды — *вращением плоскости поляризации* света — и связаны термины «гиротропная среда» и «вектор гирации». Угол поворота направления поляризации пропорционален, как мы видим, длине пути, пройденного светом в гиротропной среде. Для единичного пути он составляет

$$\kappa = \frac{\omega}{2c} (N_+ - N_-) = \frac{\omega \mathbf{g} \cos \theta}{2cN_0}.$$

19.2. Естественная и искусственная гиротропия

Выяснив физический смысл вектора гирации, рассмотрим простой физический пример того, как возникает гиротропия среды — холодной плазмы во внешнем постоянном и однородном магнитном поле с индукцией \mathbf{B}_0 . Гиротропия этой среды в отличие

от большинства других сред может быть объяснена в рамках классической физики; что же касается гиротропии кристаллов, то в основе ее лежат квантовые эффекты, например зеемановское расщепление уровней энергии в магнитном поле.

Радиус-вектор какого-либо электрона плазмы \mathbf{r} определяется, очевидно, уравнением движения

$$\ddot{\mathbf{r}} + \frac{e}{m_e} [\dot{\mathbf{r}} \mathbf{B}_0] = \frac{e}{m_e} \mathbf{E},$$

где \mathbf{E} — напряженность переменного электрического поля волны (мы не учитываем здесь ее магнитного поля). Направляя ось z вдоль \mathbf{B}_0 , перепишем это уравнение в компонентах:

$$\ddot{r}_z = \frac{e}{m_e} E_z, \quad \ddot{r}_\pm \mp i\omega_L \dot{r}_\pm = \frac{e}{m_e} E_\pm,$$

где $r_\pm = r_x \pm ir_y$, $E_\pm = E_x \pm iE_y$, $\omega_L = eB_0/m_e$. Если падающая волна является монохроматической с частотой ω , то

$$-\omega^2 z = \frac{e}{m_e} E_z, \quad -\omega(\omega \pm \omega_L) r_\pm = \frac{e}{m_e} E_\pm.$$

Возникающий в волне дипольный момент единичного объема есть, очевидно, $\mathbf{P} = e\mathbf{r}n_0$, где n_0 — плотность электронов. Поэтому, вводя плазменную частоту $\omega_e = \sqrt{e^2 n_0 / (\epsilon_0 m_e)}$, получим

$$P_x = -\frac{\epsilon_0 \omega_e^2}{2\omega} \left(\frac{E_+}{\omega + \omega_L} + \frac{E_-}{\omega - \omega_L} \right),$$

$$P_y = -\frac{\epsilon_0 \omega_e^2}{2\omega i} \left(\frac{E_+}{\omega + \omega_L} - \frac{E_-}{\omega - \omega_L} \right), \quad P_z = -\frac{\epsilon_0 \omega_e^2}{\omega} E_z.$$

Нам осталось представить с помощью этих соотношений вектор \mathbf{P} в виде $P_i = \epsilon_0 \alpha_{ij} E_j$ и найти тензор диэлектрической проницаемости $\epsilon_{ij} = \delta_{ij} + \alpha_{ij}$. В результате получим

$$\epsilon_{ij} = \begin{pmatrix} \epsilon_1 & i\epsilon_2 & 0 \\ -i\epsilon_2 & \epsilon_1 & 0 \\ 0 & 0 & \epsilon_3 \end{pmatrix},$$

где

$$\epsilon_1 = 1 - \frac{\omega_e^2}{\omega^2 - \omega_L^2}, \quad \epsilon_3 = 1 - \frac{\omega_e^2}{\omega^2}, \quad \epsilon_2 = -\frac{\omega_e^2 \omega_L}{\omega(\omega^2 - \omega_L^2)}.$$

Мы видим, что плазма в магнитном поле является анизотропной ($\epsilon_3 \neq \epsilon_1$) и гиротропной ($\epsilon_2 \neq 0$). Легко проверить, что недиагональные матричные элементы в выражении для тензора ϵ эквивалентны вектору гирации:

$$\mathbf{g} = \epsilon_2 \mathbf{i}_B = -\frac{\omega_e^2 \omega_L \mathbf{i}_B}{\omega(\omega^2 - \omega_L^2)}, \quad \mathbf{i}_B = \frac{\mathbf{B}_0}{B_0}.$$

Но вращение направления поляризации света может иметь место и в отсутствие внешних полей. В частности, этим свойством (называемым *естественной оптической активностью*) обладают кристаллы определенных симметрий, например кристаллы кварца. Именно в кварце это явление наблюдалось впервые (Араго, 1811). Оптическая активность кварца относительно велика (хотя меньше, например, чем у киновари): для желтого света, распространяющегося вдоль оптической оси кристалла, она составляет $\alpha = 21,7$ град·мм⁻¹.

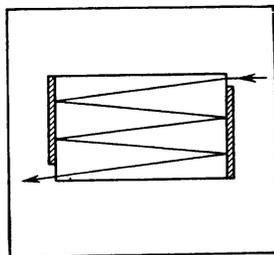
Естественная оптическая активность присуща не только кристаллам, но и растворам и парам большого числа органических соединений, например сахаров или камфоры (это явление впервые наблюдал Био в 1815 г.). У растворов вращательная способность пропорциональна концентрации x ; она зависит от природы растворителя (и, как у всех веществ, от температуры и длины волны). Для водного раствора сахара при комнатной температуре она составляет $6,65$ град·см⁻¹ x , где x — массовая концентрация.

Гиротропия жидкостей и газов является следствием так называемой *оптической изомерии*. Дело в том, что молекулы большого числа органических соединений не симметричны; в простейшем случае у такого соединения возможны два *стереоизомера*, являющихся зеркальным отражением друг друга. Один из изомеров вращает плоскость поляризации направо, другой — налево. Если вещество получено синтетически, то оно представляет собой смесь изомеров в равных количествах и поэтому не вращает плоскость поляризации. Но очень любопытен следующий факт: в живых организмах такие вещества, как правило, представлены только одним изомером. Природная глюкоза, например, вращает плоскость поляризации направо. Оптическая активность растворов широко используется на практике, например в *сахариметрах* — приборах для определения концентрации сахара в растворе.

Естественной гиротропией обладают ферромагнетики. Это явление впервые наблюдал Керр (1876), обнаруживший вращение плоскости поляризации при отражении света от ферромагнетика (в то время прозрачные ферромагнетики не были известны). Реально для наблюдения такого явления (*магнитного эффекта Керра*) кристалл нужно намагнитить, чтобы он был однодоменным или по меньшей мере содержал доменов с одной ориентацией спонтанной намагниченности больше, чем с другой. Однако в этом нет принципиальной необходимости: например, пропуская узкий пучок света через отдельный домен, можно было бы наблюдать вращение плоскости поляризации в ненамагниченном образце.

Во внешнем магнитном поле оптическую активность приобретают и неактивные вещества (*эффект Фарадея*, 1846). В этом случае вращательная способность $\alpha = CH_0$ пропорциональна напряженности поля H_0 ; коэффициент C , называемый *постоянной Верде*, зависит от частоты света и температуры и может быть как положительным, так и отрицательным. Например, для стекла (тяжелый флинт) в желтых лучах $C = 1,25 \div 2$ град·кА⁻¹, для жидкого сероуглерода $C = 0,9$ град·кА⁻¹. Однако, несмотря на малость C , эффект Фара-

дея легко наблюдать даже в не очень сильных полях благодаря остроумному приему, примененному Фарадеем. Дело в том, что знак κ меняется при изменении направления поля на противоположное, так что при прохождении луча через одну и ту же пластинку вперед и назад эффекты складываются. Поэтому суммарный эффект оказывается многократно усиленным, если использовать многократное отражение (например, посеребрив тыловые поверхности пластинки, кроме мест входа и выхода луча; рис. 19.1).



19.1.
Схема усиления эффекта Фарадея

19.3. Спиновые волны и ферромагнитный резонанс

До сих пор мы изучали среды, анизотропные или гиротропные по отношению к электрическим свойствам. Теперь мы рассмотрим среду анизотропную и гиротропную по отношению к магнитным свойствам — ферромагнитный кристалл. Остановимся прежде всего на колебаниях намагниченности в таком кристалле, так называемых *спиновых волнах*.

В гл. 12 мы видели, что магнитный момент \mathbf{m} атома во внешнем поле \mathbf{H} прецессирует около \mathbf{H} , т. е. вращается по закону

$$\frac{\partial \mathbf{m}}{\partial t} = \mu_0 g [\mathbf{mH}]$$

где g — гиромангнитное отношение (если магнитный момент имеет чисто спиновую природу, то $g = e/m_e$, где m_e — масса электрона). В ферромагнетике магнитные моменты электронов находятся под действием эффективного поля \mathbf{H}^{eff} , поэтому процессия i -го момента должна описываться уравнением

$$\frac{\partial \mathbf{m}_i}{\partial t} = \mu_0 g [\mathbf{m}_i \mathbf{H}_i^{\text{eff}}],$$

где $\mathbf{H}_i^{\text{eff}}$ — эффективное поле, действующее на i -й момент.

Если моменты всех атомов параллельны друг другу, то \mathbf{H}^{eff} в отсутствие внешнего поля и магнитной анизотропии совпадает с обменным полем Вейсса $\mathbf{H}_W = \nu n_0 \mathbf{m}$, где n_0 — плотность магнитных электронов и ν — коэффициент пропорциональности (см. § 12.7). Мы видим, что при однородном распределении намагниченности все магнитные моменты параллельны обменному полю, поэтому векторные произведения $[\mathbf{m}_i \mathbf{H}_i^{\text{eff}}]$ обращаются в нуль и моменты не прецессируют.

Выясним, чему равно эффективное поле, если направление намагниченности медленно меняется в пространстве. Будем трактовать моменты как классические векторы и представим обменную энергию взаимодействия электрона i -го атома с электронами всех остальных атомов в виде

$$[U = - \sum_j I_{ij} s_i s_j = - \frac{1}{\mu_B^2} \sum_j I_{ij} m_i m_j.]$$

Обменный интеграл I быстро падает с расстоянием, поэтому в сумме по j достаточно учесть лишь ближайших соседей i -го атома.

Пусть направление магнитных моментов \mathbf{m} медленно меняется вдоль оси x и пусть для простоты ось x совпадает с одной из кристаллографических осей. Тогда если ввести энергию U_0 обменного взаимодействия электрона i -го атома с остальными электронами для однородного распределения моментов, то U представится в виде

$$U = U_0 - \frac{1}{\mu_B^2} I (m_i m_{i+1} + m_i m_{i-1} - 2m_i^2),$$

где $I = I_{i, i+1} = I_{i, i-1}$ — обменный интеграл между ближайшими соседями в направлении оси x .

Если направление \mathbf{m} медленно меняется от атома к атому, то можно считать, что \mathbf{m} — непрерывная функция координат, а \mathbf{m}_i — значение этой функции в точке $\mathbf{r} = \mathbf{r}_i$, где расположен i -й атом: $\mathbf{m}_i = \mathbf{m}(\mathbf{r}_i)$. Тогда

$$\mathbf{m}_{i+1} = \mathbf{m} + \frac{\partial \mathbf{m}}{\partial x} a + \frac{1}{2} \frac{\partial^2 \mathbf{m}}{\partial x^2} a^2 + \dots,$$

$$\mathbf{m}_{i-1} = \mathbf{m} - \frac{\partial \mathbf{m}}{\partial x} a + \frac{1}{2} \frac{\partial^2 \mathbf{m}}{\partial x^2} a^2 + \dots,$$

где a — постоянная решетки. Поэтому

$$U = U_0 - \frac{1}{\mu_B^2} I a^2 \frac{\partial^2 \mathbf{m}}{\partial x^2} \cdot \mathbf{m}.$$

Таким образом, в случае неоднородного распределения магнитных моментов к энергии U_0 электрона i -го атома прибавляется слагаемое

$$\delta U = - \frac{1}{\mu_B^2} I a^2 \frac{\partial^2 \mathbf{m}}{\partial x^2} \cdot \mathbf{m}.$$

Энергия же магнитного момента \mathbf{m} в поле \mathbf{H}^{eff} составляет $\delta U = -\mu_0 \mathbf{m} \mathbf{H}^{\text{eff}}$. Таким образом, величина

$$\mathbf{H}^{\text{eff}} = \frac{1}{\mu_0 \mu_B^2} I a^2 \frac{\partial^2 \mathbf{m}}{\partial x^2}$$

и представляет собой эффективное поле *магнитной неоднородности*, т. е. добавку к обменному полю, обусловленную неоднородностью распределения намагниченности.

Если теперь сложить уравнения прецессии отдельных магнитных моментов по очень малому, но макроскопическому объему, то мы получим уравнение вращения вектора плотности магнитного момента \mathbf{J} :

$$\frac{\partial \mathbf{J}}{\partial t} = \mu_0 g [\mathbf{J} \mathbf{H}^{\text{eff}}].$$

Эффективное же поле \mathbf{H}^{eff} окажется пропорциональным второй производной от вектора \mathbf{J} :

$$\mathbf{H}^{\text{eff}} = \alpha \frac{\partial^2 \mathbf{J}}{\partial x^2}, \quad \alpha \equiv \frac{l\alpha^2}{n_0 \mu_0 \mu_B^2}.$$

Величина α совпадает с введенной в § 12.8 постоянной неоднородного обменного взаимодействия.

Предположим теперь, что равновесное направление магнитного момента совпадает с осью z . Тогда от координат и времени зависят компоненты J_x и J_y вектора \mathbf{J} . Если при этом J_x и J_y малы, то J_z очень мало отличается от равновесной намагниченности J_0 , а именно $J_z \approx J_0[1 - (J_x^2 + J_y^2)/(2J_0^2)]$. Поэтому векторное уравнение вращения магнитного момента можно представить в виде двух скалярных уравнений:

$$\frac{\partial J_x}{\partial t} = -\mu_0 g J_0 \alpha \frac{\partial^2 J_y}{\partial x^2}; \quad \frac{\partial J_y}{\partial t} = \mu_0 g J_0 \alpha \frac{\partial^2 J_x}{\partial x^2}.$$

Если обозначить \mathbf{J}_\perp проекцию намагниченности на плоскость, перпендикулярную \mathbf{J}_0 (плоскость x, y), то при вращении намагниченности длина вектора \mathbf{J}_\perp не меняется, а сам этот вектор равномерно вращается. Таким образом, если $J_x = J_\perp \cos \varphi$ и $J_y = J_\perp \sin \varphi$, то во времени и пространстве меняется только угол φ , но не величина J_\perp . Такие колебания называют, как мы знаем, поляризованными по кругу или циркулярно. Вводя вместо J_x и J_y величину $J_+ = J_x + iJ_y$, мы получим одно уравнение:

$$\frac{\partial J_+}{\partial t} = i\alpha \mu_0 g J_0 \frac{\partial^2 J_+}{\partial x^2}.$$

Будем искать решение этого уравнения в виде плоской монохроматической волны $J_+ \sim \exp(i\mathbf{k}\mathbf{r} - i\omega t)$. Подстановка этого выражения в уравнение для J_+ приводит к следующей зависимости частоты волны ω от волнового вектора \mathbf{k} :

$$\omega(\mathbf{k}) = \alpha \mu_0 g J_0 k^2$$

Отклонение магнитного момента от равновесного направления распространяется в ферромагнетике в виде колебаний, частота которых квадратично зависит от волнового вектора. Эти колебания называют спиновыми волнами.

Существование спиновых волн было предсказано Блохом.

Мы знаем, что каждому колебанию с частотой ω и волновым вектором \mathbf{k} сопоставляется частица с энергией $\varepsilon = \hbar\omega$ и импульсом $\mathbf{p} = \hbar\mathbf{k}$; например, звуковым волнам сопоставляются частицы — *фононы*, у которых $\varepsilon = v_s p$ (v_s — скорость звука). Аналогично, спиновой волне сопоставляется частица, называемая *магноном*. Для магнона

$$\varepsilon = \frac{1}{\hbar} \alpha g \mu_0 J_0 p^2.$$

Мы видели в гл. 12, что у многих ферромагнетиков намагниченность при нуле температур равна максимально возможному значению $J_\infty = \mu_B n_0$, где n_0 — плотность магнитных электронов. При очень низких температурах намагниченность уменьшается с ростом температуры, поскольку из-за тепловых колебаний магнитные моменты случайно отклоняются от равновесия. Но такие отклонения обязательно распространяются в виде спиновых волн, поэтому уменьшение намагниченности пропорционально числу тепловых магнонов — спиновых волн, обусловленных, как уже говорилось, тепловыми колебаниями магнитных моментов. Определим число тепловых магнонов.

Магноны, подобно фононам, подчиняются статистике Бозе—Эйнштейна, поэтому их распределение описывается планковской функцией

$$N(\mathbf{p}) = \frac{1}{e^{\varepsilon/(k_B T)} - 1}$$

(k_B — постоянная Больцмана). Таким образом, средняя плотность магнонов

$$N_m = \int \frac{1}{e^{\varepsilon/(k_B T)} - 1} \frac{d^3 p}{(2\pi\hbar)^3}.$$

Если вместо энергии ввести безразмерную переменную $x = \varepsilon/(k_B T)$, то импульс p пропорционален \sqrt{T} ; поэтому интеграл в выражении для N_m пропорционален $T^{3/2}$. Иными словами, плотность тепловых магнонов возрастает с температурой как $T^{3/2}$; по этому же закону возрастает разность $J_\infty - J_s$:

$$(J_\infty - J_s)/J_\infty \sim T^{3/2}/T_c^{3/2}.$$

Это и есть упоминавшийся ранее (см. гл. 12) закон Блоха для температурной зависимости намагниченности.

В гл. 12 мы видели, что наряду с обменной энергией спины электронов обладают еще дополнительной энергией, зависящей от их ориентации относительно кристаллографических осей, — так называемой энергией магнитной анизотропии. В случае одноосных кристаллов мы использовали следующее выражение для плотности этой энергии: $\omega_a = K \sin^2 \theta$, где θ — угол между направлением намагниченности и осью кристалла (ось z). Отбрасывая несущественное постоянное слагаемое (что эквивалентно просто переопределению начала отсчета энергии) и вводя безразмерную постоянную $\beta = 2K/(\mu_0 J_0^2)$, запишем плотность энергии магнитной анизотропии в виде

$$\omega_a = -\frac{\beta}{2} \mu_0 J_z^2.$$

Постоянную β называют *постоянной магнитной анизотропии*; она зависит от температуры кристалла и внешних напряжений, но не от намагниченности. Будем считать ее для определенности положительной; тогда в равновесии магнитные моменты направлены по оси z (*анизотропия типа легкой оси*).

Сравним теперь формулу для ω_a с выражением $W = -\mu_0 \int \mathbf{H} d\mathbf{J}$ для энергии момента \mathbf{J} в магнитном поле \mathbf{H} . Мы видим, что существование энергии магнитной анизотропии эквивалентно введению некоторого эффективного поля \mathbf{H}_a , называемого *полем магнитной анизотропии*. Оно направлено по оси z и равно по модулю βJ_z , т. е.

$$\mathbf{H}_a = \beta \mathbf{J}_0 \frac{(\mathbf{J}\mathbf{J}_0)}{J_0^2},$$

где \mathbf{J}_0 — равновесная намагниченность.

Поле магнитной анизотропии значительно меньше обменного поля, поэтому при рассмотрении спиновой волны с немалыми \mathbf{k} ее можно было не учитывать. Но для очень длинных волн, как мы видели, вклад обменного поля (поле магнитной неоднородности) обращается в нуль, поэтому учет \mathbf{H}_a становится необходимым.

Итак, для изучения колебаний намагниченности во всем диапазоне длин волн следует исходить из уравнения

$$\frac{\partial \mathbf{J}}{\partial t} = \mu_0 g [\mathbf{J} \mathbf{H}^{eff}],$$

где

$$\mathbf{H}^{eff} = \alpha \frac{\partial^2 \mathbf{J}}{\partial x^2} + \beta \mathbf{J}_0 \frac{(\mathbf{J}\mathbf{J}_0)}{J_0^2} + \mathbf{H}$$

и \mathbf{H} — напряженность магнитного поля в кристалле. Добавляя к этому уравнению уравнения магнитостатики

$$\text{div}(\mathbf{H} + \mathbf{J}) = 0, \quad \text{rot} \mathbf{H} = 0,$$

можно найти зависимость частоты спиновой волны от волнового вектора при всех \mathbf{k} :

$$\omega(\mathbf{k}) = \mu_0 g J_0 \sqrt{\left(\alpha k^2 + \beta + \frac{H_0}{J_0}\right) \left(\alpha k^2 + \beta + \frac{H_0}{J_0} + \sin^2 \theta\right)},$$

где H_0 — напряженность невозмущенного магнитного поля (которая считается постоянной, однородной и направленной вдоль оси анизотропии z) и θ — угол между \mathbf{k} и \mathbf{H}_0 .

С существованием спиновых волн связано наблюдающееся в ферромагнетиках явление *ферромагнитного резонанса*. Оно заключается в следующем. Поместим ферромагнитный образец в однородное высокочастотное поле $\sim \exp(-i\omega t)$. Если частота ω совпадает с частотой ω_0 спиновой волны при $k = 0$, то энергия высокочастотного поля интенсивно поглощается, превращаясь в энергию колебания спинов. Это и есть ферромагнитный резонанс.

Определим частоту ω_0 ферромагнитного резонанса. Для этого в выражении для $\omega(\mathbf{k})$ положим $\mathbf{k} = 0$. При этом, однако, возникает неопределенность в значении $\sin\theta$. Эта неопределенность раскрывается в зависимости от геометрии образца. Например, для плоской пластины, поверхность которой параллельна оси легкого намагничивания, $\sin\theta = 1$ и

$$\omega_0 = \mu_0 g J_0 \sqrt{\left(\beta + \frac{H_0}{J_0}\right) \left(\beta + \frac{H_0}{J_0} + 1\right)}.$$

Аналогичное явление имеет место и в парамагнетиках (*парамагнитный резонанс*). Заключается оно в интенсивном поглощении энергии высокочастотного магнитного поля, если его частота совпадает с частотой прецессии ω_0 атомных магнитных моментов около внешнего магнитного поля. Если, например, постоянное однородное магнитное поле H_0 параллельно поверхности образца в форме плоской пластинки, то

$$\omega_0 = \mu_0 g H_0 \sqrt{\mu},$$

где μ — статическая магнитная проницаемость парамагнетика.

19.4. Связанные электромагнитно-спиновые волны

Перейдем к определению тензора высокочастотной магнитной восприимчивости ферромагнетика. Пусть ферромагнетик (для определенности имеющий анизотропию типа легкая ось) находится в магнитном поле $\mathbf{H} = \mathbf{H}_0 + \mathbf{h}$, где \mathbf{H}_0 — напряженность постоянного однородного магнитного поля и $\mathbf{h} = \mathbf{h}_0 e^{-i\omega t}$ — напряженность слабого высокочастотного магнитного поля. Согласно уравнению прецессии магнитного момента, $\partial \mathbf{J} / \partial t = \mu_0 g [\mathbf{J} \mathbf{H}^{\text{eff}}]$, где \mathbf{H}^{eff} — напряженность эффективного поля:

$$\mathbf{H}^{\text{eff}} = \alpha \frac{\partial^2 \mathbf{J}}{\partial x^2} + \beta \mathbf{J}_0 (\mathbf{J} \mathbf{J}_0) J_0^{-2} + \mathbf{H}_0 + \mathbf{h}.$$

Подставляя сюда \mathbf{J} в виде $\mathbf{J} = \vec{\mathbf{J}}_0 + \tilde{\mathbf{J}}$, где $\tilde{\mathbf{J}} = \tilde{\mathbf{J}}_0 \exp(-i\omega t)$ — высокочастотная составляющая намагниченности, и линеаризуя это уравнение относительно $\tilde{\mathbf{J}}$, получим связь между векторами $\tilde{\mathbf{J}}$ и \mathbf{h} :

$$\tilde{\mathbf{J}}_i = \chi_{ij} h_j.$$

Входящий в это соотношение тензор χ_{ij} и представляет собой тензор высокочастотной магнитной восприимчивости ферромагнетика. Он имеет следующие компоненты:

$$\chi_{ij}(\omega, \mathbf{k}) = \begin{pmatrix} \chi_{\perp} & \frac{i\tilde{\mu}}{4\pi} & 0 \\ -\frac{i\tilde{\mu}}{4\pi} & \chi_{\perp} & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

где

$$\chi_{\perp} \equiv \chi_{\perp}(\omega, k) = \frac{\mu_0 g J_0 \Omega_0}{\Omega_0^2 - \omega^2}, \quad \tilde{\mu} \equiv \tilde{\mu}(\omega, k) = \frac{\mu_0 g J_0 \omega}{\Omega_0^2 - \omega^2},$$

и $\Omega_0 = \mu_0 g J_0 (\alpha k^2 + \beta + H_0/J_0)$ (внешнее постоянное однородное магнитное поле \mathbf{H}_0 направлено вдоль оси магнитной анизотропии — оси z).

Тензор высокочастотной магнитной проницаемости $\mu \equiv \mu(\omega, \mathbf{k})$ связан с тензором χ общим соотношением

$$\mu_{ij}(\omega, \mathbf{k}) = \delta_{ij} + \chi_{ij}(\omega, \mathbf{k}).$$

Мы видим, что ферромагнетик представляет собой анизотропную и гиротропную среду. Связь между магнитными индукцией и напряженностью поля для него имеет вид

$$\mathbf{B} = \mathbf{J}_0 \frac{(\mathbf{h}\mathbf{J}_0)}{J_0^2} + \mu_{\perp} \left(\mathbf{h} - \mathbf{J}_0 \frac{(\mathbf{h}\mathbf{J}_0)}{J_0^2} \right) + i [\mathbf{h}\mathbf{g}_m],$$

где $\mu_{\perp} = 1 + \chi_{\perp}$ и \mathbf{g}_m — вектор магнитной гирации:

$$\mathbf{g}_m = \tilde{\mu} \frac{\mathbf{J}_0}{J_0} = \frac{\mu_0 g \mathbf{J}_0 \omega}{\Omega_0^2 - \omega^2}.$$

При известной магнитной проницаемости можно исследовать взаимодействие между спиновыми и собственно электромагнитными волнами. Такое взаимодействие особенно интенсивно, если частоты и волновые векторы тех и других волн близки между собой, т. е. находятся вблизи резонансного значения волнового вектора k_0 , определяющегося из уравнения

$$\omega_s(\mathbf{k}_0) = ck_0/\sqrt{\epsilon},$$

где $\omega_s(\mathbf{k}_0)$ — частота спиновой волны и ϵ — диэлектрическая проницаемость (по отношению к электрическим свойствам ферромагнетик считается изотропным). Частоту $\omega_s(\mathbf{k}_0)$ называют *частотой электромагнитно-спинового резонанса*.

Замечая, что $\mu_0 g J_0 \sim 3 \cdot 10^{10} \text{ с}^{-1}$ и $\alpha \sim \Theta_C a^2 / (\mu_0 \mu_B J_0) \sim 10^{-16} \text{ м}^2$, мы видим, что интенсивно взаимодействуют только очень длинноволновые колебания с длиной волны порядка нескольких сантиметров (Θ_C — температура Кюри).

Рассмотрим взаимодействие спиновых и собственно электромагнитных волн. Для этого запишем уравнения Максвелла для плоской электромагнитной волны и исключим из них электрические поле и индукцию. В результате получим

$$\mathbf{B} = \frac{\mu_0 c^2}{\epsilon \omega^2} [k^2 \mathbf{h} - \mathbf{k}(\mathbf{k}\mathbf{h})].$$

С другой стороны, $B_i = \mu_0 \mu_{ij} h_j$. Приравняв оба выражения, придем к системе уравнений

$$\left\{ \frac{c^2}{\epsilon \omega^2} (k^2 \delta_{ij} - k_i k_j) + \mu_{ij} \right\} h_j = 0.$$

Условие существования у полученной системы уравнений нетривиального решения, т. е. равенство нулю определителя

$$\det \left\{ (k^2 \delta_{ij} - k_i k_j) + \frac{\omega^2}{\varepsilon c^2} \mu_{ij}(\omega, \mathbf{k}) \right\} = 0,$$

и определяет зависимость ω от \mathbf{k} для всех рассматриваемых ветвей колебаний.

Если бы магнитная проницаемость ферромагнетика не зависела от частоты, то это уравнение свелось бы к квадратному уравнению относительно переменной $\omega^2/(c^2 k^2)$. Иными словами, мы получили бы две ветви собственно электромагнитных волн, частота которых была бы (при заданном направлении распространения) пропорциональна волновому числу. Но компоненты тензора магнитной проницаемости вносят дополнительную зависимость от частоты, причем ω^2 начинает входить в уравнение не только в комбинации $\omega^2/(c^2 k^2)$, но и независимо. В результате появляется еще одна ветвь функции $\omega(\mathbf{k})$, или, что то же самое, еще одна, третья ветвь колебаний.

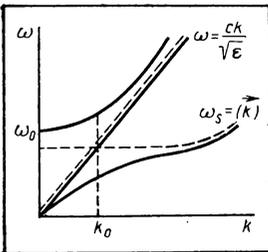
Общие формулы для частот трех ветвей громоздки, и мы не будем их приводить. Они упрощаются только вдали от резонанса, т. е. при $k \gg k_0$ и при $k \ll k_0$, и в окрестности резонанса, когда $|k - k_0| \ll k_0$ (напомним, что для резонанса $\omega_s(k_0) = ck_0/\sqrt{\varepsilon}$).

При $k \gg k_0$ мы имеем две ветви собственно электромагнитных волн с частотой $ck/\sqrt{\varepsilon}$ и одну ветвь спиновых волн с частотой $\omega_s(\mathbf{k})$. В области $k \ll k_0$ частота одной из ветвей колебаний много больше, чем двух других, и при $k \rightarrow 0$ стремится к ω_0 , где

$$\omega_0 = \mu_0 g (H_0 + J_0 + \beta J_0);$$

частоты же двух других ветвей зависят от k по линейному закону. Наконец, в резонансной области ($|k - k_0| \ll k_0$) частоты двух ветвей колебаний отличаются от $\omega_s(k_0)$ на величины, пропорциональные $\pm \sqrt{|k - k_0|}$, в то время как частота третьей ветви отличается от $\omega_s(k_0)$ на величину, пропорциональную $|k - k_0|$. Зависимость частоты от волнового вектора для трех ветвей колебаний представлена схематически на рис. 19.2.

Остановимся кратко на вращении плоскости поляризации света в ферромагнетиках. Пусть линейно поляризованная электромагнитная волна падает нормально на ферромагнитный образец и ось анизотропии (ось z) ферромагнетика совпадает с направлением нормали к образцу поверхности. Будем для простоты считать $\omega \gg ck_0$; тогда внутри образца мы получим две волны с равными амплитудами, но разными волновыми числами:



19.2.

Зависимость частоты от волнового вектора для трех ветвей электромагнитно-спиновых волн в ферромагнетике

Если h_0 — значение амплитуды волны при входе в образец, то напряженность поля в образце определится формулами

$$h_x = h_0 \cos \chi z e^{i(kz - \omega t)}, \quad h_y = h_0 \sin \chi z e^{i(kz - \omega t)},$$

где $k = 1/2(k_+ + k_-)$ и $x = 1/2(k_+ - k_-)$ (ось x выбрана в направлении магнитного поля падающей волны). Очевидно, при выходе из образца направление поляризации волны оказывается повернутым на угол

$$\varphi = xl = \frac{\omega l}{2c} \left(\sqrt{\mu_{\perp}(\omega) + \tilde{\mu}(\omega)} + \sqrt{\mu_{\perp}(\omega) - \tilde{\mu}(\omega)} \right),$$

где l — путь света в кристалле.

ОСНОВНЫЕ ФОРМУЛЫ

Угол поворота плоскости поляризации электромагнитной волны в гиротропной среде на единичном пути

$$x = \frac{\omega g \cos \theta}{2cN_0}$$

Вектор гирации плазмы в магнитном поле

$$g = - \frac{\omega_e^2 \omega_L}{\omega (\omega^2 - \omega_L^2)} \frac{B_0}{B_0}$$

Частота спиновой волны

$$\omega(\mathbf{k}) = \mu_0 \alpha g J_0 k^2, \quad g = e/m_e$$

Закон Блоха

$$\frac{J_{\infty} - J_s}{J_{\infty}} \sim T^{3/2}$$

Частота ферромагнитного резонанса пластинки

$$\omega_0 = \mu_0 g J_0 \sqrt{\left(\beta + \frac{H_0}{J_0}\right) \left(\beta + \frac{H_0}{J_0} + 1\right)}$$



Глава 20. ГЕОМЕТРИЧЕСКАЯ ОПТИКА

20.1. Уравнение эйконала

Уравнения Максвелла охватывают все электромагнитные явления, в том числе и оптические, и поэтому любая задача, относящаяся к кругу этих явлений, может быть, в принципе, сформулирована на языке уравнений Максвелла как строгая математическая задача. Однако такой общий подход, как правило, очень сложен, потому что при этом нужно учитывать как граничные, так и начальные условия. Точные решения уравнений Максвелла можно поэтому получить только в ряде простейших случаев, когда задача характеризуется значительной симметрией. Мы убедились в этом уже в случае электростатики, задачи же в случае переменных полей гораздо сложнее статических задач.

Между тем имеется обширная область переменных полей, в которой уравнения Максвелла могут быть значительно упрощены. Это

область очень коротких волн, настолько коротких, что длина волны в первом приближении вообще не фигурирует. Для этого нужно, чтобы длина волны λ была значительно меньше всех характерных размеров задачи: $\lambda \ll a$, где a — наименьший из этих размеров. В этом случае распространение электромагнитного поля происходит вдоль определенных геометрических линий, называемых *лучами*, и свойства этих лучей не зависят в первом приближении от длины волны. Иначе говоря,

при $\lambda \ll a$ законы электромагнитного поля могут быть сформулированы на языке геометрии. Поэтому случай $\lambda \ll a$ носит название геометрической (или лучевой) оптики

— оптики потому, что видимый свет характеризуется очень малой длиной волны (от $4 \cdot 10^{-7}$ до $8 \cdot 10^{-7}$ м) по сравнению с размерами оптических приборов.

В следующем приближении по малому параметру λ/a начнет проявляться волновая природа света. Эти проявления носят общее название *дифракционных явлений*.

Прежде всего мы разъясним, как вводится понятие луча, затем установим законы распространения лучей и, наконец, рассмотрим дифракционные явления.

Будем исходить из уравнений Максвелла для среды, которую будем считать изотропной и характеризовать зависящими от координат диэлектрической и магнитной проницаемостями $\epsilon = \epsilon(\mathbf{r})$, $\mu = \mu(\mathbf{r})$. Считая поле монохроматическим:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}(\mathbf{r}) e^{-i\omega t}, \quad \mathbf{H}(\mathbf{r}, t) = \mathbf{H}(\mathbf{r}) e^{-i\omega t},$$

имеем

$$\begin{aligned} \operatorname{rot} \mathbf{H} + i\epsilon_0 \epsilon \omega \mathbf{E} &= 0, & \operatorname{rot} \mathbf{E} - i\mu_0 \mu \omega \mathbf{H} &= 0, \\ \operatorname{div} \epsilon \mathbf{E} &= 0, & \operatorname{div} \mu \mathbf{H} &= 0. \end{aligned}$$

Для учета временной дисперсии проницаемостей достаточно, очевидно, считать ϵ и μ зависящими, как от параметра, от частоты ω .

Если проницаемости не зависят от координат, то полученные уравнения допускают решения в виде плоских волн:

$$\mathbf{E}(\mathbf{r}) = \mathbf{A} e^{i\mathbf{k}\mathbf{r}}, \quad \mathbf{H}(\mathbf{r}) = \mathbf{B} e^{i\mathbf{k}\mathbf{r}},$$

где \mathbf{A} и \mathbf{B} — постоянные векторы, $\mathbf{k} = (\omega/c) N \mathbf{n}$ — волновой вектор (не зависящий от координат), $N = \sqrt{\epsilon \mu}$ — показатель преломления и \mathbf{n} — единичный вектор в направлении \mathbf{k} .

В приближении геометрической оптики поля $\mathbf{E}(\mathbf{r})$ и $\mathbf{H}(\mathbf{r})$ имеют структуру, близкую к структуре полей в плоской волне, а именно

$$\mathbf{E}(\mathbf{r}) = \mathbf{A}(\mathbf{r}) e^{(i\omega/c) L(\mathbf{r})}, \quad \mathbf{H}(\mathbf{r}) = \mathbf{B}(\mathbf{r}) e^{(i\omega/c) L(\mathbf{r})},$$

где $L(\mathbf{r})$ — некоторая функция точки \mathbf{r} , заменяющая входившую в выражение для плоской волны функцию $N \mathbf{n} \mathbf{r}$, и $\mathbf{A}(\mathbf{r})$, $\mathbf{B}(\mathbf{r})$ — медленно

меняющиеся в пространстве величины. Функцию $L(\mathbf{r})$ называют *оптическим путем* или *эйконалом* («эйкон» — по-гречески «изображение», откуда наше слово «икона»).

Рассмотрим небольшие пространственные участки Δr , такие, что $\lambda_0 \ll \Delta r \ll a$ (здесь $\lambda_0 = 2\pi c/\omega$ — длина волны в вакууме). На таких участках функция $L(\mathbf{r})$ мало отличается от линейной, поэтому имеет смысл ввести локальный волновой вектор $\mathbf{k} = \mathbf{k}(\mathbf{r})$, связанный с $L(\mathbf{r})$ соотношением

$$\mathbf{k}(\mathbf{r}) = \frac{\omega}{c} \frac{\partial L}{\partial \mathbf{r}}.$$

Линия, касательная к кривой $\mathbf{k}(\mathbf{r})$ в каждой точке \mathbf{r} , носит название луча. Уравнение для луча может быть, очевидно, записано в форме

$$N(\mathbf{r}) \frac{d\mathbf{r}}{ds} = \text{grad } L$$

(s — длина дуги луча). При зависящем от координат показателе преломления $N(\mathbf{r})$ луч не является прямой линией.

Поверхности $L(\mathbf{r}) = \text{const}$ носят название *волновых поверхностей*. При N , не зависящем от \mathbf{r} , волновые поверхности — это плоскости, перпендикулярные волновому вектору. В общем случае $N = N(\mathbf{r})$ волновые поверхности уже не являются плоскими.

Лучи образуют семейство линий, ортогональных к волновым поверхностям

Локальный волновой вектор $\mathbf{k}(\mathbf{r})$ связан с частотой и показателем преломления таким же соотношением $k = (\omega/c)N$, как и в случае плоской волны. Различие состоит лишь в том, что теперь k и N зависят от координат точки. Вспоминая связь между \mathbf{k} и L , мы видим, что $(\text{grad } L)^2 = N^2$, т. е.

$$\left(\frac{\partial L}{\partial x}\right)^2 + \left(\frac{\partial L}{\partial y}\right)^2 + \left(\frac{\partial L}{\partial z}\right)^2 = N^2.$$

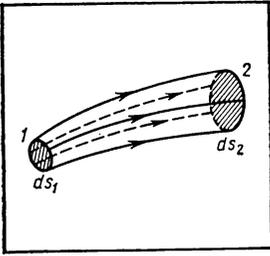
Это и есть основное уравнение геометрической оптики, называемое уравнением эйконала.

20.2. Интенсивность света

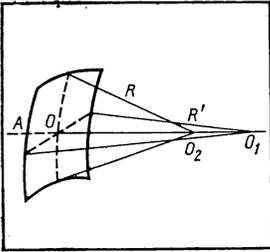
Как и в случае плоской монохроматической волны, в геометрической оптике плотности электрической и магнитной энергий (усредненные по времени) равны между собой ($\langle w_e \rangle = \langle w_m \rangle$) и среднее значение вектора Пойнтинга \mathbf{S} определяется формулой

$$\langle \mathbf{S} \rangle = v_{\phi} \langle w \rangle \mathbf{n},$$

где $\langle w \rangle = \langle w_e \rangle + \langle w_m \rangle$, $v_{\phi} = c/N$ и \mathbf{n} — единичный вектор вдоль луча (среда предполагается бездисперсной). Таким образом, усредненный по времени вектор Пойнтинга направлен по лучу, т. е. ортогонально волновому фронту, скорость же распространения энергии равна c/N .



20.1.
Трубка световых лучей



20.2.
К выводу закона интенсивности в геометрической оптике

Интенсивность света определяется как $I = v_{\Phi} \langle w \rangle$. Из закона сохранения энергии вытекает дифференциальное соотношение

$$\operatorname{div} I n = 0.$$

Это соотношение допускает простую геометрическую интерпретацию. Рассмотрим узкую трубку, образованную световыми лучами (рис. 20.1). Пусть ds_1 и ds_2 — элементы площадей двух замыкающих трубку волновых фронтов $L(\mathbf{r}) = L_1$ и $L(\mathbf{r}) = L_2$ (L_1, L_2 — постоянные) и пусть I_1 и I_2 — соответствующие значения интенсивности. Тогда потоки энергии через поверхности 1 и 2, т. е. величины $I_1 ds_1$ и $I_2 ds_2$, должны быть равны между собой. Учитывая, что площадь ds пропорциональна произведению главных радиусов кривизны R и R' соответствующей поверхности, получим $I_1 R_1 R'_1 = I_2 R_2 R'_2$, откуда

$$I = C / (RR'),$$

где C — константа для данной лучевой трубки. Иными словами, интенсивность света в какой-либо точке O луча A (рис. 20.2) обратно пропорциональна произведению главных радиусов кривизны (OO_1 и OO_2) волновой поверхности, проходящей через эту

точку, причем коэффициент пропорциональности одинаков для всех точек одного и того же луча (но не для разных точек волновой поверхности).

Если свет испускается точечным источником, то оба центра кривизны и, следовательно, оба радиуса кривизны совпадают ($R = R'$), волновыми поверхностями при этом являются концентрические сферы с центром в источнике света. Таким образом,

интенсивность света, испускаемого точечным источником, обратно пропорциональна квадрату расстояния до источника.

Полевые величины f изменяются при этом вдоль луча по закону

$$f = \frac{\operatorname{const}}{R} e^{ikR},$$

где kR определяет фазу волны.

Если хотя бы один из радиусов кривизны равен нулю, то интенсивность, согласно геометрической оптике, обращается в бесконечность. Геометрические места всех центров кривизны волновой поверхности образуют две поверхности (различные для точек O и O'), называемые *каустиками*. Для сферической волновой поверхности обе каустики сливаются в одну точку — *фокус*. Отметим, что лучи касаются каустик. В действительности, конечно, интенсивность на каустике

не обращается в бесконечность — этот вывод связан лишь с приближением геометрической оптики ($\lambda \rightarrow 0$). При учете конечности λ интенсивность на каустиках оказывается пропорциональной $\lambda^{-1/3}$.

В заключение этого параграфа дадим определения принятых в СИ единиц для выражения световых величин. В качестве основной единицы принята **кандела** (кд) — сила света, излучаемого абсолютно черным телом с поверхности площадью $1/600\,000$ м² в перпендикулярном направлении при температуре затвердевания платины и давлении 101 325 Па. Размерность силы света имеет также $dI/d\Omega$; если в телесный угол, равный 1 ср, излучается мощность в 1 кд, то говорят, что световой поток равен одному **люмену** (лм); таким образом, 1 лм = = 1 кд/ср. В качестве единицы освещенности принят **люкс**: 1 лк = = 1 лм/м².

20.3. Принцип Ферма

Найдем время t_{AB} , затрачиваемое светом, чтобы пройти путь между двумя точками A и B одного и того же луча. Вспоминая, что $v_\phi = c/N$, имеем

$$t_{AB} = \frac{1}{c} \int N ds,$$

где ds — элемент длины луча. Эту же формулу можно представить в виде $t_{AB} = (1/c)L(AB)$, где

$$L(AB) = \int_A^B dr \operatorname{grad} L \equiv L(B) - L(A)$$

и $L(A)$, $L(B)$ — значения эйконала $L(r)$ в точках A и B . В самом деле, $\operatorname{grad} L = nN$, а при движении вдоль луча $dr = n ds$. Величину $L(AB)$ называют *оптической длиной луча* между точками A и B .

Оптическую длину можно ввести не только для реального луча, но и для произвольной кривой C , соединяющей точки A и B :

$$L_C(AB) = \left(\int_A^B N dl \right)_C,$$

где dl — элемент длины кривой C . Чем же отличается реальный луч от произвольной кривой, соединяющей те же крайние точки A и B ? Принцип Ферма утверждает, что

оптическая длина реального луча всегда меньше оптической длины любой другой кривой, соединяющей те же точки. Иными словами, свет распространяется таким образом, чтобы время его распространения было минимальным.

Для того чтобы убедиться в этом, заметим, что интеграл $\int_A^B dr \operatorname{grad} L$, взятый вдоль некоторой кривой C , соединяющей точки A и B , не зависит от вида кривой и всегда равен $L(B) - L(A)$. С другой стороны, так как $\operatorname{grad} L = Nn$, то его можно записать в виде

$$\left(\int_A^B dl \cos \alpha N \right)_C = L(B) - L(A),$$

где dl — элемент длины кривой C и α — угол между dl и n . Для истинного луча $\alpha = 0$, т. е.

$$\left(\int_A^B dLN \right)_{\text{луч}} = L(B) - L(A).$$

Пусть теперь кривая C близка к лучу, так что $\cos \alpha > 0$. Тогда

$$\left(\int_A^B dLN \right)_C > \left(\int_A^B dl \cos \alpha N \right)_C,$$

т. е.

$$\left(\int_A^B dLN \right)_C > \left(\int_A^B dLN \right)_{\text{луч}}.$$

Физический смысл этого замечательного принципа крайне прост.

Действительно, речь идет о сравнении интегралов $\int_A^B dl/v_\Phi$ для истинного луча и для произвольной кривой. Но при заданной частоте света выписанный интеграл пропорционален $\int_A^B dl/\lambda$, где $\lambda = 2\pi v_\Phi/\omega$ — длина волны; иными словами, согласно принципу Ферма,

реальный луч отличается тем, что вдоль него укладывается минимальное (по сравнению с любой другой кривой, соединяющей те же точки) число длин волн.

Но так и должно быть, потому что в этом случае мы получаем положительную интерференцию, или главный максимум в волновой оптике.

Можно поставить вопрос о форме луча в произвольной неоднородной среде, когда показатель преломления N является произвольной функцией радиуса-вектора \mathbf{r} . Для этого нужно использовать уравнение луча $N(d\mathbf{r}/ds) = \text{grad}L$, преобразовав его так, чтобы в него входила не функция $L(\mathbf{r})$, определяющая волновые поверхности $L(\mathbf{r}) = \text{const}$, а только показатель преломления $N(\mathbf{r})$. Продифференцируем с этой целью уравнение луча по s . Учитывая, что

$$\frac{d}{ds} \left(N \frac{d\mathbf{r}}{ds} \right) = \frac{d}{ds} \nabla L = \frac{d\mathbf{r}}{ds} \nabla(\nabla L) = \frac{1}{N} \nabla L \nabla(\nabla L) = \frac{1}{2N} \nabla(\nabla L)^2,$$

и используя уравнение эйконала $(\nabla L)^2 = N^2$, получим дифференциальное уравнение для определения формы луча:

$$\frac{d}{ds} \left(N \frac{d\mathbf{r}}{ds} \right) = \text{grad} N.$$

Из этого уравнения, в частности, видно, что в однородной среде, т. е. при $N = \text{const}$, лучи — прямые линии: $d\mathbf{r}/ds = \text{const}$.

Это уравнение можно переписать в другой форме с помощью используемой в дифференциальной геометрии тройки ортов (единичных векторов касательной $\tau = dr/ds$, главной нормали $\mathbf{a} = R \frac{d\tau}{ds}$ и би-нормали \mathbf{b} , здесь R — радиус кривизны). Раскрывая производную в левой части уравнения и учитывая, что $dN/ds = \tau \nabla N$, получим

$$\frac{\mathbf{a}}{R} = \frac{1}{N} \{ \nabla N - \tau (\tau \nabla N) \}.$$

Из такой формы уравнения луча вытекает важный вывод о том, что в неоднородной среде луч изгибается в сторону увеличения показателя преломления.

В самом деле, умножая обе стороны на орт \mathbf{a} и учитывая, что $\mathbf{a} \perp \tau$ и $a^2 = 1$, имеем

$$\frac{1}{R} = \frac{1}{N} \mathbf{a} \nabla N.$$

20.4. Теория оптических отображений

Вектор $\mathbf{n}N$, равный градиенту скалярной величины — эйконала, является, очевидно, безвихревым:

$$\text{rot } \mathbf{n}N = 0,$$

отсюда (как в электростатике для электрического поля) может быть получено граничное условие для этого вектора при переходе из одной среды (1) в другую (2). Тангенциальная составляющая вектора $\mathbf{n}N$ должна быть непрерывна при переходе через поверхность раздела, или, что то же самое, вектор $(\mathbf{n}_2 N_2 - \mathbf{n}_1 N_1)$ должен быть перпендикулярен поверхности раздела сред. Обозначив \mathbf{n}_{12} единичный вектор нормали к поверхности, имеем, следовательно,

$$[\mathbf{n}_{12} (\mathbf{n}_2 N_2 - \mathbf{n}_1 N_1)] = 0 \text{ или } N_2 [\mathbf{n}_{12} \mathbf{n}_2] = N_1 [\mathbf{n}_{12} \mathbf{n}_1],$$

откуда

$$N_2 \sin \theta_2 = N_1 \sin \theta_1,$$

где θ_1 и θ_2 — углы, образуемые падающим и преломленным лучами с нормалью к поверхности раздела (углы падения и преломления). Иными словами,

отношение синусов углов преломления и падения равно отношению N_1/N_2 . Кроме того, ясно, что векторы $\mathbf{n}_1 N_1$, $\mathbf{n}_2 N_2$ и \mathbf{n}_{12} лежат в одной плоскости.

Отметим, наконец, что

при отражении луча от поверхности углы падения и отражения должны быть равны.

Полученные граничные условия совпадают формально с граничными условиями для плоской волны, претерпевающей отражение и преломление на плоской границе, разделяющей две однородные среды. Сделанные выше выводы являются, однако, более общими, так как

они справедливы для лучей любой формы, а не только для прямолинейных лучей;

необходимо лишь, чтобы выполнялось условие применимости геометрической оптики, т. е. условие малости длины волны.

Граничные условия для вектора nN лежат в основе теории оптических приборов (линз, зеркал и их различных комбинаций), элементы которой мы здесь изложим. Основным при этом является вопрос о нахождении отображений различных предметов с помощью лучей, проходящих через *оптические системы*, т. е. прозрачные тела, преломляющие и отражающие световые лучи.

Дадим сперва определение *оптического изображения*. Рассмотрим распространение света от точечного источника P_0 , находящегося в среде с показателем преломления $N(\mathbf{r})$. Из P_0 выходит бесконечное число лучей, но через любую другую точку их проходит только конечное число. Может, однако, существовать и такая точка P_1 , в которую придет бесконечное число лучей; такую точку называют *стигматическим* или *резким изображением* точки P_0 . *Идеальный оптический прибор* (система) для каждой точки P_0 трехмерной области, называемой *пространством предмета*, дает стигматическое изображение P_1 . Совокупность точек P_1 называют *пространством изображения*, а точки P_0 и P_1 — *сопряженными*. Если P_0 описывает некоторую кривую C_0 , то P_1 опишет сопряженную кривую C_1 . Эти кривые не обязательно подобны друг другу, но если такое подобие имеет место для любой кривой C_0 , то говорят об *идеальном отображении областей*.

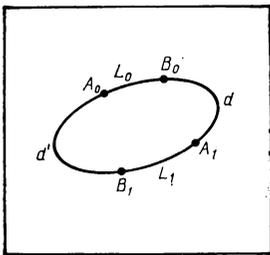
Оптическую систему, дающую стигматическое (но не обязательно идеальное) изображение трехмерного предмета, называют *абсолютным прибором*.

Для абсолютного прибора справедлива следующая теорема, принадлежащая Максвеллу:

в таком приборе оптическая длина любой кривой C_0 в пространстве предмета равна оптической длине изображения C_1 этой кривой:

$$\int_{C_0} N_0 ds_0 = \int_{C_1} N_1 ds_1,$$

где индексы 0 и 1 относятся соответственно к пространствам предмета и изображения.



20.3.
К доказательству теоремы Максвелла для абсолютного прибора

Чтобы убедиться в справедливости этой теоремы, рассмотрим две пары сопряженных точек (A_0, A_1) и (B_0, B_1) (рис. 20.3). Согласно предположению, луч A_0B_0 должен пройти через точки A_1 и B_1 и луч B_0A_0 также должен пройти через эти точки. Оптические пути — это контурные интегралы от безвихревого вектора nN , поэтому они не зависят от формы пути (принцип равного оптического пути) и, следовательно, $L(A_0A_1) = L(A_1A_0)$. Вводя обозначения

$$L(A_0B_0) = L_0, \quad L(A_1B_1) = L_1, \quad L(B_0A_1) = d,$$

$$L(B_1A_0) = d',$$

имеем

$$L_0 + d = d' + L_1, \quad d + L_1 = L_0 + d',$$

откуда сразу следует теорема Максвелла: $L_0 = L_1$. Остается лишь заметить, что до сих пор кривая считалась частью луча; чтобы обобщить теорему на случай произвольной кривой, достаточно заменить ее ломаной, состоящей из отрезков лучей, а затем устремить число этих отрезков к бесконечности.

Из теоремы Максвелла вытекает важное следствие. Рассмотрим бесконечно малый треугольник со сторонами $ds_0^{(1)}$, $ds_0^{(2)}$ и $ds_0^{(3)}$. Тогда согласно этой теореме

$$N_1 ds_1^{(1)} = N_0 ds_0^{(1)}, \quad N_1 ds_1^{(2)} = N_0 ds_0^{(2)}, \quad N_1 ds_1^{(3)} = N_0 ds_0^{(3)},$$

где $ds_1^{(1)}$, $ds_1^{(2)}$, $ds_1^{(3)}$ — изображения этих сторон и N_0 , N_1 — показатели преломления. Значит, треугольник и его изображение в абсолютном приборе подобны друг другу, или, что то же самое, их соответственные углы равны.

Образования, сохраняющие углы, называют конформными, причем существует общая теорема, согласно которой конформное отображение трехмерного пространства может быть только двух видов — либо это проективное преобразование, либо инверсия, либо инверсия. Поэтому отображение, создаваемое абсолютным прибором, представляет собой либо проективное преобразование, либо инверсию, либо их комбинацию.

Прежде чем разъяснять, что означают эти преобразования, рассмотрим тот случай, когда отображение является не только стигматическим, но и идеальным, т. е. изображение геометрически подобно исходному предмету. В общем случае стигматического отображения подобие имеет место только в бесконечно малых областях пространства, причем коэффициент подобия N_1/N_0 меняется от точки к точке; в случае же идеального отображения должно быть $N_1/N_0 = \text{const}$. В частности, если $N_1 = N_0 = \text{const}$, то идеальное отображение тривиально в том смысле, что изображение полностью конгруэнтно предмету; такое отображение может быть осуществлено только плоским зеркалом или системой плоских зеркал.

Напомним теперь, что понимается под *инверсией* и *проективным преобразованием*. Инверсия относительно точки O , называемой *центром инверсии*, — это преобразование, не меняющее произведения длин отрезков $P_0O \cdot OP_1$ (P_0 , P_1 — пара сопряженных точек). В частности, точкой, сопряженной центру O , является бесконечно удаленная точка.

Проективное преобразование, или *коллинеация*, переводит точку P с координатами x, y, z в точку P' с координатами

$$x' = F_1/F_0, \quad y' = F_2/F_0, \quad z' = F_3/F_0,$$

где F_i — линейные функции координат x, y, z , т. е.

$$F_i = a_i x + b_i y + c_i z + d_i$$

(a_i, b_i, c_i, d_i — константы). Если разрешить эти соотношения относительно x, y, z , то мы получим полностью аналогичные выражения:

$$x = F'_1/F_0, \quad y = F'_2/F_0, \quad z = F'_3/F_0,$$

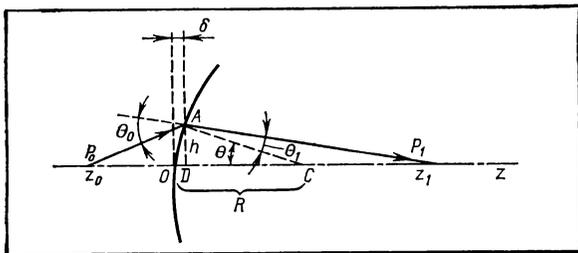
где F'_i — линейные функции x', y', z' (с отличными от a, b, c, \dots коэффициентами a', b', c', d').

Мы видим, что изображение любой точки, лежащей в плоскости $F_0 = 0$, находится на бесконечности и что, наоборот, бесконечно удаленные точки проецируются на плоскость $F'_0 = 0$. Плоскость $F_0 = 0$ называют *фокальной плоскостью пространства предмета*, а плоскость $F'_0 = 0$ — *фокальной плоскостью пространства изображения*. Лучи, параллельные в пространстве предмета, преобразуются в лучи, пересекающиеся в некоторой точке, лежащей в фокальной плоскости $F'_0 = 0$; лучи, выходящие из точки на плоскости $F_0 = 0$, преобразуются в пучок параллельных лучей. Если обе фокальные плоскости находятся на бесконечности, то преобразование называют *аффинным* или *телескопическим*; в этом случае все шесть коэффициентов $a_0, b_0, c_0, a'_0, b'_0, c'_0$ равны нулю.

20.5. Параксиальная оптика

Возникает вопрос: можно ли достичь проективного преобразования с помощью реального оптического прибора, состоящего из совокупности линз и зеркал, т. е. можно ли построить абсолютный оптический прибор? Ответ гласит, что *этого можно добиться только приближенно*, используя узкие пучки лучей, распространяющиеся вблизи оптической оси. Теорию отображения такими лучами называют *параксильной оптикой*.

Разъясним, как возникает в этом случае стигматическое изображение. Для этого рассмотрим ход лучей вблизи оси преломляющей поверхности вращения, которая разделяет две однородные среды с показателями преломления N_0 и N_1 (рис. 20.4). Пусть луч исходит из точки P_0 , лежащей на оси z поверхности, встречает поверхность в точке A , преломляется и пересекает ось в точке P_1 . В треугольнике P_0AC , где C — центр кривизны поверхности вращения, соответствующий точке O , отношение сторон P_0C и P_0A составляет $\sin \theta_0/\sin \theta$; далее, в треугольнике P_1AC отношение сторон P_1C и P_1A равно $\sin \theta_1/\sin \theta$. Мы интересуемся лучами, идущими близко к оси, и поэтому должны считать, что $P_0A \approx P_0O$ и $P_1A \approx P_1O$. Поместив начало координат в точке O и обозначив z_0 и z_1 координаты точек P_0 и P_1 , получим



20.4.

Возникновение стигматического изображения при преломлении

$$\frac{-z_0 + R}{z_0} = \frac{\sin \theta_0}{\sin \theta}, \quad \frac{z_1 - R}{z_1} = \frac{\sin \theta_1}{\sin \theta},$$

где R — радиус кривизны в точке O (заметим, что $z_1 > 0$, $z_0 < 0$). Разделив первое равенство на второе и учитывая, что $\sin \theta_0 / \sin \theta_1 = N_1 / N_0$, получим

$$N_0 \left(\frac{1}{z_0} - \frac{1}{R} \right) = N_1 \left(\frac{1}{z_1} - \frac{1}{R} \right).$$

Это соотношение однозначно определяет координату z_1 точки изображения P_1 как функцию координаты z_0 точки предмета P_0 независимо от выбора точки A , в которой исходящий из P_0 луч пересекает преломляющую поверхность. Иными словами, изображение оказывается стигматическим, но, разумеется, только в *параксиальном приближении*, т. е. в пренебрежении различием между отрезками P_0A , P_1A и отрезками P_0O , P_1O , пропорциональным θ^2 .

Правую и левую части написанного равенства называют *инвариантом Аббе*. Это равенство можно переписать в виде

$$\frac{N_1}{z_1} - \frac{N_0}{z_0} = \Phi, \quad \Phi \equiv \frac{N_1 - N_0}{R}.$$

Величину Φ называют *оптической силой* преломляющей поверхности.

Легко видеть, что если $z_1 \rightarrow \infty$, то $z_0 \rightarrow -N_0/\Phi$; если же $z_0 \rightarrow \infty$, то $z_1 \rightarrow N_1/\Phi$. Таким образом, фокусами оказываются точки на оси с координатами $(-N_0/\Phi)$ и (N_1/Φ) и фокусные расстояния составляют

$$f_0 = \frac{N_0 R}{N_1 - N_0}, \quad |f_1| = \frac{N_1 R}{N_1 - N_0}.$$

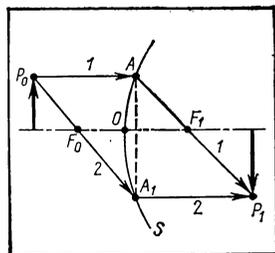
Поэтому

$$f_0/z_0 - |f_1|/z_1 = -1.$$

Используя свойства фокусов, легко получить изображение P_1 точки P_0 , не лежащей на оси преломляющей поверхности S . Для этого нужно из P_0 провести два луча (1 и 2) до пересечения с преломляющей поверхностью — первый, параллельный оси, и второй, проходящий через фокус F_0 . Первый луч после преломления должен пройти через фокус F_1 , а второй — параллельно оси (рис. 20.5). На рис. 20.6 это построение сделано в предположении, что S — плоскость, проходящая через O (это приближение соответствует случаю, когда расстояние от O до AA_1 мало по сравнению с фокусными расстояниями). Из этого построения, рассматривая подобные треугольники $F_1P_1P'_1$ и F_1AO , $F_0P_0P'_0$ и F_0A_1O , легко заключить, что

$$-Y_1/Z_1 = Y_0/|f_1|, \quad Y_0/-Z_0 = Y_1/f_0,$$

где Z_0 , Y_0 и Z_1 , Y_1 — координаты точек P_0 и



20.5.

Построение изображения точки, не лежащей на оси преломляющей поверхности

цательным. Это значит, что отображение диоптрическое ($f_0 f_1 < 0$): если поверхность обращена своей выпуклой стороной к предмету ($R > 0$), то при $N_0 < N_1$ оно сходящееся, а при $N_0 > N_1$ — расходящееся.

Оптическую силу Φ легко выразить через фокусные расстояния:

$$\Phi = N_0/f_0 = -N_1/f_1.$$

Эта формула определяет оптическую силу любой оптической системы при любых знаках f_0 и f_1 . Величина Φ положительна для собирающей ($f_0 > 0$) и отрицательна для рассеивающей системы ($f_0 < 0$). Единицей оптической силы является *диоптрия*, соответствующая оптической силе сферической поверхности с $f_0 = 1$ м при $N_0 = 1$.

Так же как для преломляющей поверхности вращения, легко получить формулу, связывающую координаты z_0 , z_1 точки и ее изображения в случае отражающей поверхности вращения. Она имеет вид

$$\frac{1}{z_0} + \frac{1}{z_1} = \frac{2}{R},$$

где R — радиус кривизны. Последний считается положительным, если поверхность обращена выпуклой стороной к падающему пучку, а отрицательной — если вогнутой.

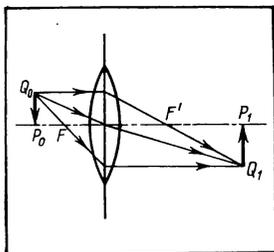
Умея находить изображения для одной преломляющей или отражающей поверхности вращения, можно найти изображения, создаваемые сколь угодно сложной оптической системой, состоящей из многих преломляющих и отражающих (коаксиальных) поверхностей вращения. Для этого следует найти изображение P_1 исходной точки P_0 , создаваемое первой поверхностью; взяв затем точку P_1 в качестве исходной, найти ее изображение P_2 , создаваемое второй поверхностью, и т. д. Общие формулы, получающиеся при этом, очень сложны, и мы ограничимся здесь случаем *тонкой линзы* — оптической системы, состоящей из двух преломляющих поверхностей вращения с общей осью, расположенных так, что расстоянием между ними (толщиной линзы) можно пренебречь. Оптическая сила такой системы равна сумме сил обеих поверхностей:

$$\Phi = \Phi_1 + \Phi_2 = \frac{N_1 - N_0}{R_1} + \frac{N_2 - N_1}{R},$$

где N_0 , N_1 , N_2 — показатели преломления трех сред (перед линзой, внутри нее и позади линзы) и R_1 , R_2 — радиусы кривизны поверхностей линзы. При этом R считается положительным или отрицательным в зависимости от того, выпуклой или вогнутой стороной обращена поверхность по отношению к падающему пучку лучей. Уравнение, определяющее координату z_2 изображения, имеет вид

$$\frac{N_1}{z_1} - \frac{N_0}{z_0} + \frac{N_2}{z_2} - \frac{N_1}{z_1} = \Phi, \quad \frac{N_2}{z_2} - \frac{N_0}{z_0} = \Phi,$$

где z_0 , z_1 , z_2 — координаты точек P_0 , P_1 , P_2 . Если показатели преломления сред по обе стороны линзы одинаковы ($N_0 = N_2$), то



20.7. Построение изображения в тонкой линзе

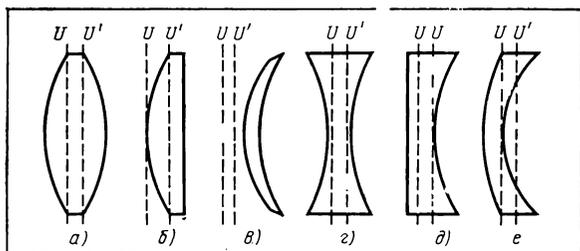
$$\frac{1}{z_2} - \frac{1}{z_0} = (N - 1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right),$$

где $N = N_1/N_0$. Координаты фокусов линзы определяются формулами

$$\frac{1}{z_0} = -(N - 1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right), \quad z_2 \rightarrow \infty;$$

$$\frac{1}{z_2} = (N - 1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right), \quad z_0 \rightarrow \infty,$$

т. е. фокусы находятся на одинаковом расстоянии по обе стороны линзы.



20.8.

Различные типы линз (U, U' — главные плоскости, свет падает слева)

собирающие: двояковыпуклая (а), плосковыпуклая (б), собирательный мениск (в); рассеивающие: двояковогнутая (г), плосковогнутая (д), рассеивающий мениск (е)

На рис. 20.7 показано построение изображения с помощью двояковыпуклой линзы. Линзы различных типов изображены на рис. 20.8.

До сих пор мы рассматривали только параксиальные лучи. Однако роль непараксиальных лучей в оптических приборах весьма существенна — они нарушают точность изображения и создают вместо точечного изображения пятно. Это явление называют *абберацией*. Более того, из-за зависимости показателя преломления от частоты света изображение белой точки имеет вид цветного пятна; это так называемая *хроматическая абберация*.

ОСНОВНЫЕ ФОРМУЛЫ

Уравнение эйконала $\left(\frac{\partial L}{\partial x} \right)^2 + \left(\frac{\partial L}{\partial y} \right)^2 + \left(\frac{\partial L}{\partial z} \right)^2 = N^2$

Оптическая длина дуги C между точками A и B $L_C(AB) = \left(\int_A^B N dl \right)_C$

Уравнение для определения формы луча $\frac{d}{ds} \left(N \frac{dr}{ds} \right) = \text{grad } N$

Связь между координатами точки и ее изображения $\frac{N_1}{z_1} - \frac{N_0}{z_0} = \frac{N_1 - N_0}{R}$

Фокусные расстояния $f_0 = \frac{N_0 R}{N_1 - N_0}, \quad |f_1| = \frac{N_1 R}{N_1 - N_0}$

Оптическая сила

$$\Phi = \frac{N_0}{f_0} = - \frac{N_1}{f_1}$$

Формула линзы

$$\frac{1}{z_2} - \frac{1}{z_0} = (N - 1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right)$$



Глава 21. ДИФРАКЦИЯ

21.1. Разрешающая сила оптических приборов

В предыдущей главе мы оперировали понятиями светового луча и светящейся точки, придавая им буквальный геометрический смысл. В действительности, однако, эти понятия являются идеализацией, связанной с самим характером геометрической оптики, которая справедлива лишь в том случае, когда длина волны значительно меньше всех встречающихся в задаче линейных размеров. Поэтому геометрические понятия луча и точки, строго говоря, относятся к предельному случаю $\lambda \rightarrow 0$.

Вследствие волновой природы света, т. е. конечности λ , не может существовать светящейся точки в геометрическом смысле слова и вместо точки мы всегда имеем дело с некоторым волновым пакетом, т. е. суперпозицией плоских монохроматических волн с различными частотами и волновыми векторами.

В § 16.5 мы рассмотрели волновые пакеты и, в частности, показали, что размеры пакета не могут быть значительно меньше длины волны. Точнее, были получены соотношения

$$\Delta k_x \cdot \Delta x \geq 2\pi, \quad \Delta k_y \cdot \Delta y \geq 2\pi, \quad \Delta k_z \cdot \Delta z \geq 2\pi,$$

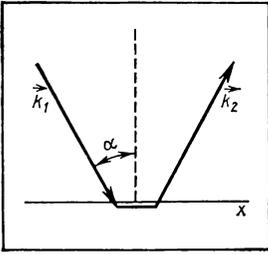
где Δx , Δy , Δz — размеры пакета в направлениях x , y , z и Δk_x , Δk_y , Δk_z — разбросы в пакете по проекциям волнового вектора. Из этих соотношений вытекает, что если мы хотим иметь дело со световой волной, близкой к плоской волне, т. е. хотим, чтобы выполнялись неравенства $\Delta k_x \ll k_x$, $\Delta k_y \ll k_y$, $\Delta k_z \ll k_z$ (\mathbf{k} — среднее значение волнового вектора в пакете), то мы обязательно получим пакет больших размеров.

Аналогичное соотношение было выведено для частотного разброса $\Delta \omega$ пакета:

$$\Delta \omega \cdot \Delta t \geq 2\pi,$$

где Δt — время существования пакета (в определенной точке пространства). Из этого соотношения вытекает ограничение на возможность использовать пакеты, близкие к монохроматическим волнам ($\Delta \omega \ll \ll \omega$); такой пакет существует «почти бесконечное» время и поэтому не может служить несущим информацию сигналом.

Выписанные соотношения находят важное применение при анализе вопроса о разрешающей силе оптических приборов. Рассмотрим, например, вопрос о получении стигматического изображения с по-



21.1.

К выводу формулы разрешающей силы прибора

мощью сходящегося пучка лучей (рис. 21.1). Пусть k_1 и k_2 — волновые векторы крайних лучей, используемых для получения точечного изображения, причем $k_1 = k_2 = k = 2\pi/\lambda$ (λ — длина волны используемого света). Тогда для интервала, в котором лежат x -проекции волновых векторов, образующих пучок, имеем

$$\Delta k_x = k_{1x} - k_{2x} = 2k \sin \alpha = (4\pi/\lambda) \sin \alpha,$$

где α — половина угла между крайними лучами (*угол апертуры*).

При известном Δk_x легко оценить величину Δx — минимальный размер изображения в направлении x :

$$\Delta x \sim \frac{2\pi}{\Delta k_x} \sim \frac{\lambda}{2 \sin \alpha}.$$

Мы видим, что изображение не является строго стигматическим — оно было бы таким только в предельном случае $\lambda \rightarrow 0$. В действительности же размеры «светящейся точки» составляют $\lambda/(2 \sin \alpha)$. Ее минимальные размеры составляют $\sim \lambda/2$ и соответствуют углу апертуры $\alpha = \pi/2$. Таковы и минимальные размеры предметов, которые можно различать приборами, в частности микроскопами (точнее, таковы минимальные расстояния между различаемыми деталями предметов). Для уменьшения этих размеров нужно использовать свет меньшей длины волны;

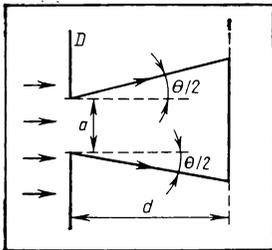
для волн определенной длины нужно стремиться к скользящему падению света на поверхность предмета.

Но в любом случае точки пересечения лучей — это не геометрические точки, а протяженные объекты с размерами порядка $\lambda/2$. Соответственно и лучи — не чисто геометрические линии, а скорее «жирные» линии, имеющие толщину.

В качестве другого примера рассмотрим получение плоскопараллельного пучка лучей конечной ширины (рис. 21.2). Допустим, например, что с помощью диафрагмы D , в которой сделано отверстие размера a , мы хотим вырезать из параллельного пучка лучей узкий строго плоскопараллельный пучок лучей ширины a . Итак, $\Delta x = a$ и, следовательно, $\Delta k_x \geq 2\pi/a$. С другой стороны,

$$\Delta k_x = 2k \sin(\theta/2) \approx 2\pi\theta/\lambda,$$

где θ — угол расхождения пучка. Мы видим, что строгая плоскопараллельность ($\theta = 0$) недостижима: обязательно $\theta \geq \lambda/a$; мы тем больше приблизимся к ней, чем короче волны используемого света.



21.2.

Расхождение плоскопараллельного пучка

Ширина пучка на расстоянии d от отверстия составит

$$a' = a + \theta d = a + \lambda d/a;$$

эта величина минимальна, если $a = \sqrt{\lambda d}$.

Обратим внимание на то, что источником света определяется частота (а не длина волны $\lambda = 2\pi/(\omega N)$, где N — показатель преломления среды); поэтому использование сред с большим показателем преломления позволяет получить от того же источника более коротковолновый свет.

Заметим также, что все наши выводы относятся не только к свету (и другим электромагнитным колебаниям), но и к любому волновому процессу, для которого справедлив принцип суперпозиции, т. е. *возможно построение волновых пакетов*. В следующем параграфе мы остановимся на этом подробнее.

21.2. Оптико-механическая аналогия

Световой луч в геометрической оптике подобен траектории материальной частицы в механике. Не случайно многие физики, в том числе и великий Ньютон, склонялись к корпускулярной, а не волновой теории света, считая его состоящим из отдельных частиц — корпускул. Это одно из проявлений аналогии между оптикой и механикой, которая, как оказывается, имеет очень глубокий физический смысл.

В основе ее лежит аналогия между принципом Ферма и известным из механики принципом Мопертюи. Как мы видели, принцип Ферма утверждает минимальность оптического пути света, т. е. интеграла

$$\int_1^2 \mathbf{k}(\mathbf{r}) d\mathbf{r} \rightarrow \min,$$

взятого вдоль луча, по сравнению с оптическим путем, вычисленным вдоль любой другой кривой, соединяющей те же точки 1 и 2. Принцип Мопертюи утверждает минимальность *интеграла действия*

$$\int_1^2 \mathbf{p}(\mathbf{r}) d\mathbf{r} \rightarrow \min,$$

взятого вдоль истинной траектории частицы, по сравнению с таким же интегралом, взятым вдоль любой другой кривой, соединяющей те же точки 1 и 2. Здесь $\mathbf{k}(\mathbf{r})$ — волновой вектор, связанный с частотой соотношением $\mathbf{k}(\mathbf{r}) = \omega/v_\Phi(\mathbf{r})$; $\mathbf{p}(\mathbf{r})$ — импульс частицы, связанный с ее энергией E соотношением $\mathbf{p}(\mathbf{r}) = \sqrt{2m[E-U(\mathbf{r})]}$, $v_\Phi(\mathbf{r})$ — фазовая скорость волны; $U(\mathbf{r})$ — потенциальная энергия частицы и m — ее масса. Векторы \mathbf{k} и \mathbf{p} направлены по касательным к кривым. В принципе Ферма считается заданной постоянной частота света ω , в принципе Мопертюи — энергия частицы E .

Чтобы убедиться в справедливости принципа Мопертюи, найдем вариацию интеграла при бесконечно малом варьировании кривой, соединяющей точки 1 и 2, и приравняем ее нулю:

$$\delta \int_1^2 \sqrt{2m(E-U)} ds = \sqrt{2m} \int_1^2 \left\{ -\frac{\delta U ds}{2\sqrt{E-U}} + \sqrt{E-U} \delta ds \right\} = 0,$$

где ds — элемент длины вдоль кривой. Здесь первое слагаемое связано с вариацией потенциальной энергии $\delta U = (\partial U/\partial r)\delta r$, а второе — с вариацией участка пути, причем так как $(\Delta s)^2 = (\Delta r)^2$ для малых участков пути Δs и малых Δr , то $\delta ds = (dr/ds)\delta dr$. Интегрируя второе слагаемое по частям и учитывая произвольность вариации δr , а также закрепленность точек 1 и 2, получим

$$\frac{d}{ds} \left(\sqrt{E-U} \frac{dr}{ds} \right) = -\frac{1}{2} \frac{1}{\sqrt{E-U}} \frac{\partial U}{\partial r}.$$

Заметим теперь, что величина $v = \sqrt{(2/m)(E-U)}$ представляет собой скорость частицы, а $ds/v = dt$ есть время, за которое частица пройдет путь ds . Поэтому мы можем переписать последнее равенство в виде

$$m \frac{d^2 r}{dt^2} = -\frac{\partial U}{\partial r}.$$

Таким образом, из принципа Мопертюи мы получили правильный закон движения частицы.

Из аналогии между принципом Ферма и принципом Мопертюи вытекает аналогия между волновым вектором $\mathbf{k} \equiv \mathbf{k}(\mathbf{r})$ в оптике и импульсом частицы $\mathbf{p} \equiv \mathbf{p}(\mathbf{r})$ в механике ($\mathbf{k} \rightarrow \mathbf{p}$) и, следовательно, аналогия между лучом (касающимся в каждой точке \mathbf{r} вектора $\mathbf{k}(\mathbf{r})$) и траекторией (касающейся в каждой точке \mathbf{r} вектора $\mathbf{p}(\mathbf{r})$).

Из аналогии между принципами Ферма и Мопертюи следует далее, что показателю преломления (или, что эквивалентно, фазовой скорости) как функции координат соответствует в механике кинетическая энергия $T \equiv E - U$. Если эти величины от координат не зависят, то и траектория, и луч — прямые.

Основным понятием механики является понятие движущейся материальной точки. В оптике, как мы видели, из-за волновой природы света невозможно движение точечных световых сигналов, так что аналогом материальной точки приходится считать волновой пакет. Поэтому скорость v частицы в механике должна сопоставляться групповой скорости v_g в оптике:

$$v_g \rightarrow v.$$

Вспоминая, что $v_g = d\omega/dk$ и $v = dE/dp$, мы приходим к аналогии между частотой в оптике и энергией частицы в механике:

$$\omega \rightarrow E.$$

Напомним в связи с этим, что в принципах Ферма и Мопертюи заданными считаются как раз эти величины — частота и энергия.

Таким образом, можно провести аналогию между основными понятиями геометрической оптики и механики материальной частицы.

Но геометрическая оптика — это предельный (при $\lambda \rightarrow 0$) случай более общей волновой оптики. Что же соответствует волновой оптике в механике? Ответ гласит: квантовая механика. (Напомним, что в эпоху создания этой науки едва ли реже термина «квантовая механика» применялся термин «волновая механика».)

Поясним, с какими понятиями оперирует квантовая механика. Для этого оптико-механической аналогии следует придать точную количественную формулировку. Эта формулировка гласит:

частице с энергией E и импульсом p сопоставляется волна с частотой ω и волновым вектором k так, что

$$E = \hbar\omega, \quad p = \hbar k,$$

где $\hbar = 1,0546 \cdot 10^{-34}$ Дж·с — универсальная постоянная, называемая *квантовой постоянной* или *постоянной Планка*. Отсюда следует выражение

$$\lambda = 2\pi\hbar/p$$

для длины волны и выражение $v_\phi = \omega\lambda/(2\pi) = \hbar\omega/p$ для фазовой скорости волны, связанной с частицей. Величину λ называют *длиной волны де Бройля*.

Теперь можно, не вникая поначалу в смысл волновой функции $\psi(\mathbf{r}, t)$, связанной с частицей, написать для нее по аналогии с волновой оптикой волновое уравнение

$$\Delta\psi - \frac{1}{v_\phi^2} \frac{\partial^2\psi}{\partial t^2} = 0.$$

Так как речь идет о монохроматической волне

$$\psi(\mathbf{r}, t) = \psi(\mathbf{r}) e^{-i\omega t}, \quad \omega = E/\hbar,$$

то волновое уравнение можно переписать в виде $\Delta\psi + (\omega^2/v_\phi^2)\psi = 0$, а так как $\omega^2/v_\phi^2 = p^2/\hbar^2$ и $p^2 = 2m(E - U)$, где U — потенциальная энергия частицы, то волновое уравнение окончательно приобретает вид

$$-\frac{\hbar^2}{2m} \Delta\psi + U(\mathbf{r})\psi = E\psi.$$

Это и есть уравнение, определяющее волновую функцию частицы с энергией E , движущейся в поле с потенциальной энергией $U(\mathbf{r})$. Его называют уравнением Шредингера.

Важнейшим свойством этого уравнения является то, что при заданной потенциальной энергии $U(\mathbf{r})$ оно допускает решение не при всех, а только при вполне определенных значениях энергии E . Этот набор энергий, называемый по аналогии с оптикой *спектром энергии*, может быть как дискретным, так и непрерывным (ср. с набором час-

тот волновода или резонатора, § 14.5, 14.6), причем каждому допустимому значению энергии соответствует своя волновая функция ψ . В частности, если в качестве U взять потенциальную энергию $U = -e^2/(4\pi\epsilon_0 r)$ электрона в кулоновском поле протона, то мы получим энергетический спектр атома водорода.

Волновая функция ψ имеет вероятностный смысл, но мы не будем разбирать здесь этот вопрос, а остановимся лишь на соотношениях, аналогичных соотношениям $\Delta x \cdot \Delta k_x \geq 2\pi$, $\Delta t \cdot \Delta \omega \geq 2\pi$ для волнового пакета.

Так как $k = p/\hbar$ и $\omega = E/\hbar$, то мы приходим к соотношениям

$$\Delta x \cdot \Delta p_x \geq 2\pi\hbar, \quad \Delta y \cdot \Delta p_y \geq 2\pi\hbar, \quad \Delta z \cdot \Delta p_z \geq 2\pi\hbar,$$

$$\Delta t \cdot \Delta E \geq 2\pi\hbar.$$

Это так называемые соотношения неопределенностей Гейзенберга. Смысл их заключается в следующем:

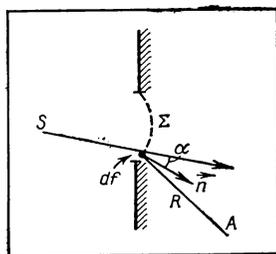
если частица локализована в интервале Δx по оси x , то проекция ее импульса на эту ось не может иметь определенного значения, а лежит в некотором интервале Δp_x , таком, что $\Delta x \cdot \Delta p_x \geq 2\pi\hbar$ (и аналогично для осей y и z), если частица находится в некотором нестационарном состоянии в течение конечного времени Δt , то это состояние не может быть строго моноэнергетическим, причем мера немонотонности ΔE связана с Δt соотношением $\Delta t \cdot \Delta E \geq 2\pi\hbar$.

Оптико-механическая аналогия позволяет без труда проанализировать разрешающую способность аналогов оптических приборов, использующих вместо света другие частицы, — достаточно в соотношения предыдущего параграфа подставить длину волны де Бройля соответствующей частицы вместо длины световой волны. Например, в электронном микроскопе используются электроны, соответствующие с точки зрения длины волны рентгеновским лучам ($\lambda \sim 10^{-10}$ м); такой микроскоп позволяет увидеть объекты, размеры которых в несколько раз превосходят размеры атомов. Еще меньшие объекты, например атомные ядра (размеры их составляют от 10^{-15} м и меньше), наблюдают с помощью ускорителей заряженных частиц, в которых получают частицы высоких энергий. Чем выше энергия, тем больше импульс частиц, т. е. меньше длина волны де Бройля и, следовательно, тем меньшие объекты можно «увидеть» с помощью такого прибора.

21.3. Принцип Гюйгенса и дифракция Френеля

Геометрическая оптика, как мы знаем, справедлива в предельном случае исчезающе малой длины волны: $\lambda \rightarrow 0$. При очень малых, но конечных λ возникают отклонения от геометрической оптики, которые носят общее название *дифракции*. Если, например, на пути распространения света стоит непрозрачный экран, то согласно геометрической оптике за экраном должна возникать резкая граница между светом и тенью. В действительности, однако, такая граница не возникает и свет проникает в область тени, а в области света появляются участки с ослабленной освещенностью.

Дифракционные явления при малых λ могут быть исследованы общим методом, который мы здесь рассмотрим. Он основывается на принципе Гюйгенса, заключающемся в следующем. Пусть на пути распространения света от источника S находится непрозрачный экран с отверстием и мы хотим найти поле (электрическое или магнитное) в некоторой точке A за экраном (рис. 21.3). Тогда, согласно принципу Гюйгенса, нужно представить себе, что отверстие покрыто некоторой произвольной поверхностью Σ , и считать эту поверхность источником вторичных волн, распространяющихся за экраном, в частности в точку A . Что же касается истинного источника S , то он только создает вторичные источники на поверхности Σ , но сам непосредственно не вносит вклада в поле в точке A .



21.3.
Принцип Гюйгенса

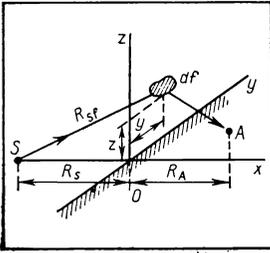
Вторичным источником является каждый элемент df поверхности Σ . Поле, создаваемое df , пропорционально полю $ue^{i\omega t}$, создаваемому истинным источником S на элементе поверхности df в том случае, если бы никакого экрана не было. Кроме того, поле, создаваемое df , пропорционально $df \cos \alpha$, где α — угол между нормалью \mathbf{n} к df и направлением падающего на df луча (в частности, если площадка параллельна лучу, но луч «скользит» вдоль нее, не оказывая на нее действия).

Наконец, чтобы найти поле в точке A на расстоянии R от df , нужно учесть изменение амплитуды, обратно пропорциональной R , и эффект запаздывания. Последний заключается в том, что для нахождения поля в точке A в момент времени t нужно взять поле на площадке df в момент времени $t - R/c$. Так как поле предполагается монохроматическим, то окончательно поле, порождаемое df в точке A , можно записать в виде $(au/R)\exp[-i\omega t - R/c]$, где a — некоторая константа. Складывая поля в точке A , создаваемые всеми элементами df поверхности Σ , найдем результирующее поле:

$$u_A = a \int_{\Sigma} \frac{u}{R} e^{ikR} \cos \alpha \, df,$$

где $k = \omega/c$ (вектор $\exp(-i\omega t)$ опущен). Это и есть математическая формулировка принципа Гюйгенса. Мы не будем вычислять постоянную a , значение которой несущественно для дальнейшего (она равна $k/(2\pi i)$).

Начнем рассмотрение дифракционных явлений с простейшего случая — так называемой *дифракции Френеля*, когда и источник, и точка наблюдения находятся на конечных (а не бесконечных!) расстояниях от экрана. Так как речь идет о малых отклонениях от геометрической оптики, то главную роль играют те участки поверхности Σ , которые расположены вблизи прямой, соединяющей S и A . Иными словами, существенную роль играют лишь небольшие участки края экрана,



21.4.
Дифракция Френеля

которые можно считать прямолинейными. Наконец, существенны только фазовые, а не амплитудные изменения полей. Это значит, что в случае дифракции Френеля поле в точке A можно записать в виде

$$u_A = \text{const} \int_{\Sigma} \exp [ik(R_{Sf} + R_{fA})] \cdot df,$$

где R_{Sf} и $R_{fA} \equiv R$ — расстояния от источника S до элемента поверхности df и от этого элемента до точки наблюдения A . (Множитель $1/R$, входящий в формулировку принципа Гюйгенса, а также амплитуда поля, создаваемого источником, в том числе и $\cos \alpha$, включены в const .)

Пусть экран совпадает с нижней полуплоскостью yz , а поверхность Σ — с верхней полуплоскостью yz , так что ось y направлена по краю экрана (на рис. 21.4 экран заштрихован). Пусть далее ось x направлена по нормали к плоскости экрана и на ней на расстоянии R_S слева от начала координат O находится источник S , координаты же точки наблюдения A равны $R_A, 0, d$. Тогда $R_{Sf} = \sqrt{R_S^2 + y^2 + z^2}$, $R_{fA} = \sqrt{R_A^2 + y^2 + (z-d)^2}$, где (y, z) — координаты участка df . Так как существенны, как уже указывалось, только близкие к точке O участки df , то эти выражения можно упростить, считая величины $y, z, (z-d)$ малыми по сравнению с R_A и R_S :

$$R_{Sf} = R_S + \frac{y^2 + z^2}{2R_S}; \quad R_{fA} = R_A + \frac{y^2 + (z-d)^2}{2R_A}.$$

Подставляя эти выражения в формулу для u_A и учитывая, что $df = dydz$, мы получим произведение двух интегралов — одного по y и другого по z . Нас интересует зависимость поля u_A от координаты d точки наблюдения A . Эта зависимость содержится лишь в интеграле по z , только который мы поэтому и выпишем:

$$u_A = \text{const} \int_0^{\infty} \exp \left\{ \frac{ik}{2} \left[\frac{z^2}{R_S} + \frac{(z-d)^2}{R_A} \right] \right\} \cdot dz.$$

Нижний предел интегрирования равен, очевидно, $z = 0$; верхний же мы положили равным ∞ , так как область больших z все равно несущественна. Вводя вместо z новую переменную

$$\eta = \sqrt{\frac{k(R_A + R_S)}{2R_A R_S}} \left(z - \frac{R_S d}{R_A + R_S} \right),$$

легко преобразовать этот интеграл к виду

$$u_A = \text{const} \int_{-w}^{\infty} e^{i\eta^2} d\eta, \quad w = d \sqrt{\frac{kR_S}{2R_A(R_A + R_S)}}.$$

Последний интеграл носит название *интеграла Френеля*.

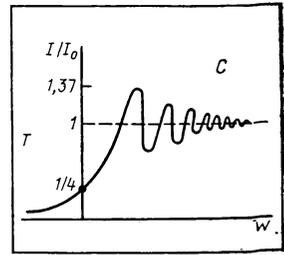
Так как ω содержит в виде множителя $k^{1/2} \sim \lambda^{-1/2}$, а мы рассматриваем область очень коротких волн, то $|\omega| \gg 1$. При этом следует различать два случая: $d > 0$, т. е. $\omega > 0$, и $d < 0$, т. е. $\omega < 0$. В первом случае точка A лежит в освещенной области, а во втором — в области тени. В первом случае

$$\int_{-\omega}^{\infty} e^{i\eta^2} d\eta \approx (1+i) \sqrt{\frac{\pi}{2}} + \frac{1}{2i\omega} e^{i\omega^2} \quad (\omega > 0),$$

во втором же случае

$$\int_{-\omega}^{\infty} e^{i\eta^2} d\eta \approx \frac{e^{i\omega^2}}{2i\omega} \quad (\omega < 0).$$

Квадрат модуля интеграла Френеля определяет интенсивность I света. Графически зависимость интенсивности от ω представлена на рис. 21.5 (I_0 — интенсивность вдали от экрана; области света и геометрической тени обозначены соответственно C и T). Мы видим, что интенсивность не спадает резко до нуля в области геометрической тени; напротив, этот спад происходит монотонно. В освещенной же области интенсивность не всюду равна I_0 ; вблизи экрана она в четыре раза меньше I_0 , затем становится больше I_0 и, постепенно осциллируя, стремится к I_0 .

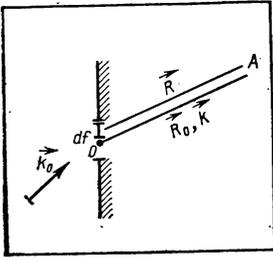


21.5. Интенсивность при дифракции Френеля от края экрана

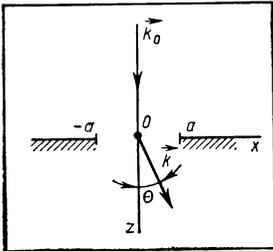
21.4. Дифракция Фраунгофера

В предыдущем параграфе, рассматривая дифракцию Френеля, мы предполагали, что источник света S и точка наблюдения A находятся на конечных расстояниях от экрана. Рассмотрим теперь дифракционные явления в том случае, когда и S , и A находятся на бесконечности. Практически это означает, что источник света помещен в одном из фокусов оптической системы, пройдя через которую свет выходит в виде плоскопараллельного пучка, падающего на экран с одним или несколькими отверстиями. Направление распространения света, прошедшего через отверстия в экране, отличается от направления распространения падающего пучка. В этом и заключается явление дифракции. Таким образом, речь идет о дифракции около экрана направлений световых лучей; такую дифракцию называют *дифракцией Фраунгофера*.

Найдем поле дифрагировавшего света. Для этого по-прежнему будем исходить из принципа Гюйгенса, взяв в качестве поверхности Σ плоскость, закрывающую отверстие. Выбирая некоторую точку O на этой плоскости в качестве начала координат (рис. 21.6) и обозначая \mathbf{r} радиус-вектор, проведенный из O в элемент поверхности df , можно



21.6. Дифракция Фраунгофера



21.7. Дифракция на длинной щели

представить поле падающего света в точке \mathbf{r} поверхности Σ в виде $u = U \exp(i\mathbf{k}_0 \mathbf{r})$, где \mathbf{k}_0 — волновой вектор падающей волны (плоской по определению) и U — ее амплитуда в точке O .

Для того чтобы применить принцип Гюйгенса, мы должны теперь найти $\exp(i\mathbf{k}R)$. Учтем, что точка наблюдения A находится очень далеко от экрана, поэтому направление дифрагировавшего луча совпадает с направлением радиуса-вектора \mathbf{R}_0 этой точки; иными словами, волновой вектор дифрагировавшего света \mathbf{k} направлен по \mathbf{R}_0 . Далее, $\mathbf{R} = \mathbf{R}_0 - \mathbf{r}$, причем $r \ll R_0$, так что $R = \sqrt{(\mathbf{R} - \mathbf{r})^2} \approx R_0 - \mathbf{r}\mathbf{R}_0/R_0 = R_0 - \mathbf{r}\mathbf{k}/k$ и, следовательно,

$$\exp(i\mathbf{k}R) = \exp(i\mathbf{k}_0 R_0) \cdot \exp(-i\mathbf{k}\mathbf{r}).$$

Подставляя это выражение и выражение для поля u на поверхности Σ в интеграл, входящий в формулировку принципа Гюйгенса, получим

$$u = \frac{a \exp(i\mathbf{k}_0 R_0 \cos \alpha)}{R_0} \int_{\Sigma} \exp[i(\mathbf{k}_0 - \mathbf{k}) \mathbf{r}] df$$

(α — по-прежнему угол падения луча на площадку df ; мы заменили в знаменателе R на R_0 и вынесли $1/R_0$ из-под интеграла).

В экспоненту под знаком интеграла входит разность волновых векторов падающей и дифрагировавшей волн. Эти векторы по модулю одинаковы и равны ω/c . Дифракция направлений лучей определяется зависимостью u_A от $(\mathbf{k}_0 - \mathbf{k})$, мы запишем поэтому u_A в виде

$$u_A = \text{const} \int_{\Sigma} \exp[i(\mathbf{k}_0 - \mathbf{k}) \mathbf{r}] df.$$

Заметим, что в отличие от дифракции Френеля, в которой был существен вклад лишь небольшого участка поверхности Σ , при дифракции Фраунгофера существенна вся эта поверхность.

Рассмотрим несколько конкретных примеров. Начнем с дифракции на бесконечно длинной щели с параллельными границами в непрозрачном экране (рис. 21.7). Волновой вектор \mathbf{k}_0 падающей волны направлен нормально к плоскости экрана (x, y); волновой вектор \mathbf{k} дифрагировавшего луча лежит, как ясно из соображений симметрии, в плоскости (z, x), образуя с \mathbf{k}_0 угол θ (угол дифракции). Если ширина щели равна $2a$, то

$$u_A = \text{const} \int_{-a}^a \exp(-ik\theta x) dx$$

(мы учли, что $k_x = k \sin \theta \sim k \theta$, и включили в const интеграл по y , равный длине щели). Выполняя интегрирование, получим

$$u_A = \text{const} \frac{\sin k \theta a}{k \theta}.$$

Интенсивность дифрагировавшей в интервал углов $d\theta$ волны пропорциональна $|u_A|^2$, $dI = \text{const} |u_A|^2 d\theta$. Найдем коэффициент пропорциональности. Для этого проинтегрируем dI по углу θ и приравняем полученную суммарную интенсивность интенсивности I_0 света, падающего на щель; в результате мы получим $\text{const} = I_0 k / (\pi a)$. Таким образом,

$$dI = \frac{I_0}{\pi a k} \frac{\sin^2(k \theta a)}{\theta^2} d\theta.$$

Зависимость интенсивности дифрагировавшего луча от угла дифракции показана на рис. 21.8, где введено обозначение $\eta = k a \theta$. Мы видим, что интенсивность имеет ряд быстро убывающих максимумов, разделенных минимумами при $\theta = \pi n / (ka)$ ($n = \pm 1, \pm 2, \dots$), в которых она обращается в нуль.

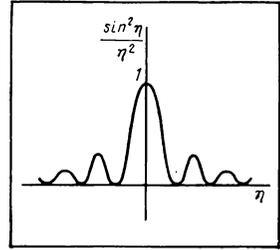
Установив на опыте положение максимумов или минимумов, можно, в принципе, найти величину $ka = 2\pi a / \lambda$; иными словами, зная a , можно найти длину волны λ . Однако точность определения длины волны по дифракции на одной щели невелика, так как максимумы и минимумы получаются не особенно резкими.

Значительно более резкая дифракционная картина получается при дифракции не на одной щели, а на совокупности одинаковых параллельных щелей, расположенных на равных расстояниях друг от друга. Такая совокупность щелей носит название *дифракционной решетки* (рис. 21.9). В этом случае под поверхностью Σ следует понимать совокупность плоских поверхностей, закрывающих все отверстия, а так как все они одинаковы и $\theta \ll 1$, то

$$u_A = u_A^{(1)} \sum_{n=0}^{N-1} \exp(-2ink \theta d) = u_A^{(1)} \frac{1 - \exp[-2i(N-1)k \theta d]}{1 - \exp(2ik \theta d)},$$

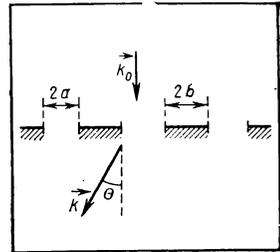
где $u_A^{(1)}$ — поле в точке A , создаваемое одной щелью, N — число щелей и $d = a + b$. Входящая сюда сумма определяет интерференцию лучей, дифрагировавших у отдельных щелей.

Интенсивность света, дифрагировавшего в интервал углов $d\theta$, пропорциональна $|u_A|^2$; коэффициент пропорциональности найдем, приравняв I_0 суммарную интенсивность, создаваемую всеми щелями. В результате получим



21.8.

Угловое распределение интенсивности при дифракции на длинной щели



21.9.

Дифракционная решетка

$$dI = \frac{I}{\pi N a k} \frac{\sin^2(ka\theta)}{\theta^2} \left(\frac{\sin Nk\theta d}{\sin k\theta d} \right)^2 d\theta.$$

Первый множитель здесь определяет распределение интенсивности, соответствующее одной щели, а второй обусловлен наличием решетки. Именно он и приводит к возникновению при углах дифракции $\theta_n = \pi n/(kd)$ (здесь $n = 0, \pm 1, \dots$) очень резких максимумов. Действительно функция $dI/d\theta$ имеет минимумы (обращается в нуль) при $\theta = \pi m/(Nkd)$, где $m = \pm 1, \pm 2, \dots$. Например, за первым максимумом при $\theta = 0$ следует первый минимум при $\theta = \pi/(Nkd)$. Но этот угол очень мал, если только число щелей N велико. Отсюда можно сделать вывод, что ширина максимума имеет порядок $\pi/(Nkd)$.

Вблизи максимумов, т. е. при $\theta = \pi n/(kd) + \vartheta$, $|\vartheta| \ll \pi n/(kd)$, интенсивность имеет вид

$$dI = \frac{I_0}{\pi a k} \left(\frac{kd}{\pi n} \right)^2 \sin^2 \left(\frac{\pi n a}{d} \right) \frac{1}{N} \left(\frac{\sin Nk\vartheta d}{\sin \vartheta k d} \right)^2 d\vartheta.$$

Произведение двух последних множителей имеет при большом N резкий максимум при $\vartheta = 0$. В пределе $N \rightarrow \infty$ ширина максимума вообще обращается в нуль, а высота — в бесконечность, так как

$$\lim_{N \rightarrow \infty} \frac{\sin^2 Nx}{\pi N x^2} = \delta(x).$$

С помощью этой формулы легко вычислить интенсивность света в n -м максимуме:

$$I^{(n)} = I_0 \frac{d}{\pi^2 a} \frac{\sin^2(\pi n a/d)}{n^2}.$$

Интенсивность падает, как мы видим, с ростом номера максимума, как n^{-2} .

Резкость дифракционной картины, создаваемой решеткой с большим числом щелей, позволяет измерять длину волны света с очень большой степенью точности.

Дифракция на двух щелях применяется для измерения угловых диаметров звезд и угловых расстояний между двойными звездами. В этом случае получаются две дифракционные картины, сдвинутые относительно друг друга на угол φ , равный углу между крайними лучами, идущими от источника (или между лучами, идущими от двух близких источников). Изменяя расстояние между серединами щелей d , можно добиться изменения расстояния между максимумами в каждой из дифракционных картин. Обычно, изменяя d , добиваются того, чтобы максимумы одной картины совпали с минимумами другой; при этом $\varphi = \lambda/(2d)$.

Мы рассмотрели дифракцию на системе параллельных щелей. Аналогичным образом можно рассмотреть дифракцию на двумерной решетке одинаковых и одинаково расположенных прямоугольных отверстий в непрозрачном экране. В этом случае нужно найти дифракцию от одного такого отверстия и затем учесть интерференцию полей, создаваемых всеми отверстиями. Если радиусы-векторы центров

отверстий равны $(na + mb)$, где \mathbf{a} , \mathbf{b} — постоянные решетки и n , m — целые числа, то возникающий при такой интерференции дополнительный множитель пропорционален, очевидно, $\sum_{n,m} \exp[i(na + mb)\mathbf{k}]$.

Трехмерные дифракционные решетки дает нам сама природа — это кристаллы (впервые примененные с этой целью Лауэ в 1913 г.). В роли рассеивающих отверстий при этом выступают атомы, расположенные в узлах кристаллической решетки. Зная закон рассеяния лучей атомом и учтя интерференцию лучей, рассеянных отдельными атомами, можно найти дифракцию от всего кристалла. Так как постоянная кристаллической решетки $a \sim 10^{-10}$ м, то кристалл оказывается хорошей дифракционной решеткой для электромагнитных волн рентгеновского диапазона ($\lambda \sim 10^{-10}$ м). В самом деле, разность хода двух лучей 1 и 2, отраженных от двух соседних кристаллических плоскостей, равна $2a \sin \theta$ (рис. 21.9); поэтому положительная интерференция будет иметь место при $2d \sin \theta = n\lambda$, где n — целое число (у с л о в и е Б р э г г а — В у л ь ф а).

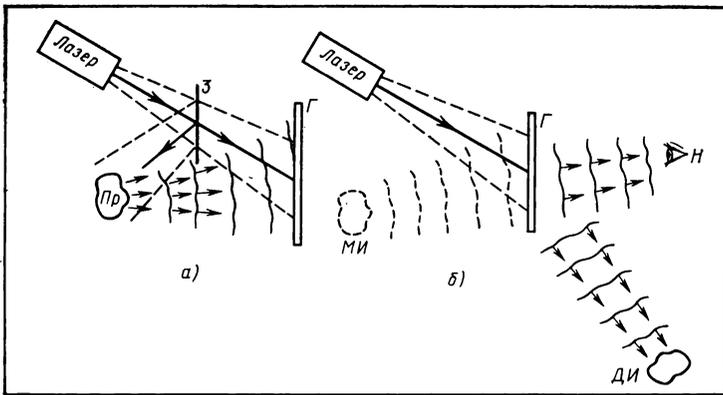
В заключение этого параграфа остановимся на важном свойстве *дополнительных экранов*, т. е. двух экранов с такими отверстиями, что при наложении их друг на друга получается полностью непрозрачный экран. Свойство таких экранов заключается в том, что даваемые ими дифракционные картины совпадают. В частности, например, дифракция от круглого отверстия в непрозрачном экране такая же, как от непрозрачного кружка или шарика такого же радиуса. В обоих случаях интенсивность света, дифрагировавшего в элемент do телесных углов,

$$dI = I_0 \frac{J_1^2(k\theta R)}{\pi\theta^2} d\theta,$$

где R — радиус отверстия или шарика (кружка), θ — угол дифракции, I_0 — полная интенсивность света, падающего на отверстие (шарик, кружок), и $J_1(x)$ — функция Бесселя. Интенсивность имеет главный максимум при $\theta = 0$ и обращается в нуль при $k\theta R \approx 1, 2; 2,2; \dots$ (свет предполагается падающим нормально к плоскости отверстия).

Аналогичной формулой, согласно оптико-механической аналогии, определяется дифракция дебройлевских волн от отверстия или шарика, в частности дифракция или рассеяние быстрых нейтронов на непрозрачном ядре радиуса R . Необходимо лишь, чтобы дебройлевская длина волны нейтрона была малой по сравнению с радиусом ядра R . Изучая рассеяние таких нейтронов ядрами, можно экспериментально находить радиусы тяжелых ядер.

Обратим внимание на то, что согласно той же оптико-механической аналогии имеет место дифракция нейтронов в кристаллах, полностью аналогичная дифракции рентгеновских лучей. Из условия Брэгга—Вульфа следует, что этот эффект существует для медленных нейтронов (со скоростью порядка 10^3 м/с, что соответствует длине волны $\lambda \sim 10^{-10}$ м), в то время как для дифракции на отдельном ядре ($R \sim 10^{-15}$ м) требуются в 10^5 раз более быстрые нейтроны.



21.10.

Схема получения голограммы (а) и восстановления изображения (б)

З — полупрозрачное зеркало, Пр — предмет, Г — голограмма, МИ — мнимое и ДИ — действительное изображения, Н — наблюдатель

21.5. Голография

Дифракция и интерференция лежат в основе так называемого *голографического метода* получения оптических изображений (открытого Д. Габором в 1947 г.). В этом методе записывается не двумерное распределение яркости излучения, рассеянного предметом (как в обычной фотографии), а регистрируются амплитуда и фаза рассеянной световой волны. В результате при восстановлении голограммы получается точная копия записанной волны, которая несет информацию об изображаемом предмете. При этом если предмет непрозрачен, то получается информация о характере его поверхности, в случае же прозрачного предмета дополнительно можно получить информацию о распределении показателя преломления внутри него. С этой целью изображаемый предмет освещается светом от лазера и рассеянная волна попадает на фотопластинку. Вместе с этой волной на пластинку попадает часть света от того же лазера. Эту не испытывавшую рассеяния волну называют *опорной* волной (рис. 21.10, а). Существенно, что рассеянная и опорная волна являются когерентными. В результате записи их интерференции на пластинке создается некоторая картина, которая фиксируется в виде своеобразной, нерегулярной дифракционной решетки. Это и есть голограмма изображаемого предмета.

Если теперь, убрав предмет, осветить голограмму одной только опорной волной, то последняя испытывает дифракцию на голограмме, как на дифракционной решетке. Дифрагировавшая волна совпадает с волной, рассеянной изображаемым предметом, при получении голограммы (так называемой *предметной волной*). Таким образом,

трехмерное изображение предмета можно увидеть, просветив диапозитив только одним опорным пучком света; изображение возникнет за голограммой, если смотреть под углом к опорному лучу (рис. 21.10, б).

Итак, суть голографии заключается в том, что фотопластинка с зарегистрированной на ней картиной интерференции нескольких световых волн освещается только некоторыми из этих волн; тогда в результате дифракции возникают остальные волны.

Проиллюстрируем это следующим образом. Представим себе, что голографическая решетка создается двумя пучками света 1 и 2 (рис. 21.11). Пусть A и B — две соседние дифракционные полосы, находящиеся на расстоянии a друг от друга. Ясно, что $a(\sin\varphi_1 + \sin\varphi_2) = \lambda$, где λ — длина волны. Осветим теперь дифракционную решетку только одним пучком 1. Если β_1 — угол падения света, то ближайший к нему минимум образуется под углом дифракции β_2 , для которого $\sin\beta_1 + \sin\beta_2 = \lambda/a$. В нашем случае $\beta_1 = \varphi_1$; следовательно, $\beta_2 = \varphi_2$. Иными словами, осветив голограмму одной волной (1), мы восстанавливаем участвовавшую в создании голограммы вторую волну (2).

Легко понять, что если бы мы, наоборот, использовали пучок 2, то восстановили бы пучок 1. Это общее свойство любых голограмм обычно формулируют так:

опорный и предметный пучки полностью равноправны.

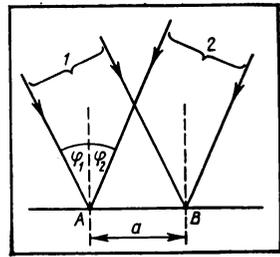
На этом свойстве основан *метод коррекции волновых фронтов*: в нем лазерный пучок со сложным волновым фронтом с помощью изготовленной заранее голограммы (пучок является для голограммы предметным пучком) преобразуется в пучок с плоским волновым фронтом, совпадающим с фронтом того опорного пучка, который использовался при изготовлении голограммы. Дальнейшее развитие этот метод получил при использовании в качестве преобразователя динамической голограммы, отслеживающей изменение во времени волнового фронта преобразуемого лазерного пучка.

В качестве другого примера рассмотрим голограмму точки при падении на нее плоской волны (рис. 21.12). На рис. 21.12, *a* изображены опорный (ОП) и рассеянный (РП) пучки света. Видно, что голограмма точки представляет собой совокупность колец, радиусы которых r_j ($j = 0, 1, \dots$) определяются уравнением

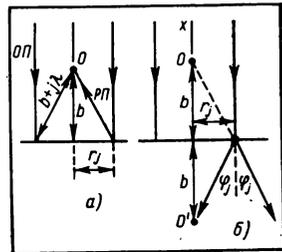
$$r_j^2 = (b + j\lambda)^2 - b^2 = 2jb\lambda + j^2\lambda^2$$

(так называемая *зонная решетка Френеля*).

Пусть теперь на такую решетку падает плоская волна и пусть φ_j обозначает угол дифракции на j -м кольце (рис. 21.12, *б*). Тогда $\Delta r_j \sin\varphi_j = n\lambda$, где $n = 0, \pm 1, \dots$, и $\Delta r_j = (b\lambda + j^2\lambda^2)/r_j$ — расстоя-



21.11. Голографическая решетка от двух когерентных пучков



21.12. Схема получения (а) и восстановления (б) голограммы точки при плоской опорной волне

ние между соседними кольцами. При $n = \pm 1$ дифрагировавшие лучи пересекают ось симметрии x в двух точках $x = \pm b$, так как $x = r_j \operatorname{ctg} \varphi_j = \pm b$. Таким образом, дифрагировавшие волны при $n = \pm 1$ сводятся к расходящейся сферической волне с центром в точке O (при создании голограммы она была рассеивающей точкой) и к сходящейся сферической волне с центром в точке O' , являющейся зеркальным изображением исходной точки O (рис. 21.12). Мы видим, что при освещении голограммы точки опорной волной возникают два изображения точки — мнимое и действительное.

Наряду с уже отмеченным свойством обратимости пучков укажем еще два важных свойства голограммы. Во-первых, если рассеянная каждой точкой предмета волна при записи заполняет всю апертуру голограммы, то каждый участок голограммы способен восстановить изображение всего предмета (при уменьшении участков ослабевает только четкость изображения), поэтому голограмма дает изображение предмета не только трехмерное, но и по сравнению с фотографией более надежное с точки зрения сохранения информации о нем. Во-вторых, если снять с голограммы (как с негатива) контактный отпечаток (аналог позитива), то изображение, восстанавливаемое по копии, тождественно с изображением, восстанавливаемым по оригиналу (это свойство является следствием отмеченного в предыдущем параграфе свойства дополнительных экранов).

Особыми свойствами обладают отражательные голограммы, предложенные Ю. Н. Денисюком. В силу достаточно высокой спектральной селективности их можно восстанавливать некогерентным белым светом. При этом в изображении восстанавливается не только структура предмета, но и информация о его цвете. Голограммы могут быть записаны на самых разнообразных светочувствительных материалах, включая и различные нелинейные среды (см. гл. 22), оптические характеристики которых зависят от интенсивности падающего излучения. Голограммы такого рода называют *динамическими*, поскольку они существуют только при воздействии записывающих пучков и самопроизвольно стираются при выключении света.

В отличие от обычной голографии, целью которой является получение изображения объекта, целью динамической голографии является преобразование самих записывающих пучков (либо по интенсивности, либо по пространственному спектру). Методами динамической голографии можно проводить корреляционный анализ непрерывно меняющихся образов, осуществлять простейшие (сложение, вычитание) и интегральные (фурье-преобразование, свертки) операции с двумерными образами, создавать оптические усилители световых пучков и изображений.

ОСНОВНЫЕ ФОРМУЛЫ

Разрешающая сила
оптического прибора

$$\Delta x = \frac{2\pi}{\Delta k_x} \sim \frac{\lambda}{2 \sin \alpha}$$

Угол расхождения
пучка

$$\theta \geq \lambda/a$$

Оптико-механическая
аналогия

$$k \rightarrow p$$

$$\omega \rightarrow E$$

$$v_g \rightarrow v$$

Принцип Гюйгенса

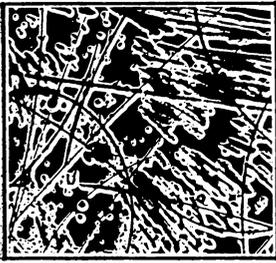
$$u_A = \frac{k}{2\pi i} \int_{\Sigma} \frac{u}{R} e^{ikR} \cos \alpha \, df$$

Дифракция от щели

$$dI = \frac{I_0}{\pi N \alpha k} \frac{\sin^2(k\alpha\theta)}{\theta^2} \left(\frac{\sin Nk\theta d}{\sin k\theta d} \right)^2 d\theta$$

Дифракция от круглого отверстия (или диска)

$$dI = I_0 \frac{J_1^2(k\theta R)}{\pi\theta^2} d\theta$$



IV. ВЗАИМОДЕЙСТВИЕ ЭЛЕКТРОМАГНИТНЫХ ВОЛН С ЧАСТИЦАМИ И ВОЛНАМИ

Глава 22. НЕЛИНЕЙНАЯ ОПТИКА

22.1. Нелинейные поляризованности и восприимчивости

В уравнении Максвелла для материальных сред входят две пары векторов: векторы \mathbf{E} и \mathbf{D} , характеризующие электрическое поле, и векторы \mathbf{H} и \mathbf{B} , характеризующие магнитное поле. Эти векторы связаны соотношениями

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}, \quad \mu_0^{-1} \mathbf{B} = \mathbf{H} + \mathbf{J},$$

где \mathbf{P} — поляризованность и \mathbf{J} — намагниченность. В случае достаточно слабых полей \mathbf{P} и \mathbf{J} являются линейными функциями напряженностей полей \mathbf{E} и \mathbf{H} ; вообще же \mathbf{P} и \mathbf{J} представляют собой сложные функции напряженностей. Линейный характер связи между \mathbf{J} и \mathbf{H} и между \mathbf{P} и \mathbf{E} быстро нарушается уже при сравнительно слабых полях в ферромагнетиках, сегнетоэлектриках и плазме, в большинстве же веществ нелинейные связи поляризацій с полями проявляются, как правило, в области сильных полей. В частности, нелинейные эффекты в диэлектриках в оптическом диапазоне начинают проявляться при $E \sim 10^8$ В/м. Такие поля могут создаваться лазерами. До появления лазеров нелинейные оптические эффекты практически нельзя было наблюдать, с помощью же лазеров это стало возможным, и в первую очередь был обнаружен эффект удвоения частоты при прохождении луча лазера через кристалл кварца.

Покажем, как возникают нелинейные оптические эффекты. Для этого рассмотрим воздействие световых волн на осциллятор, который послужит для нас моделью атомного электрона. Мы уже пользовались такой моделью раньше, считая осциллятор гармоническим; теперь же для описания нелинейных эффектов мы должны считать осциллятор ангармоническим. Выбирая описывающее ангармонизм слагаемое в простейшем виде, будем исходить из уравнения движения свободного одномерного осциллятора

$$\ddot{x} + \gamma \dot{x} + \omega_0^2 x + \beta x^2 = 0,$$

где ω_0 — собственная частота колебаний осциллятора, γ — коэффициент затухания колебаний и величина βx^2 описывает в простейшем случае ангармонический характер собственных сил, действующих на осциллятор.

Нас интересует воздействие на осциллятор электрического поля световой волны. Гармонический осциллятор удовлетворял, как мы знаем, принципу суперпозиции, поэтому достаточно было рассматривать воздействие на него только одной плоской монохроматической волны. Для ангармонического осциллятора принцип суперпозиции не имеет места, поэтому даже в простейшей ангармонической задаче следует считать, что внешнее поле — поле световой волны — содержит по крайней мере две гармоники с разными частотами. Мы будем, таким образом, исходить из следующего уравнения движения осциллятора во внешнем поле:

$$\ddot{x} + \gamma \dot{x} + \omega_0^2 x + \beta x^2 = \frac{2e}{m_e} \operatorname{Re} \left\{ E_1 e^{i(k_1 z - \omega_1 t)} + E_2 e^{i(k_2 z - \omega_2 t)} \right\},$$

где $2E_1$ и $2E_2$ — проекции на направление x амплитуд электрических полей обеих волн, e и m_e — заряд и масса электрона, ω_1 , ω_2 и k_1 , k_2 — частоты и волновые векторы волн, распространяющихся вдоль оси z (отличающаяся в два раза от обычной нормировка амплитуд полей удобна тем, что позволяет избежать усложняющих числовых множителей в выражениях для нелинейных восприимчивостей).

Так как нелинейное слагаемое βx^2 мало, то это уравнение можно решать по методу теории возмущений. В пренебрежении этим слагаемым (нулевое приближение) получим суперпозицию колебаний $x = x(\omega_1) + x(\omega_2)$ с частотами ω_1 и ω_2 , где

$$x(\omega) = \frac{(e/m_e) E}{-\omega^2 + \omega_0^2 - i\gamma\omega} e^{i(kz - \omega t)}$$

(под E , k понимается E_1, k_1 , если $\omega = \omega_1$, и E_2, k_2 , если $\omega = \omega_2$). В нелинейном приближении низшего порядка, т. е. в первом приближении по βx^2 , решение уравнения содержит колебания с частотами вторых гармоник $2\omega_1, 2\omega_2$, с комбинационными частотами $\omega_1 + \omega_2, \omega_1 - \omega_2$ и, наконец, слагаемое с нулевой частотой, соответствующее «выпрямлению» света за счет квадратичной нелинейности βx^2 .

Колебание на второй гармонике $2\omega_1$ определяется формулой

$$x(2\omega_1) = \frac{-e^2 \beta E_1^2}{\Delta^2(\omega_1) \Delta(2\omega_1)} e^{2i(k_1 z - \omega_1 t)},$$

где $\Delta(\omega) = -\omega^2 + \omega_0^2 - i\gamma\omega \equiv \Delta^*(-\omega)$.

Колебание с разностной частотой $\omega_1 - \omega_2$ имеет вид

$$x(\omega_1 - \omega_2) = \frac{-e^2 \beta E_1 E_2^*}{\Delta(\omega_1) \Delta^*(\omega_2) \Delta(\omega_1 - \omega_2)} e^{i[(k_1 - k_2)z - (\omega_1 - \omega_2)t]}.$$

Остальные три колебания описываются аналогичными формулами.

Исходя из этих выражений, можно ввести нелинейные поляризованности и восприимчивости. Поляризованность \mathbf{P} (отнесенная к одному электрону) есть, очевидно, ег. При учете ангармонизмов она содержит как линейное ($\mathbf{P}^{(L)}$), так и нелинейные (\mathbf{P}^{NL}) по полю слагаемые, причем нелинейных поляризованностей (в отличие от одной линейной) оказывается много: каждая из них относится к определенной генерации новых частот с помощью заданных исходных частот. В частности, генерации с помощью ангармонического осциллятора с нелинейностью вида βx^2 второй гармоники соответствует нелинейная поляризованность

$$P_x^{NL}(2\omega) = \chi_{xx}^{NL}(2\omega; \omega, \omega) E_x(\omega) E_x(\omega),$$

где χ_{xxx} — соответствующая нелинейная восприимчивость:

$$\chi_{xxx}(2\omega; \omega, \omega) = \frac{-n_0 e^3 \beta}{m_e^2 \Delta^2(\omega) \Delta(2\omega)}$$

(n_0 — плотность электронов). Индексы x, x, x написаны здесь для того, чтобы напомнить о тензорном характере соотношений, связывающих компоненты вектора P^{NL} с компонентами полей $E(\omega_1)$, $E(\omega_2)$, участвующих в нелинейном процессе. В общем случае такое соотношение имеет вид

$$P_i^{NL} = \chi_{ijl} E_j E_l,$$

где индексы i, j, l нумеруют оси координат (по дважды встречающемуся индексу подразумевается суммирование). Поля E_j , E_l могут содержать одинаковые или разные частоты ω_1 , ω_2 ; поляризованность же P_i^{NL} относится к какой-либо из частот $2\omega_1$, $2\omega_2$, $\omega_1 \pm \omega_2$ (или к нулевой частоте). Эти частоты должны быть указаны в качестве аргументов величин P_i^{NL} , E_j , E_l , а также величин χ_{ijl} , образующих тензор третьего ранга.

Заметим, что полученное выше выражение для $x(\omega_1 - \omega_2)$ позволяет сразу выписать формулы для поляризованности $P_x^{NL}(\omega_1 - \omega_2)$ и восприимчивости $\chi_{xxx}(\omega_1 - \omega_2; \omega_1, \omega_2)$; они имеют ту же структуру, что и приведенные формулы для $P_x^{NL}(2\omega)$ и $\chi_{xxx}(2\omega; \omega, \omega)$. Расчет нелинейных эффектов следующего порядка по βx^2 приведет к появлению колебаний с частотами $2\omega_1 \pm \omega_2$, $\omega_1 \pm 2\omega_2$ и т. д., а также к колебаниям с исходными частотами ω_1 , ω_2 ; в этом процессе в простейшем случае участвуют две исходные волны (с разными или совпадающими частотами).

Легко убедиться, что отношение нелинейных поляризованностей низшего порядка к линейной поляризованности составляет

$$\frac{P^{NL}}{P^L} \sim \frac{eE}{m_e \Delta} \frac{\beta}{\Delta}.$$

Величину β можно оценить, основываясь на том, что нелинейная сила $m\beta x^2$ должна быть того же порядка, что и линейная сила $m\omega_0^2 x$, если $x \sim a$, где a — радиус равновесной орбиты электрона. Вводя внутриатомное электрическое поле E_a так, что $m\omega_0^2 a \equiv eE_a$, получим $\beta/\Delta \sim \beta/\omega_0^2 \sim a^{-1}$ и

$$P^{NL}/P^L \sim E/E_a.$$

Типичное значение $E_a \sim 3 \cdot 10^{10}$ В/м, поэтому даже при напряженностях волны $E \sim 10^8$ В/м отношение P^{NL}/P^L составит примерно $3 \cdot 10^{-3}$. Это отношение увеличивается в ω_0/γ раз, если один из сомножителей, входящих в выражение для x , становится резонансным.

Итак, простейшая нелинейная поляризованность квадратична по полю волны и, следовательно, определяется тензором третьего ранга χ_{ijl} . Существенно иметь в виду, что в системах с центром инверсии не может существовать тензора третьего ранга, поэтому для наблюдения билинейных по полю эффектов должны использоваться кристаллы без центра инверсии. К таким кристаллам относятся, в частности, пьезоэлектрики; на одном из пьезоэлектриков — кварце — впервые наблюдались нелинейные эффекты в оптическом диапазоне (1961), а именно было зарегистрировано ультрафиолетовое излучение ($\lambda = 347$ нм) при прохождении через кристалл кварца луча от рубино-

вого лазера ($\lambda = 694 \text{ нм} \equiv 2 \cdot 347 \text{ нм}$; рис. 22.1). Если в числе элементов симметрии кристалла имеется центр инверсии, то оптическая нелинейность низшего порядка, которой может обладать кристалл, не квадратична, а кубична.

Остановимся на соотношениях симметрии, которым подчиняется нелинейная восприимчивость. Пусть имеются три волны: $\mathbf{E}_1(\omega_1)$, $\mathbf{E}_2(\omega_2)$ и $\mathbf{E}_3(\omega_3)$, где $\omega_3 = \omega_1 + \omega_2$, и пусть мы рассматриваем квадратичные по полям нелинейные поляризованности вида

$$P_i^{NL}(\omega_3) = \chi_{ijl}(\omega_3 = \omega_1 + \omega_2) E_j(\omega_1) E_l(\omega_2).$$

Мы используем здесь обозначение $\chi(\omega_3 = \omega_1 + \omega_2)$, что означает $\chi(\omega; \omega_1; \omega_2)$ с $\omega = \omega_1 + \omega_2$. Тогда имеют место следующие соотношения симметрии:

$$\chi_{ijl}(\omega_3 = \omega_1 + \omega_2) = \chi_{jil}(\omega_1 = \omega_3 - \omega_2) = \chi_{lij}(\omega_2 = \omega_3 - \omega_1).$$

Подчеркнем, что при изменении порядка координатных индексов (i, j, l) должен меняться порядок частотных аргументов. Если дисперсия во всей области частот $\omega_1, \omega_2, \omega_3$ пренебрежимо мала, то χ_{ijl} не меняется при изменении порядка индексов i, j, l и при неизменном порядке аргументов.

Мы рассмотрели простейшую нелинейность — квадратичную — в векторе нелинейной поляризованности, но она, как уже отмечалось, может существовать только у таких систем, которые не имеют центра симметрии. Нелинейность следующего порядка — кубичная — может существовать у всех кристаллов и даже у изотропной жидкости. Нелинейность третьего порядка означает, что обусловленная ею часть нелинейной поляризации имеет следующую структуру:

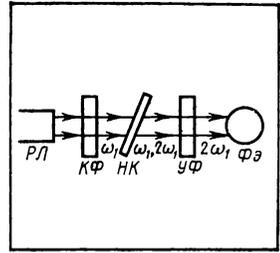
$$P_i^{NL}(\omega_1 + \omega_2 + \omega_3) = \hat{\chi}^{(4)} E_1 E_2 E_3 \exp \{ i(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \mathbf{r} - i(\omega_1 + \omega_2 + \omega_3)t \},$$

где $\hat{\chi}^{(4)}$ — нелинейная восприимчивость третьего порядка, представляющая собой тензор четвертого ранга. В развернутом виде i -компонента этого выражения (если опустить экспоненту) есть

$$P_i^{NL} = \chi_{ijkl}^{(4)} E_{1j} E_{2k} E_{3l}.$$

При $\omega_1 = \omega_2 = \omega_3$ эта нелинейная поляризованность обуславливает генерацию третьей гармоники.

Как только что указывалось, тензор нелинейной восприимчивости $\hat{\chi}^{(4)}$ существует и у изотропной жидкости; при этом он имеет две независимые компоненты: $\chi_{xyxy} \equiv \chi_{44}$ и $\chi_{xxyy} \equiv \chi_{12}$. Поэтому в случае трех одинаковых ($\mathbf{E}_1 = \mathbf{E}_2 = \mathbf{E}_3$) полей P_i^{NL} имеет фурье-составляющую



22.1.

Схема эксперимента по генерации второй гармоники света:

РЛ — рубиновый лазер,
 НК — нелинейный кристалл,
 КФ и УФ — красный и ультрафиолетовый фильтры,
 ФЭ — фотоэлемент

$$P^{NL}(\omega) = \chi_{44}(\omega) E(\omega) (E(\omega) E^*(-\omega)) + \chi_{12}(\omega) E^*(-\omega) (E(\omega) E(\omega)).$$

Если $\omega_1 > \omega_2$, $\omega_2 = \omega_3$ и $E_2 = E_3$, причем E_2 линейно поляризовано вдоль оси x , то нелинейная поляризованность с частотой ω_1 имеет компоненты

$$P_x^{NL}(\omega_1) = \chi_{11} E_{1x} |E_{2x}|^2, \quad P_y^{NL}(\omega_1) = \chi_{21} E_{1y} |E_{2x}|^2,$$

$$P_z^{NL}(\omega_1) = \chi_{21} E_{1z} |E_{2x}|^2.$$

Эти формулы имеют простой физический смысл: они описывают изменение показателя преломления волны с частотой ω_1 , обусловленное волной с частотой ω_2 .

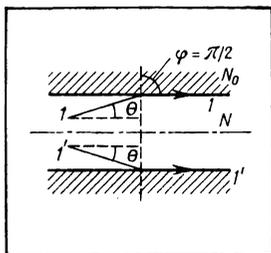
Если $\omega_2 = 0$ (т. е. поле E_2 постоянно), то эти формулы определяют изменение показателя преломления волны с частотой ω_1 , обусловленное наличием постоянного поля (эфф е к т К е р р а).

22.2. Самофокусировка и самомодуляция светового пучка

Нелинейности в поляризации приводят, как уже говорилось, к зависимости диэлектрической проницаемости и коэффициента преломления от напряженности электрического поля волны. Основными физическими эффектами, обуславливающими эту зависимость, являются эффект Керра для переменного поля, при котором происходит переориентация молекул, имеющих анизотропные поляризуемости, и эффект э л е к т р о с т р и к ц и и, приводящий к изменению плотности среды под действием поля волны. Ограничиваясь рассмотрением изотропных сред и учетом нелинейности низшего порядка, мы можем записать показатель преломления среды в виде

$$N = N_0 + N_2 |E|^2,$$

где E — напряженность электрического поля волны, N_0 — линейный показатель преломления и N_2 — некоторая величина, которая зависит от поля и может быть как положительной, так и отрицательной. Наиболее интересен случай $N_2 > 0$; в этом случае показатель преломления больше в тех местах, где интенсивность волн выше, поэтому в областях высокой интенсивности создается линзовый эффект, приводящий к формированию интенсивных световых нитей. Это явление носит название *самофокусировки* или *самоканалирования*.



22.2.

Самоканалирование при полном внутреннем отражении

Попытаемся разъяснить, как оно возникает. Для этого обратимся к рис. 22.2, на котором изображен расходящийся пучок света (угол расходимости 2θ , крайние лучи обоз-

начены I и I'), движущийся в области с показателем преломления N , резко отделенной от области с показателем преломления $N_0 < N$. Самоканалирование в этом случае происходит вследствие полного внутреннего отражения. Напомним, что согласно законам преломления угол преломления φ связан с углом падения $(\pi/2 - \theta)$ соотношением

$$\sin \varphi / \sin (\pi/2 - \theta) \equiv \sin \varphi / \cos \theta = N/N_0.$$

Считая величины θ и $N - N_0 = N^2 |E|^2$ малыми и замечая, что при полном внутреннем отражении $\varphi = \pi/2$, найдем минимальное критическое поле, приводящее к каналированию пучка с углом расходимости 2θ :

$$|E| = \sqrt{N_0/(2N_2)} \theta.$$

(Напомним, что минимальная угловая расходимость пучка радиуса r_0 , обусловленная дифракцией, есть $\theta \sim \lambda/(4N_0 r_0)$.)

Если поле превосходит критическое значение, то пучок лучей отразится от границы области с меньшим показателем преломления к оси пучка. Поэтому возникает вопрос о *длине фокусировки* z_f , т. е. расстоянии, пройденном лучом до пересечения с осью пучка. На рис. 22.3 показана траектория луча, начинающего свой путь на расстоянии r_1 от оси пучка.

Пусть квадрат напряженности поля вдоль поперечного сечения пучка имеет гауссово распределение:

$$|E|^2 = E_0^2 \exp \left(-2r^2/r_0^2 \right),$$

где E_0 — поле на оси пучка и r_0 — радиус пучка. Показатель преломления в этом случае определяется формулой

$$N \equiv N(r) = N_0 + 1/2 N_2 E_0^2 \exp \left(-2r^2/r_0^2 \right)$$

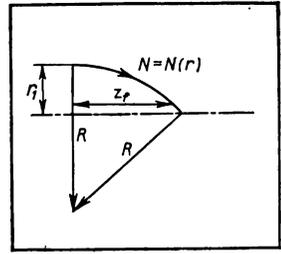
(множитель $1/2$ возникает потому, что поле является переменным и должно быть произведено усреднение по времени). При известном значении $N(r)$ можно найти радиус кривизны R луча:

$$\frac{1}{R} = \frac{1}{N} \frac{dN}{dr} \cong \frac{1}{N_0} \frac{d}{dr} \left[\frac{1}{2} N_2 E_0^2 \exp \left(-\frac{2r^2}{r_0^2} \right) \right].$$

Для $r \ll a$ получим отсюда

$$R = N_0 r_0^2 / (4N_2 E_0^2 r_1).$$

С другой стороны, из рис. 22.3 видно, что фокусное расстояние $z_f = \sqrt{2rR}$ при $r \ll R$. Подставляя сюда найденное выражение для R , получим



22.3.
Самофокусировка

$$z_f = \frac{r_0}{2} \frac{(2N_0/N_2)^{1/2}}{\varepsilon_0}.$$

Мы видим, что z_f не зависит от r_1 , т. е. фокусное расстояние одинаково для всех лучей. Отметим, однако, что наш вывод основывается на предположении о первоначальной параллельности лучей в пучке, в действительности же гауссов пучок не является параллельным, а стремится разойтись вследствие дифракции, которая противодействует фокусировке.

Остановимся еще на явлении *самоодуляции*, или самосжатия, пучка. Предварительно напомним, что в геометрической оптике, которой мы пользуемся, всякая величина f , описывающая поле волны, выражается формулой

$$f = a \exp(i\psi),$$

где a — амплитуда волны, ψ — ее фаза, или эйконал. Частота ω и волновой вектор \mathbf{k} связаны с ψ соотношениями

$$\omega = -\partial\psi/\partial t, \quad \mathbf{k} = \text{grad } \psi.$$

Величины a , ω , \mathbf{k} в плоской монохроматической волне постоянны, но вообще они могут зависеть от координат и времени, только изменения их на расстояниях порядка длины волны λ и в течение промежутков времени порядка периода волны T должны быть малыми по сравнению с самими величинами. В этом случае можно пользоваться геометрической оптикой, причем она может быть как линейной, так и нелинейной в зависимости от того, является ли поляризация линейной или нелинейной функцией поля.

Если волна мало отличается от плоской монохроматической и нелинейность мала, то

$$\psi(\mathbf{r}, t) = \mathbf{k}_0 \mathbf{r} - \omega_0 t + \varphi(\mathbf{r}, t),$$

где \mathbf{k}_0 и $\omega_0 \equiv \omega_0(\mathbf{k}_0)$ — волновой вектор и частота плоской монохроматической волны; $\varphi(\mathbf{r}, t)$ — малая добавка к эйконалу $|\varphi| \ll |\psi|$, обусловленная нелинейностью. Используя это выражение для ψ , получим

$$\omega(\mathbf{r}, t) = \omega_0 - \frac{\partial\varphi}{\partial t}, \quad \mathbf{k} = \mathbf{k}_0 + \text{grad } \varphi.$$

Именно последние слагаемые определяют зависимость ω и \mathbf{k} от \mathbf{r} и t . Ясно, что

$$\frac{\partial \mathbf{k}}{\partial t} = -\text{grad } \omega.$$

Так как φ определяется нелинейностью поляризации, то частота, вообще говоря, зависит от амплитуды волны a , которая, в свою очередь, есть функция координат и времени. Мы рассмотрим простейший случай, когда $\omega = \omega(\mathbf{k}, a^2)$, т. е. частота является функцией \mathbf{k} и a^2 . Если волна распространяется вдоль оси x и амплитуда волны мала, то

$$\omega = \omega_0(k_0) + \alpha a^2,$$

где $\alpha \equiv (\partial\omega/\partial a^2)_{a=0}$ [и a и k являются функциями x]. Поэтому соотношение $\partial k/\partial t = -\partial\omega/\partial x$ приобретет вид

$$\frac{\partial k}{\partial t} = -v_g \frac{\partial k}{\partial x} - \alpha \frac{\partial a^2}{\partial x},$$

где $v_g = \partial\omega/\partial k$ — групповая скорость.

Учтем теперь закон сохранения энергии. Так как плотность энергии пропорциональна a^2 и энергия переносится в волновом пакете с групповой скоростью $v_g(\mathbf{k}) = v_g$, то имеет место уравнение непрерывности

$$\frac{\partial a^2}{\partial t} + \frac{\partial}{\partial x} (v_g a^2) = 0.$$

Полученные уравнения для k и a^2 допускают, очевидно, стационарное решение $k = k_0$, $a = a_0$, соответствующее плоской монохроматической волне. Возникает, однако, вопрос об устойчивости этого решения.

Для исследования этого вопроса нужно «включить» малое возмущение и выяснить, как оно развивается в пространстве и во времени. Положим с этой целью $k = k_0 + \delta k$; $a = a_0 + \delta a$ и будем искать δk и δa в виде волн

$$\delta k = k' e^{i\eta x - i\nu t}, \quad \delta a = a' e^{i\eta x - i\nu t},$$

где η , ν , a' — волновое число, частота и амплитуда (или, как говорят, глубина) модуляции исходного решения; все эти величины будем считать малыми; $\nu \ll \omega_0$, $\eta \ll k_0$, $a' \ll a$, $k' \ll k_0$. Подстановка этого возмущения в уравнения для k и a^2 приводит (в линейном приближении) к следующей зависимости ν от η :

$$\nu = \eta \left(v_g \pm \sqrt{\alpha \frac{\partial v_g}{\partial k}} \right).$$

Мы видим, что возникают две волны модуляции (соответственно двум знакам перед корнем). Эти волны ведут себя совершенно по-разному в зависимости от того, положительна или отрицательна величина $\alpha(\partial v_g/\partial k)$. Если эта величина положительна, то фазовые скорости ν/η обеих модуляционных волн вещественны и близки к групповой скорости $v_g = (\partial\omega/\partial k)_0$ исходной волны. Если же $\alpha(\partial v_g/\partial k) < 0$, то фазовые скорости имеют малые мнимые добавки, из-за чего одна из волн нарастает в пространстве (т. е. с ростом x при заданном ν) по закону

$$\exp \left\{ \frac{\nu x}{v_g^2} \sqrt{|\alpha \partial v_g / \partial k|} \right\}.$$

Таким образом, исходная плоская волна становится неустойчивой и разбивается на отдельные волновые пакеты. Такую неустойчивость называют *модуляционной*.

Постулированное нами малое возмущение может иметь не только шумовую природу. Так, помимо сильного светового пучка в ту же область нелинейной среды под некоторым углом может быть введена слабая световая волна той

же частоты. Эта слабая сигнальная волна вместе с сильной запишет в нелинейной среде динамическую решетку изменения показателя преломления и при выполнении приведенного выше условия будет усиливаться за счет откачки интенсивности от сильной волны вследствие дифракции на решетке. Именно этот процесс лежит в основе динамической голографии.

22.3. Нелинейное взаимодействие электромагнитных волн

Важнейшим свойством оптически нелинейных сред является то, что в них

одна монохроматическая волна не может распространяться сама по себе, вместе с ней всегда должны распространяться волны с высшими гармониками и комбинационными частотами.

Действительно, запишем уравнения Максвелла для фурье-компонент электрического и магнитного полей с частотой ω :

$$\operatorname{rot} \mathbf{E}(\omega) = i\mu_0 \omega \mathbf{H}(\omega), \quad \operatorname{rot} \mathbf{H}(\omega) = -i\omega \mathbf{D}(\omega),$$

где $\mathbf{D}(\omega)$ — фурье-компонента вектора электрической индукции: $\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}^L + \mathbf{P}^{NL}$. Вспомогая, что $\mathbf{E}(\omega) + \epsilon_0^{-1} \mathbf{P}^L(\omega) = \epsilon(\omega) \times \mathbf{E}(\omega)$, перепишем второе из этих уравнений в виде

$$\operatorname{rot} \mathbf{H}(\omega) = -i\omega \epsilon_0 \epsilon(\omega) \mathbf{E}(\omega) - i\omega \mathbf{P}^{NL}(\omega).$$

Исключив $\mathbf{H}(\omega)$, приходим к волновому уравнению для $\mathbf{E}(\omega)$:

$$\operatorname{rot} \operatorname{rot} \mathbf{E}(\omega) - \left(\frac{\omega}{c}\right)^2 \epsilon(\omega) \mathbf{E}(\omega) = \left(\frac{\omega}{c}\right)^2 \epsilon_0^{-1} \mathbf{P}^{NL}(\omega).$$

В случае линейной среды $\mathbf{P}^{NL} = 0$ и мы получаем однородное волновое уравнение:

$$\Delta \mathbf{E}(\omega) + \frac{\omega^2}{c^2} \epsilon(\omega) \mathbf{E}(\omega) = 0.$$

В случае же линейной среды $\mathbf{P}^{NL} \neq 0$ и мы получаем неоднородное волновое уравнение, в котором нелинейная поляризованность \mathbf{P}^{NL} выступает в роли источника полей. Так как \mathbf{P}^{NL} выражается в виде суммы различных степеней поля, то фурье-компоненты нелинейной поляризованности $\mathbf{P}^{NL}(\omega)$ содержат произведения фурье-компонент полей с частотами, отличными от ω (алгебраическая сумма которых дает ω). Для этих компонент также должны быть записаны уравнения Максвелла; они также содержат нелинейные поляризации и, следовательно, фурье-компоненты полей с новыми частотами. Таким образом, мы приходим к системе связанных уравнений Максвелла для различных фурье-компонент поля. Система эта «развязывается» только в пренебрежении нелинейной поляризацией — в этом случае мы получаем независимые уравнения для каждой из фурье-компонент: иными словами, в этом случае волны с различными частотами распространяются независимо друг от друга. С учетом же \mathbf{P}^{NL} волны ста-

новятся связанными друг с другом и независимое их существование делается невозможным.

Например, в случае простейшей нелинейности низшего порядка ($P^{NL} \sim EE$) возникает взаимодействие трех волн с частотами ω_1 , ω_2 и $\omega_3 = \omega_1 + \omega_2$ и мы получаем систему связанных нелинейных волновых уравнений для фурье-компонент полей $E_1(\omega_1)$, $E_2(\omega_2)$, $E_3(\omega_3)$ вида

$$\begin{aligned} \text{rot rot } E_1(\omega_1) - \left(\frac{\omega_1}{c}\right)^2 \epsilon(\omega_1) E_1(\omega_1) = \\ = \epsilon_0^{-1} \left(\frac{\omega_1}{c}\right)^2 \chi(\omega_1 = \omega_3 - \omega_2) E_3(\omega_3) E_2^*(-\omega_2) \end{aligned}$$

(мы для компактности используем обозначение $\widehat{\chi EE}$ для вектора с компонентами $\chi_{ijl} E_j E_l$). Эти уравнения должны решаться с учетом граничных условий: непрерывности на границе двух сред тангенциальных составляющих всех фурье-компонент векторов напряженности полей и нормальных составляющих всех фурье-компонент векторов индукций (напоминаем, что фурье-компоненты векторов электрических индукций содержат нелинейные слагаемые P^{NL}).

Ввиду малости нелинейности часто оказывается разумной задача о генерации волн с комбинационными частотами, заданными гармоническими полями. При этом предполагается, что гармонические поля создают волну с комбинационной частотой, но ее поле не влияет на исходные гармонические поля, иными словами, не учитывается взаимодействие волн.

Разъясним теперь эффект взаимодействия волн. Рассмотрим, например, задачу о генерации второй гармоники. Границу раздела нелинейной и линейной сред выберем в качестве плоскости $z = 0$. Ограничимся рассмотрением взаимодействия только двух волн: исходной с частотой $\omega_1 = \omega$ и второй гармоники $\omega_2 = 2\omega$. Полагая

$$E_1(\omega) = e_1 A_1(z) e^{ik_1 r}, \quad E_2(2\omega) = ie_2 A_2(z) e^{ik_2 r},$$

где $e_{1,2}$ — векторы поляризации волн и $A_{1,2}(z)$ — их амплитуды (зависящие от z), получим два дифференциальных уравнения для $A_1(z)$ и $A_2(z)$, которые в случае $k_2 = 2k_1$ имеют вид

$$\frac{dA_1}{dz} = -CA_1 A_2, \quad \frac{dA_2}{dz} = CA_1^2,$$

где $C = \omega^2 [e_1 \chi(2\omega) e_2 e_1] / (2\epsilon_0 k_1 c^2)$. Нас интересует генерация второй гармоники, т. е. решение, при котором $A_2(0) = 0$, $A_1(0) \neq 0$. Оно имеет вид

$$A_2(z) = A_1(0) \text{th} [CA_1(0)z], \quad A_1(z) = A_1(0) \text{ch}^{-1} [CA_1(0)z].$$

Характерная длина нелинейного взаимодействия может быть определена как

$$l^{NL} = [CA_1(0)]^{-1}.$$

Ее числовое значение вычисляется по формуле

$$I^{NL} = [CA_1(0)]^{-1} = \frac{\varepsilon_0 N \lambda}{2\pi \chi(2\omega) A_1(0)},$$

где λ — длина волны в вакууме и N — показатель преломления (линейный). Для кристалла KH_2PO_4 $\chi(2\omega) \approx 1,77 \cdot 10^{-24}$ Ф·В⁻¹. Считая начальную амплитуду $A_1(0) = 1/2 \cdot 2,7 \cdot 10^7$ В/м, получим $I^{NL} = 0,005$ м при $\lambda = 694,3$ нм.

На опыте было получено 80%-ное преобразование мощности излучения рубинового лазера в излучение второй гармоники. Заметим, что в среде без потерь возможно полное преобразование мощности основного излучения во вторую гармонику.

22.4. Эффект Манделъштама — Бриллюэна

Наряду с нелинейным взаимодействием между только электромагнитными волнами имеет место и взаимодействие между электромагнитными волнами и волнами других типов — акустическими, а в ферромагнетиках также и со спиновыми. Начнем с рассмотрения взаимодействия электромагнитных волн с акустическими волнами.

При упругих деформациях твердых тел и при изменении плотности газа или жидкости происходит изменение коэффициента преломления вещества. Поэтому меняющиеся в пространстве и времени упругие деформации — акустические волны — оказывают влияние на распространение электромагнитных волн. Фазовая скорость акустических колебаний мала по сравнению со скоростью света (для твердых тел и жидкостей $v_s/c \sim 10^{-5}$, для газов это отношение еще меньше), поэтому это влияние можно трактовать как дифракцию электромагнитной волны на почти покоящейся решетке с переменным показателем преломления, образованной акустическими колебаниями. Рассеяние световой волны на акустической носит название **э ф ф е к т а М а н д е л ь ш т а м а — Б р и л л ю э н а**.

Проще всего этот эффект можно интерпретировать на квантовом языке, т. е. считать, что свет представляет собой совокупность фотонов — частиц с энергией $\hbar\omega$ и импульсом $\hbar\mathbf{k}$ (ω и \mathbf{k} — частота и волновой вектор света), а звук есть совокупность фононов — частиц с энергией $\hbar\Omega$ и импульсом $\hbar\mathbf{q}$ (точнее, квазичастиц с квазиимпульсом $\hbar\mathbf{q}$; Ω и \mathbf{q} — частота и волновой вектор звука).

Элементарными процессами взаимодействия являются испускание и поглощение фотоном фонона. В этих процессах выполняются законы сохранения энергии и импульса:

$$\hbar\omega \pm \hbar\Omega = \hbar\omega', \quad \hbar\mathbf{k} \pm \hbar\mathbf{q} = \hbar\mathbf{k}'$$

(знаки плюс и минус относятся к поглощению и испусканию фонона; ω' и \mathbf{k}' — частота и волновой вектор рассеянного света). Учитывая, что $\omega = ck$, $\omega' = ck'$ и $\Omega = v_s q$, мы видим, что при рассеянии света на звуке должны выполняться соотношения

$$\mathbf{k}' = \mathbf{k} \pm \mathbf{q}, \quad \frac{ck'}{N'} = \frac{ck}{N} \pm v_s q,$$

где $N \equiv N(\omega)$, $N' \equiv N(\omega')$ — показатели преломления. Так как $v_s \ll c$, то $\omega' \approx \omega$ и, следовательно, $k' \approx k$, так что треугольник, образованный тремя векторами \mathbf{k} , \mathbf{k}' , \mathbf{q} , является почти равнобедренным (рис. 22.4). Поэтому волновой вектор звука связан с углом рассеяния θ соотношением $q = 2k \sin(\theta/2)$, а частота звука — соотношением

$$\Omega = 2\omega N(\omega) \frac{v_s}{c} \sin \frac{\theta}{2}$$

Максимальная частота звука соответствует углу рассеяния $\theta = \pi$, т. е. рассеянию света в обратном направлении.

Изменение частоты света представляет собой нелинейный эффект, и мы можем сопоставить ему нелинейную поляризованность $\mathbf{P}^{\text{NL}}(\omega')$. Этот вектор является билинейной функцией электрического поля $\mathbf{E}(\omega)$ волны с частотой ω и тензора деформации $u_{ij}(\Omega)$, соответствующего звуковой волне частоты Ω :

$$P_i^{\text{NL}}(\omega') = \xi_{ijlm} E_j(\omega) u_{lm}(\Omega),$$

где ξ_{ijlm} — некоторый тензор четвертого ранга, характеризующий связь между электромагнитным и акустическим полями (мы опустили фазовый множитель $\exp\{i(\mathbf{k} + \mathbf{q})\mathbf{r} - i(\omega + \Omega)t\}$).

До сих пор мы не касались вопроса о происхождении акустических фононов. Это могут быть как фононы, возбужденные внешними источниками, так и тепловые фононы, связанные с тепловыми колебаниями решетки. Если речь идет о тепловых фононах, то число их определяется функцией распределения Планка

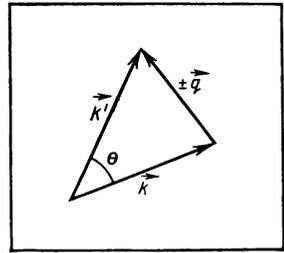
$$N_{\mathbf{q}} = \left(\exp \frac{\hbar v_s q}{T} - 1 \right)^{-1},$$

причем в твердом теле возможны три, а в жидкости и газе — одна поляризация фонона.

Фононы могут возбуждаться и той самой электромагнитной волной, которая на них рассеивается (вынужденное рассеяние Мандельштама — Бриллюэна). В основе этого эффекта лежит тот факт, что при взаимодействии электромагнитных волн, распространяющихся в твердом теле, возникает добавочное напряжение, тензор которого σ_{ij}^E является билинейной функцией напряженностей полей: ($\sigma_{ij}^E = \xi_{ijlm} E_l E_m$ (ξ — тот же тензор, что и в формуле для \mathbf{P}^{NL})). В фурье-компонентах имеем

$$\sigma_{ij}^E(\Omega) = \xi_{ijlm} E_l(\omega) E_m^*(\omega').$$

Добавочное переменное напряжение σ_{ij}^E вызовет переменную деформацию \mathbf{u} в соответствии с уравнением теории упругости:



22.4. Векторы \mathbf{k} , \mathbf{k}' , \mathbf{q} при рассеянии Мандельштама—Бриллюэна

$$\rho \frac{\partial^2 u_i}{\partial t^2} - \varepsilon \Delta u_i = \frac{\partial \sigma_{ij}^E}{\partial x_j}$$

(ρ — массовая плотность, $\varepsilon = \rho v_s^2$). Иными словами, поля с близкими частотами ω , ω' рождают акустические фононы с частотой $\Omega = \omega + \omega'$.

Если присоединить к этому уравнению волновые уравнения для полей $\mathbf{E}(\omega)$ и $\mathbf{E}(\omega')$, в которые входят нелинейные поляризованности $P_i^{\text{NL}} = \xi_{ijlm} u_{jl} E_m$, то мы получим систему трех связанных дифференциальных уравнений. Решая эту систему в приближении заданного поля и считая $\mathbf{E}(\omega)$ заданным, мы приходим к следующему выводу: амплитуды обеих вторичных волн (электромагнитной и звуковой) одновременно возрастают (естественно, за счет энергии поля исходной электромагнитной волны).

Аналогично комбинационному рассеянию на акустических колебаниях, возможно комбинационное рассеяние на спиновой волне. Максимальная частота спиновой волны $\omega_s \sim 10^{13} \text{ с}^{-1}$, поэтому в оптическом диапазоне всегда $\omega \gg \omega_s$, так что имеет место то же соотношение $q = 2k \sin(\theta/2)$ между волновыми векторами спиновой волны и света и углом рассеяния, что и для фононов ($\omega_s \equiv \omega_s(\mathbf{q})$ — частота спиновой волны). При рассеянии волны оптического диапазона главным механизмом, обеспечивающим связь волн, оказывается эффект Фарадея (в медленно меняющемся магнитном поле спиновой волны); при рассеянии же ультракоротких радиоволн главную роль играет модуляция магнитной проницаемости μ (входящей в коэффициент преломления $N = \sqrt{\varepsilon\mu}$) полем спиновой волны.

ОСНОВНЫЕ ФОРМУЛЫ

Нелинейная поляризация

$$\mathbf{D} = \varepsilon_0 \mathbf{E}_0 + \mathbf{P}, \quad \mathbf{P} = \mathbf{P}^{\text{L}} + \mathbf{P}^{\text{NL}},$$

$$P_i^{\text{L}} = \chi_{ij} E_j, \quad P_i^{\text{NL}} = \chi_{ijl} E_j E_l,$$

$$P_i^{\text{NL}}(\omega_3) = \chi_{ijl}(\omega_3 = \omega_1 + \omega_2) E_j(\omega_1) E_l(\omega_2)$$

Фокусное расстояние при самофокусировке

$$z_f = \frac{r_0}{2} \frac{(2N_0/N_2)^{1/2}}{E_0}$$

Законы сохранения при рассеянии света на звуке

$$\hbar \mathbf{k}' = \hbar \mathbf{k} \pm \hbar \mathbf{q}, \quad \hbar \omega' = \hbar \omega \pm \hbar \Omega$$

Глава 23. ИЗЛУЧЕНИЕ ЭЛЕКТРОМАГНИТНЫХ ВОЛН

23.1. Запаздывающие потенциалы

В гл. 4 показано, как по известным пространственным распределениям стационарных зарядов и токов могут быть определены порождаемые ими статические электрическое и магнитное поля. Теперь мы перейдем к решению общей задачи о нахождении полей, создаваемых произвольно движущимися зарядами и переменными токами. Естественно, для этого нужно исходить из уравнений Максвелла.

Будем сперва предполагать, что заряды и токи находятся в вакууме, и поэтому воспользуемся уравнениями Максвелла для вакуума:

$$\begin{aligned}\operatorname{rot} \mathbf{E} &= -\mu_0 \frac{\partial \mathbf{H}}{\partial t}, \quad \operatorname{div} \mathbf{H} = 0, \\ \operatorname{rot} \mathbf{H} &= \mathbf{j} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}, \quad \epsilon_0 \operatorname{div} \mathbf{E} = \rho.\end{aligned}$$

В случае статических полей мы выражали напряженности полей через скалярный φ и векторный \mathbf{A} потенциалы и получали для последних уравнения типа уравнения Пуассона. Поэтому вместо двух векторных величин \mathbf{E} и \mathbf{H} мы имели одну векторную и одну скалярную величину, т. е. число независимых величин сокращалось с шести до четырех. Попробуем теперь сделать то же самое в общем случае переменных полей и положим

$$\mu_0 \mathbf{H} = \operatorname{rot} \mathbf{A}, \quad \mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \operatorname{grad} \varphi,$$

где \mathbf{A} и φ — некоторые пока неизвестные (и можно сказать — вспомогательные) функции координат и времени. Если мы подставим эти выражения в первую пару уравнений Максвелла, то они удовлетворятся при любых \mathbf{A} и φ . Это значит, что написанные выражения эквивалентны первой паре уравнений Максвелла.

Подставив написанные выражения во вторую пару уравнений Максвелла, получим

$$\begin{aligned}\operatorname{rot} \operatorname{rot} \mathbf{A} &= \mu_0 \mathbf{j} + \frac{1}{c^2} \left(-\frac{\partial^2 \mathbf{A}}{\partial t^2} - \operatorname{grad} \frac{\partial \varphi}{\partial t} \right), \\ \epsilon_0 \operatorname{div} \left(-\frac{\partial \mathbf{A}}{\partial t} - \operatorname{grad} \varphi \right) &= \rho.\end{aligned}$$

До сих пор \mathbf{A} и φ были ничем не связаны друг с другом. Положим теперь на них дополнительное условие (у с л о в и е Л о р е н ц а)

$$\operatorname{div} \mathbf{A} + \frac{1}{c^2} \frac{\partial \varphi}{\partial t} = 0.$$

Учитывая, что

$$\operatorname{rot} \operatorname{rot} \mathbf{A} = \operatorname{grad} \operatorname{div} \mathbf{A} - \Delta \mathbf{A}, \quad \operatorname{div} \operatorname{grad} \varphi = \Delta \varphi,$$

получим

$$\Delta \varphi - \frac{1}{c^2} \frac{\partial^2 \varphi}{\partial t^2} = - \frac{\rho}{\epsilon_0},$$

$$\Delta \mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} = - \mu_0 \mathbf{j}.$$

Эти уравнения отличаются от уравнений для φ и \mathbf{A} в случае постоянных полей наличием вторых производных по времени. Мы по-прежнему будем называть φ и \mathbf{A} скалярным и векторным потенциалами электромагнитного поля, а так как полученные уравнения в отсутствие зарядов и токов переходят в известное нам волновое уравнение $\Delta f - (1/c^2) \partial^2 f / \partial t^2 = 0$, то мы назовем эти уравнения волновыми уравнениями с источниками.

Возникает теперь вопрос, как решать такие уравнения. Хорошо известно, что уравнение

$$\Delta f - \frac{1}{c^2} \frac{\partial^2 f}{\partial t^2} = -g,$$

где g — известная функция \mathbf{r} и t , имеет бесчисленное множество решений. Общее его решение имеет вид суммы какого-либо частного решения неоднородного уравнения и общего решения однородного уравнения (т. е. уравнения с $g = 0$), но последних существует бесчисленное множество и они соответствуют всевозможным волнам, способным распространяться в пустом пространстве.

Для нахождения частного решения воспользуемся принципом суперпозиции, справедливым для любых линейных уравнений, согласно которому если известны частные решения f_1 и f_2 для двух неоднородностей g_1 и g_2 , то частным решением при $g = g_1 + g_2$ будет сумма $f = f_1 + f_2$. По этой причине нам достаточно решить следующую задачу: в начале координат находится точечный заряд q , зависящий от времени; требуется определить скалярный потенциал φ , порождаемый этим зарядом. В этом случае $\rho = q(t)\delta(\mathbf{r})$, где $\delta(\mathbf{r})$ — уже неоднократно использовавшаяся нами трехмерная δ -функция, равная нулю при $\mathbf{r} \neq 0$, обращающаяся в бесконечность при $\mathbf{r} = 0$ и нормированная так, что

$$\int_V F(\mathbf{r}) \delta(\mathbf{r}) dV = F(0)$$

($F(\mathbf{r})$ — любая достаточно гладкая функция и V — произвольный объем, включающий точку $\mathbf{r} = 0$).

Так как объемная плотность заряда ρ обращается при $\mathbf{r} \neq 0$ в нуль, то при $\mathbf{r} \neq 0$ потенциал φ удовлетворяет уравнению

$$\Delta \varphi - \frac{1}{c^2} \frac{\partial^2 \varphi}{\partial t^2} = 0 \quad (\mathbf{r} \neq 0).$$

Мы должны искать сферически симметричное решение этого уравнения, т. е. считать, что φ зависит только от $r = |r|$ и t . В таком случае

$$\Delta\varphi = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial\varphi}{\partial r} \right)$$

и мы приходим к уравнению

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial\varphi}{\partial r} \right) - \frac{1}{c^2} \frac{\partial^2\varphi}{\partial t^2} = 0.$$

Если вместо φ ввести новую функцию $\chi = r\varphi$, то мы получим

$$\frac{\partial^2\chi}{\partial r^2} - \frac{1}{c} \frac{\partial^2\chi}{\partial t^2} = 0.$$

С таким уравнением мы уже встречались при изучении плоских волн — ему (если заменить r на x) удовлетворяла одномерная плоская волна, так что мы можем сразу написать его решение:

$$\chi = \chi_1(t - r/c) + \chi_2(t + r/c),$$

где χ_1, χ_2 — две произвольные функции. Функция χ_1 описывает волну, распространяющуюся от начала координат, а χ_2 — волну, распространяющуюся к началу координат; соответственно, функцию $\varphi_1 = \chi_1/r$ называют *расходящейся*, а функцию $\varphi_2 = \chi_2/r$ — *сходящейся сферическими волнами*. Источник, находящийся в начале координат, может, разумеется, создавать только расходящуюся, но не сходящуюся волну; поэтому мы можем сразу положить $\chi_2 \equiv 0$.

Заметим теперь, что на малых расстояниях от заряда поле должно иметь электростатическую структуру. Действительно, при $r \ll ct$ пространственные производные φ значительно больше временных производных и поэтому приближенное $\varepsilon_0\Delta\varphi = -\rho(r \ll ct)$. Следовательно, при $r \ll ct$ имеем $\varphi(r, t) = q(t)/(4\pi\varepsilon_0 r)$, откуда для произвольных r и t получим

$$\varphi(r, t) = \frac{1}{4\pi\varepsilon_0 r} q \left(t - \frac{r}{c} \right).$$

Мы видим, что потенциал (а следовательно, и поле), порождаемое переменным зарядом, определяется значением последнего не в момент времени t , а в более ранний момент времени $t^* = t - r/c$; величина $t - t^* = r/c$ представляет собой время, которое необходимо электромагнитной волне, чтобы пройти путь от точки нахождения заряда до точки наблюдения. Это время называют *временем запаздывания*, а найденный потенциал $\varphi = q(t^*)/(4\pi\varepsilon_0 r)$ — *запаздывающим потенциалом*.

Если бы мы положили $\chi_1 = 0$ (а не $\chi_2 = 0$), то получили бы решение в виде так называемого *опережающего потенциала*: $\varphi(r, t) = q(t + r/c)/(4\pi\varepsilon_0 r)$. Оно имеет смысл, если есть сферическая волна, сходящаяся к началу координат; поэтому опережающими потенциалами пользуются при решении задач в ограниченном пространстве, например в волноводе (но не в неограниченном свободном пространстве).

Нам остается обобщить полученный результат на случай произ-

вольного распределения зарядов. Замечая, что $\rho(\mathbf{r}, t) \equiv \int dV' \delta(\mathbf{r}-\mathbf{r}') \times \times \rho(\mathbf{r}', t)$, и воспользовавшись принципом суперпозиции, получим

$$\varphi(\mathbf{r}, t) = \int \frac{\rho\left(\mathbf{r}', t - \frac{1}{c} |\mathbf{r} - \mathbf{r}'|\right)}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}'|} dV',$$

где $dV' = dx' dy' dz'$. Аналогичную формулу можно написать и для запаздывающего векторного потенциала:

$$\mathbf{A}(\mathbf{r}, t) = \int \frac{\mu_0 \mathbf{j}\left(\mathbf{r}', t - \frac{1}{c} |\mathbf{r} - \mathbf{r}'|\right)}{4\pi |\mathbf{r} - \mathbf{r}'|} dV'.$$

23.2. Потенциалы Лиенара — Вихерга

Покажем, как с помощью общих формул для запаздывающих потенциалов найти потенциалы, создаваемые произвольным образом движущейся точечной заряженной частицей. Предварительно заметим, что запаздывающим потенциалам можно придать вид четырехмерных интегралов:

$$\varphi(\mathbf{r}, t) = \int \frac{\rho(\mathbf{r}', t')}{4\pi\epsilon_0 |\mathbf{r}' - \mathbf{r}|} \delta\left(t' - t + \frac{1}{c} |\mathbf{r}' - \mathbf{r}|\right) dV' dt',$$

$$\mathbf{A}(\mathbf{r}, t) = \int \frac{\mu_0 \mathbf{j}(\mathbf{r}', t')}{4\pi |\mathbf{r}' - \mathbf{r}|} \delta\left(t' - t + \frac{1}{c} |\mathbf{r}' - \mathbf{r}|\right) dV' dt'.$$

Введенная δ -функция может быть устранена интегрированием по dt' ; при этом, согласно определению δ -функции, в аргументах ρ и \mathbf{j} должно быть положено $t' = t - |\mathbf{r}' - \mathbf{r}|/c$ в соответствии с разъясненным в предыдущем параграфе эффектом запаздывания.

Пусть теперь имеется точечный заряд e , произвольным образом движущийся по траектории $\mathbf{r} = \mathbf{r}_e(t)$. При этом, очевидно,

$$\rho(\mathbf{r}, t) = e\delta(\mathbf{r} - \mathbf{r}_e(t)), \quad \mathbf{j} = e\mathbf{v}\delta(\mathbf{r} - \mathbf{r}_e(t)),$$

где $\mathbf{v} \equiv \mathbf{v}(t) = \dot{\mathbf{r}}_e(t)$ — скорость заряда. Подставив эти выражения в формулы для φ и \mathbf{A} и устранив одну из δ -функций интегрированием по \mathbf{r}' , получим

$$\varphi(\mathbf{r}, t) = \frac{e}{4\pi\epsilon_0} \int \frac{1}{|\mathbf{r} - \mathbf{r}_e(t')|} \delta\left(t' - t + \frac{1}{c} |\mathbf{r} - \mathbf{r}_e(t')|\right) dt',$$

$$\mathbf{A}(\mathbf{r}, t) = \frac{\mu_0 e}{4\pi} \int \frac{\mathbf{v}(t')}{|\mathbf{r} - \mathbf{r}_e(t')|} \delta\left(t' - t + \frac{1}{c} |\mathbf{r} - \mathbf{r}_e(t')|\right) dt'.$$

Интегрирование по t' легко выполнить, если воспользоваться общим правилом интегрирования выражений, содержащих δ -функцию сложного аргумента:

$$\int g(x) \delta(f(x) - \alpha) dx = \left(\frac{g(x)}{|df/dx|} \right)_{f(x)=\alpha}.$$

Полагая здесь $f(t') = t' + |\mathbf{r} - \mathbf{r}_e(t')|/c$ и замечая, что $df/dt' = 1 - \mathbf{nv}(t')/c$, получим

$$\varphi(\mathbf{r}, t) = \frac{e}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}_e(t')| (1 - \mathbf{nv}(t')/c)},$$

$$\mathbf{A}(\mathbf{r}, t) = \frac{\mu_0 e \mathbf{v}(t')}{4\pi |\mathbf{r} - \mathbf{r}_e(t')| (1 - \mathbf{nv}(t')/c)},$$

где \mathbf{n} — единичный вектор в направлении $\mathbf{r} - \mathbf{r}_e(t')$ и t' связано с t соотношением $t' = t - |\mathbf{r} - \mathbf{r}_e(t')|/c$. Эти потенциалы носят название *потенциалов Лиенара — Вихерта*. Для нерелятивистского движения ($v \ll c$) стоящие в знаменателях этих формул выражения в круглых скобках должны быть заменены единицей.

Зная потенциалы и используя общие соотношения

$$\mu_0 \mathbf{H} = \text{rot } \mathbf{A}, \quad \mathbf{E} = - \frac{\partial \mathbf{A}}{\partial t} - \text{grad } \varphi,$$

можно вычислить напряженности полей, создаваемых произвольно движущимся зарядом. (Потенциалы при этом следует дифференцировать по координатам точки наблюдения.) Мы приведем только окончательный результат:

$$\mathbf{E}(\mathbf{r}, t) = \frac{e}{4\pi\epsilon_0} \left\{ \frac{(\mathbf{n} - \mathbf{v}/c)(1 - v^2/c^2)}{(1 - \mathbf{nv}/c)^3 R^2} \right\}_{t'} +$$

$$+ \frac{\mu_0 e}{4\pi} \left[\frac{\dot{\mathbf{n}}}{(1 - \mathbf{nv}/c)^3 R} \left[(\mathbf{n} - \mathbf{v}/c) \dot{\mathbf{v}} \right] \right]_{t'}, \quad \mathbf{H} = c\epsilon_0 [\mathbf{nE}],$$

где $\mathbf{R}(t) = \mathbf{r} - \mathbf{r}_e(t)$, $\mathbf{n} = \mathbf{R}/R$, $\dot{\mathbf{v}} = d\mathbf{v}/dt$ и индекс t' у скобок означает, что все величины берутся с учетом запаздывания, т. е. в момент времени $t' = t - R(t')/c$.

Обратим внимание на то, что электрическое и магнитное поля перпендикулярны друг другу. Поля складываются из двух слагаемых: одного, убывающего с расстоянием как R^{-2} , и другого, убывающего как R^{-1} . Второе слагаемое перпендикулярно вектору \mathbf{n} и пропорционально ускорению $\dot{\mathbf{v}}$; именно с ним связано излучение энергии. Первое слагаемое носит статический характер и описывает преобразованное к движущейся системе отсчета кулоновское поле заряда.

23.3. Дипольное излучение

Определим поля вдали от порождающих их зарядов: $r \gg a$, где a — размеры области, в которой сосредоточены заряды. В знаменателе полученного в § 23.1 общего выражения для запаздывающего потенциала мы можем при этом пренебречь r' по сравнению с r . Далее, пренебрегая квадратичными по $r'^2 \sim a^2$ слагаемыми, вместо $|\mathbf{r} - \mathbf{r}'|$ можно подставить $r(1 - \mathbf{r}\mathbf{r}'/r^2)$. Таким образом,

$$\mathbf{A}(\mathbf{r}, t) = \frac{\mu_0}{4\pi r} \int \mathbf{j} \left(\mathbf{r}, t - \frac{r}{c} \left[1 - \frac{\mathbf{r}\mathbf{r}'}{r^2} \right] \right) dV'.$$

Если заряды перемещаются относительно медленно, так что $T \gg \gg a/c$ (T — характерное время их перемещения), то можно не учитывать и пропорционального r' слагаемого в квадратных скобках в аргументе \mathbf{j} ; иными словами, можно полностью пренебречь внутренним запаздыванием, т. е. различием между запаздыванием для разных зарядов системы. В случае периодического движения $T \sim \lambda/c$, где λ — длина излучаемой зарядами волны. Таким образом, если $r \gg \gg a$ и $\lambda \gg \gg a$, то

$$\mathbf{A}(\mathbf{r}, t) = \frac{\mu_0}{4\pi r} \int \mathbf{j}\left(\mathbf{r}, t - \frac{r}{c}\right) dV'$$

или, переходя от интегрирования к суммированию по отдельным зарядам,

$$\mathbf{A}(\mathbf{r}, t) = \frac{\mu_0}{4\pi r} \sum_i e_i \mathbf{v}_i \left(t - \frac{r}{c}\right).$$

Замечая, что дипольный момент системы зарядов есть $\mathbf{d}(t) = \sum_i e_i \mathbf{r}_i(t)$, получим окончательно

$$\mathbf{A}(\mathbf{r}, t) = \frac{\mu_0}{4\pi r} \dot{\mathbf{d}}\left(t - \frac{r}{c}\right).$$

Таким образом, вдали от системы медленно перемещающихся зарядов ($r \gg \gg a$, $\lambda \gg \gg a$) векторный потенциал пропорционален производной по времени от дипольного момента; такое излучение называют *дипольным*. (Если симметрия распределения зарядов такова, что $\mathbf{d} = 0$, то нельзя пренебречь внутренним запаздыванием $\mathbf{gr}'/(cr)$; к этому случаю мы вернемся ниже.)

Найдем теперь поля, создаваемые рассматриваемой системой зарядов. Напряженность магнитного поля \mathbf{H} связана с \mathbf{A} соотношением $\mu_0 \mathbf{H} = \text{rot } \mathbf{A}$, поэтому

$$\mathbf{H}(\mathbf{r}, t) = \text{rot } \frac{\dot{\mathbf{d}}(t - r/c)}{4\pi r}.$$

Воспользовавшись формулой $\text{rot } \chi \mathbf{b} = \chi \text{rot } \mathbf{b} + [\text{grad } \chi, \mathbf{b}]$, где χ — скаляр и \mathbf{b} — вектор, имеем

$$\mathbf{H} = \frac{1}{4\pi r} \left[\nabla \dot{\mathbf{d}}\left(t - \frac{r}{c}\right) \right] + \frac{1}{4\pi} \left[\text{grad } \frac{1}{r}, \dot{\mathbf{d}}\left(t - \frac{r}{c}\right) \right],$$

где ∇ — дифференциальный оператор набла (вектор с компонентами $(\partial/\partial x, \partial/\partial y, \partial/\partial z)$). Величина $\dot{\mathbf{d}}$ зависит от координат в комбинации $t^* = t - r/c$, а так как $\partial r/\partial x = x/r$ и (аналогично для $\partial r/\partial y, \partial r/\partial z$), то

$$\text{rot } \dot{\mathbf{d}}\left(t - \frac{r}{c}\right) = -\frac{1}{cr} \left[\mathbf{r} \ddot{\mathbf{d}}\left(t - \frac{r}{c}\right) \right].$$

Далее, $\text{grad } (1/r) = -\mathbf{r}/r^3$ и, следовательно,

$$\mathbf{H}(\mathbf{r}, t) = \frac{1}{4\pi cr} \left[\ddot{\mathbf{d}}\left(t - \frac{r}{c}\right), \mathbf{r} \right] + \frac{1}{4\pi r^2} \left[\dot{\mathbf{d}}\left(t - \frac{r}{c}\right), \mathbf{r} \right].$$

Второе слагаемое здесь пропорционально r^{-2} , а первое — r^{-1} , поэтому на больших расстояниях ($r \gg a$) существенно только первое слагаемое и окончательно

$$\mathbf{H} = \frac{1}{4\pi cr} \left[\ddot{\mathbf{d}} \left(t - \frac{r}{c} \right), \mathbf{r} \right].$$

Аналогично вычисляется напряженность электрического поля $\mathbf{E}(\mathbf{r}, t)$; приведем лишь окончательный результат, справедливый при $r \gg a$: по модулю $V\epsilon_0 E = V\mu_0 H$, причем \mathbf{E} направлено перпендикулярно \mathbf{H} и \mathbf{r} , так что $V\epsilon_0 \mathbf{E} = (V\mu_0/r)[\mathbf{H}, \mathbf{r}]$. Итак,

поля, создаваемые переменными зарядами и токами, на больших расстояниях убывают обратно пропорционально первой степени расстояния.

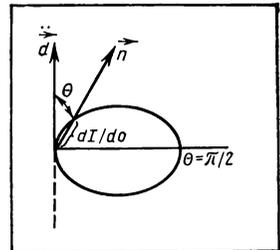
В этом отношении они существенно отличаются от статических полей. Действительно, как мы видели в гл. 1, 4, электростатические поля заряда и диполя убывают соответственно как r^{-2} и r^{-3} , а магнитное поле постоянного тока (конечных размеров) — как r^{-3} . Таким образом,

переменные поля убывают значительно медленнее статических полей.

Именно с убыванием по закону r^{-1} полей, создаваемых переменными токами и зарядами, связано важнейшее явление — возможность излучения такими токами и зарядами электромагнитных волн. Для того чтобы пояснить это, вычислим поток электромагнитной энергии \mathcal{J} через замкнутую поверхность Σ , окружающую рассматриваемую систему зарядов и токов: $\mathcal{J} = \int_{\Sigma} \mathbf{d}\mathcal{J}$, где $\mathbf{S} = [\mathbf{E}\mathbf{H}]$ — вектор Пойнтинга и $d\sigma$ — элемент поверхности. Пусть для определенности поверхности Σ есть сфера большого радиуса $R (R \gg a)$. Используя найденные выше выражения для полей и учитывая, что для сферы $d\sigma = \mathbf{n}r^2 d\omega$, где \mathbf{n} — единичный вектор нормали к сфере и $d\omega$ — элемент телесного угла, получим

$$\mathcal{J} = \int d\mathcal{J}, \quad d\mathcal{J} = \frac{1}{(4\pi)^2 \epsilon_0 c^3} [\ddot{\mathbf{d}} \mathbf{n}]^2 d\omega.$$

Мы видим, что поток энергии \mathcal{J} на больших расстояниях R не зависит от R и, следовательно, стремится при $R \rightarrow \infty$ к постоянному пределу. Этот предел называют *интенсивностью излучения* (в рассматриваемом случае — дипольного). Величину $d\mathcal{J}$ можно интерпретировать как интенсивность излучения в элемент телесного угла $d\omega$. Вводя угол θ между векторами \mathbf{n} и $\ddot{\mathbf{d}}$, мы видим, что $d\mathcal{J}$ пропорционально $\sin^2\theta$; эта угловая зависимость определяет *диаграмму направленности дипольного излучения* (рис. 23.1).



23.1. Диаграмма направленности дипольного излучения

Подчеркнем, что в случае статических полей пропорциональные $[\mathbf{E}\mathbf{H}]$ величины убывают быстрее, чем r^{-2} , и поэтому соответствующие потоки при $r \rightarrow \infty$ стремятся к нулю. По этой причине в случае статических полей нет излучения. Не следует, однако, думать, что в статическом случае величина $[\mathbf{E}\mathbf{H}]$ лишена смысла. Она всегда имеет смысл плотности потока электромагнитной энергии. Рассмотрим, например, цилиндрический проводник, по которому течет постоянный ток. Тогда, как легко убедиться, вектор \mathbf{S} направлен внутрь проводника и полный поток \mathbf{S} совпадает с джоулевой теплотой, выделяемой в проводнике.

Проинтегрировав $d\mathcal{J}$ по телесному углу, найдем полную интенсивность дипольного излучения:

$$\mathcal{J} = \frac{1}{6\pi\epsilon_0 c^3} \ddot{\mathbf{d}}^2.$$

Если дипольный момент изменяется по гармоническому закону: $\mathbf{d} = \mathbf{d}_0 \cos \omega t$, где \mathbf{d}_0 — постоянный вектор, то среднее (по периоду колебаний) значение \mathcal{J} составляет

$$\langle \mathcal{J} \rangle = \frac{\omega^4}{12\pi\epsilon_0 c^3} \mathbf{d}_0^2.$$

Мы видим, что (усредненная) интенсивность дипольного излучения растет с частотой ω пропорционально четвертой степени частоты.

Полученные формулы можно применить и в том случае, когда система состоит из одного движущегося заряда. Тогда $\mathbf{d} = e\mathbf{r}$ и $\ddot{\mathbf{d}} = e\mathbf{w}$, где \mathbf{w} — ускорение заряда, так что формула для интенсивности излучения приобретает вид

$$\mathcal{J} = \frac{e^2 w^2}{6\pi\epsilon_0 c^3}.$$

Заряд, движущийся с ускорением, обязательно излучает, причем излученная за единичное время энергия пропорциональна квадрату ускорения.

Формула эта справедлива только для нерелятивистских скоростей $v \ll c$. Действительно, в общей формуле для интенсивности дипольного излучения пренебрегают внутренним запаздыванием: $a \ll cT$. Но $T \sim a/v$, где v — скорость движения заряда, откуда и следует, что $v \ll c$.

Если заряженная частица движется в постоянном магнитном поле с индукцией \mathbf{B} , то $\mathbf{w} = e[\mathbf{v}\mathbf{B}]$ и поэтому интенсивность излучения определяется формулой

$$\mathcal{J} = \frac{e^4 v_\perp}{6\pi\epsilon_0 c^3} B^2 \sin^2 \vartheta$$

(ϑ — угол между \mathbf{v} и \mathbf{B}).

Рассмотрим еще излучение электрона, притягивающегося к ядру по закону Кулона: $v = -Ze/(4\pi\epsilon_0 r)$. В этом случае возможно движение по окружности с ускорением $w = v^2/a$, где a — радиус электронной орбиты. Замечая, что $m_e w = eE = Ze^2/(4\pi\epsilon_0 a^2)$, получим

$$\mathcal{J} = \frac{2Z^2 e^6}{3(4\pi\epsilon_0)^3 m_e^2 c^3 a^4}$$

(m_e — масса электрона).

Энергия электрона в атоме

$$W = \frac{mv^2}{2} - \frac{Ze^2}{4\pi\epsilon_0 a} = \frac{-Ze^2}{8\pi\epsilon_0 a}$$

(знак минус связан с финитностью движения). Отсюда легко оценить время τ , за которое излучится энергия $|W|$:

$$\tau \sim \frac{|W|}{\mathcal{J}} \sim \frac{ac}{Z} \left(\frac{a}{r_0}\right)^2,$$

где $r_0 = e^2/(4\pi\epsilon_0 m_e c^2)$ называют *классическим радиусом электрона*.

Применим эту формулу к классической планетарной модели атома, например атома водорода ($z = 1$). В этом случае $a \sim 10^{-10}$ м, так что $\tau \sim 10^{-9}$ с. Таким должно быть время жизни атома водорода согласно классической модели. Мы видим, что излучение атомным электроном электромагнитных волн делает невозможным объяснение устойчивости атома в рамках классической физики, основывающейся на механике Ньютона и электродинамике Максвелла. Это противоречие классической физики, выход из которого предложил Бор в своих знаменитых постулатах (1913), и было одной из главных причин создания квантовой физики.

Как мы уже говорили, в ряде случаев в силу симметрии излучающей системы зарядов дипольный момент системы тождественно равен нулю (или сохраняется), так что дипольное излучение невозможно. В связи с этим необходимо изучить кроме дипольного и другие возможные типы излучения. Для этого необходимо учесть эффект внутреннего запаздывания.

Возвращаясь к общей формуле для запаздывающего векторного потенциала

$$\mathbf{A} = \frac{\mu_0}{4\pi r} \int \mathbf{j} \left(t - \frac{1}{c} |\mathbf{r} - \mathbf{r}'| + \frac{\mathbf{r}' \cdot \mathbf{n}}{c} \right) dV',$$

где $\mathbf{n} = \mathbf{r}/r$, мы сохраним в разложении подынтегрального выражения по степеням внутреннего запаздывания $\mathbf{r}' \cdot \mathbf{n}/c$ два первых члена. В результате получим

$$\mathbf{A} = \frac{\mu_0}{4\pi r} \int \mathbf{j} dV' + \frac{\mu_0}{4\pi cr} \frac{\partial}{\partial t} \int (\mathbf{r}' \cdot \mathbf{n}) \mathbf{j} dV', \quad \mathbf{j} \equiv \mathbf{j} \left(t - \frac{1}{c} |\mathbf{r} - \mathbf{r}'| \right)$$

или, переходя к точечным зарядам,

$$\mathbf{A} = \frac{\mu_0}{4\pi r} \sum e \mathbf{v} + \frac{\mu_0}{4\pi cr} \frac{\partial}{\partial t} \sum e \mathbf{v} (\mathbf{r}' \cdot \mathbf{n})$$

(все величины в правой части берутся с учетом запаздывания, т. е. в момент времени $t' = t - |\mathbf{r} - \mathbf{r}'|/c$, где $\mathbf{r}(t)$ — радиус-вектор заряда).

Преобразуем второе слагаемое:

$$\begin{aligned} \mathbf{v}(\mathbf{r}, \mathbf{n}) &= \frac{1}{2} \frac{\partial}{\partial t} \mathbf{r}(\mathbf{n}, \mathbf{r}) + \frac{1}{2} \mathbf{v}(\mathbf{n}, \mathbf{r}) - \frac{1}{2} \mathbf{r}(\mathbf{n}, \mathbf{v}) = \\ &= \frac{1}{2} \frac{\partial}{\partial t} \mathbf{r}(\mathbf{n}, \mathbf{r}) + \frac{1}{2} [\mathbf{r}, \mathbf{v}] \mathbf{n}. \end{aligned}$$

Заметим, далее, что к \mathbf{A} можно прибавить произвольный вектор, параллельный \mathbf{n} : поля при этом не изменятся. Вспоминая определения квадрупольного и магнитного моментов системы зарядов

$$D_{ij} = \sum e (3x_i x_j - r^2 \delta_{ij}), \quad \mathbf{m} = \frac{1}{2} \sum e [\mathbf{r}\mathbf{v}],$$

легко привести выражение для векторного потенциала к виду

$$\mathbf{A} = \frac{\mu_0}{4\pi} \left(\frac{\ddot{\mathbf{d}}}{r} + \frac{\ddot{\mathbf{D}}}{6cr} + \frac{[\dot{\mathbf{m}}\mathbf{n}]}{cr} \right),$$

где \mathbf{D} — вектор с компонентами $D_i = D_{ij}n_j$.

Зная \mathbf{A} , можно найти поля и, следовательно, интенсивность излучения. Приведем лишь окончательную формулу

$$\mathcal{J} = \frac{1}{4\pi\epsilon_0} \left(\frac{2}{3c^3} \ddot{\mathbf{d}}^2 + \frac{1}{180c^5} \ddot{D}_{ij}^2 \right) + \frac{\mu_0}{6\pi c^3} \dot{\mathbf{m}}^2.$$

Первое слагаемое в правой части — уже знакомая нам формула для интенсивности дипольного излучения; второе и третье слагаемые определяют интенсивности соответственно *квадрупольного* (\mathcal{J}_Q) и *магнитно-дипольного* (\mathcal{J}_m) излучений. Заметим, что магнитно-дипольное излучение отсутствует у любой системы, состоящей только из двух частиц; нетрудно показать, что у системы из двух частиц магнитный момент пропорционален механическому и, следовательно, сохраняется ($\dot{\mathbf{m}} = 0$).

23.4. Антенна

Результаты § 23.1, 23.3 позволяют объяснить принцип действия антенны, излучающей радиоволны. Антенна представляет собой линейный проводник, вдоль которого возбуждается переменный ток с определенным распределением по проводнику. Мы конкретно рассмотрим линейную антенну с возбуждением в центре, например, с помощью коаксиальной линии (рис. 23.2). Если возбуждение происходит на определенной частоте ω , то распределение тока I вдоль антенны (ось z) синусоидальное, причем на концах антенны ($z = \pm d/2$) ток равен нулю:

$$I(z, t) = I_0 \sin\left(\frac{1}{2}kd - k|z|\right) \cdot e^{-i\omega t},$$

где $k = \omega/c$ и I_0 — максимальная сила тока в антенне, определяемая амплитудой тока в зазоре I_m : $I_0 = I_m \sin^{-1}(kd/2)$.

Найдем электромагнитное поле, порождаемое антенной. Преобразуем предварительно выражение для векторного потенциала применительно к случаю линейного тока. Для этого следует в общем выражении для \mathbf{A} положить $\mathbf{j}dV = Idl$, где dl — элемент длины линейного проводника и I — сила тока в нем. Кроме того, интересуясь полем вдали от антенны, мы должны в выражении для

$$t' = t - \frac{r}{c} + \frac{\mathbf{r}'\mathbf{n}}{c}$$

заменить внутреннее запаздывание $\mathbf{r}'\mathbf{n}/c$ на $z\cos\theta/c$, где z — координата точки антенны и θ — угол между осью z и направлением, в котором вычисляется векторный потенциал (\mathbf{n} — единичный вектор в этом направлении, φ — полярный угол). В результате мы получим следующее выражение для A_z — единственной отличной от нуля компоненты векторного потенциала:

$$A_z(\mathbf{r}, t) = \frac{\mu_0}{4\pi r} \int I \left(z, t - \frac{r}{c} + \frac{z}{c} \cos\theta \right) dz.$$

Подставляя сюда приведенное выражение для силы тока в антенне, получим

$$A_z(\mathbf{r}, t) = \frac{I_0 \mu_0}{4\pi r} e^{-i\omega(t-r/c)} \times \\ \times \int_{-d/2}^{+d/2} \sin\left(\frac{kd}{2} - k|z|\right) e^{ikz \cos\theta} dz.$$

Вычисление интеграла производится элементарно и дает

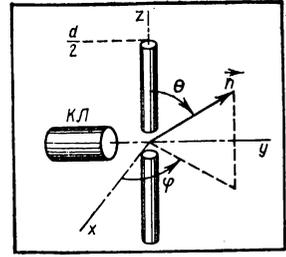
$$A_z(\mathbf{r}, t) = \frac{\mu_0 I_0 e^{i(kr-\omega t)}}{2\pi kr \sin^2\theta} \left\{ \cos\left(\frac{1}{2} kd \cos\theta\right) - \cos\frac{kd}{2} \right\}.$$

Вдали от антенны — в волновой зоне, где $kr \gg 1$, индукция магнитного поля связана с \mathbf{A} соотношением $\mathbf{B} = ik[\mathbf{n}\mathbf{A}]$, поэтому $B = k|A_z| \sin\theta$. Напряженность электрического поля $\mathbf{E} = c[\mathbf{B}\mathbf{n}]$ направлена перпендикулярно \mathbf{B} . Отсюда легко найти вектор Пойнтинга и интенсивность $d\mathcal{J}$ излучения в элемент телесного угла do около направления \mathbf{n} :

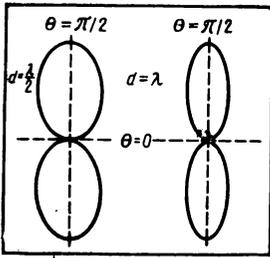
$$d\mathcal{J} = \frac{I_0^2}{8\pi^2 \epsilon_0 c \sin^2\theta} \left\{ \cos\left(\frac{1}{2} kd \cos\theta\right) - \cos\frac{kd}{2} \right\}^2 do.$$

Так как $\mathbf{E} = -ikc[\mathbf{n}[\mathbf{n}\mathbf{A}]] = -ikc\{\mathbf{n}(\mathbf{n}\mathbf{A}) - \mathbf{A}\}$, то излучение поляризовано в плоскости, проходящей через антенну и точку наблюдения.

В случае длинных волн, когда $kd \ll 1$, формула для $d\mathcal{J}$ приобретает вид



23.2. Линейная антенна с возбуждением в центре (КЛ — коаксиальная линия)



23.3.

Диаграммы направленности для полуволновой и полноволновой антенн

$$d\mathcal{J} = \frac{I_m^2}{128\pi^2\epsilon_0 c} (kd)^2 \sin^2 \theta \, d\theta, \quad I_m = I_0 \sin(kd/2).$$

Эта формула соответствует дипольному приближению, так как предполагает, что длина излучаемой волны $\lambda \gg d$. Полная интенсивность излучения получается отсюда интегрированием по $d\theta$:

$$\mathcal{J} = I_m^2 (kd)^2 / (48\pi\epsilon_0 c).$$

Как мы видим, интенсивность излучения в случае длинных волн пропорциональна квадрату отношения размеров антенны к длине волны.

Обратимся теперь к общей формуле для $d\mathcal{J}$, справедливой при любом значении kd . Наиболее интересны два случая: когда вдоль антенны располагается полволны или целая волна, т. е. когда $kd = \pi$ и $kd = 2\pi$. В этих случаях

$$\frac{d\mathcal{J}}{d\theta} = \frac{I_0^2}{8\pi^2\epsilon_0 c \sin^2\theta} \begin{cases} \cos^2 \frac{\pi \cos \theta}{2} & (kd = \pi), \\ 4\cos^4 \frac{\pi \cos \theta}{2} & (kd = 2\pi). \end{cases}$$

Определяемые этими формулами угловые зависимости $d\mathcal{J}/d\theta$, т. е. диаграммы направленности антенны, изображены на рис. 23.3. Полные интенсивности излучения даются формулами

$$\mathcal{J} = \frac{I_0^2}{8\pi\epsilon_0 c} \begin{cases} 2,44 & (kd = \pi), \\ 6,70 & (kd = 2\pi). \end{cases}$$

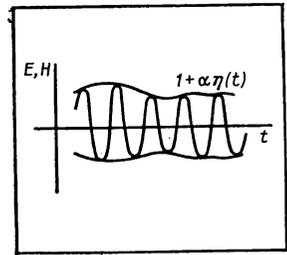
Мы видим, что интенсивность излучения полноволновой антенны $d = \lambda$ примерно в три раза больше интенсивности излучения полуволновой антенны ($d = \lambda/2$).

В заключение этого параграфа остановимся кратко на вопросе о передаче с помощью антенны сигналов, в частности человеческой речи. Обычно с этой целью модулируют амплитуду силы тока в антенне I_0 в соответствии с передаваемым сигналом. Если, например, сигнал как функция времени есть $\eta(t)$, то амплитуда тока в антенне I_0 должна меняться со временем в такт с сигналом, т. е. иметь вид

$$I_0 \equiv I_0(t) = \bar{I}_0 [1 + \alpha\eta(t)],$$

где \bar{I}_0 — постоянная и α — некоторое число, меньшее единицы (коэффициент модуляции). Ясно, что если все частоты сигнала значительно меньше частоты ω возбуждения антенны («несущей частоты»), то модуляция тока в антенне не скажется на характере излучения. Так как поле, порождаемое током, пропорционально току, то в месте наблюдения поле несет в себе информацию о сигнале, а именно поле как

функция времени имеет вид, изображенный на рис. 23.4. Здесь узкие синусоиды соответствуют частоте ω , огибающая же их кривая — зависимости $1 + \alpha\eta(t)$. Задача состоит теперь в том, чтобы сохранить только эту огибающую. Для этого поле в месте наблюдения предварительно усиливается (с помощью усилителя высокой частоты ω), а затем производится детектирование (буквально — обнаружение сигнала), т. е. срезание нижней половины кривой $E = E(t)$ каким-либо выпрямляющим устройством. После детектирования поле усиливается на низких частотах и подается на прибор, регистрирующий такие частоты, — телефон или телеграф. Мы не будем здесь рассматривать ни деталей детектирования, ни работы приемной антенны в месте наблюдения.



23.4. Модуляция высокочастотной волны

23.5. Излучение релятивистского заряда

В § 23.3 мы нашли интенсивность дипольного излучения движущейся заряженной частицы. Полученные там формулы справедливы только в нерелятивистском случае, когда $v \ll c$. Найдём теперь интенсивность излучения частицы, скорость которой сравнима со скоростью света. Это нетрудно сделать, если воспользоваться готовыми формулами для излучения нерелятивистской частицы. В самом деле, с какой бы скоростью частица ни двигалась, всегда можно перейти в такую систему отсчета, в которой она в заданный момент времени покоится, а в этой системе интенсивность излучения нам известна. Поэтому вопрос сводится к тому, как преобразуется в соответствии с преобразованиями Лоренца интенсивность излучения при переходе от одной системы отсчета к другой.

Для ответа на этот вопрос перепишем формулу для интенсивности излучения в виде

$$d\mathcal{E} = \frac{e^2}{6\pi\epsilon_0 c^3} \mathbf{w}^2 dt,$$

где $d\mathcal{E}$ — энергия, излучаемая за время dt . Вспомним теперь, что время и энергия преобразуются одинаково, поэтому квадрат ускорения есть релятивистский инвариант и его нужно лишь выразить через величины, относящиеся к исходной системе отсчета (в которой частица двигалась с большой скоростью). Замечая, что

$$\mathbf{w}^2 = \frac{1}{m^2} \left(\frac{d\mathbf{p}}{dt} \right)^2,$$

где \mathbf{p} — импульс и m — масса частицы, достаточно сделать замену $(d\mathbf{p}/dt) \rightarrow (d\mathbf{p}_\mu/d\tau)^2$, где $\mathbf{p}_\mu \equiv (\mathbf{p}, i\mathcal{E}/c)$ — 4-вектор энергии-импульса частицы и $d\tau = dt \sqrt{1 - v^2/c^2}$ — приращение собственного времени частицы, соответствующее интервалу dt в исходной системе отсчета.

Таким образом, общая формула для $d\mathcal{E}/dt$, справедливая при любых скоростях частицы, имеет вид

$$\frac{d\mathcal{E}}{dt} = \frac{e^2}{6\pi\epsilon_0 m^2 c^3} \left(\frac{d\rho_\mu}{dt} \right)^2 = \frac{e^2}{6\pi\epsilon_0 m^2 c^3} \left[\left(\frac{d\mathbf{p}}{dt} \right)^2 - \frac{1}{c^2} \left(\frac{d\mathcal{E}}{dt} \right)^2 \right],$$

или, если выразить все величины через скорость \mathbf{v} и ускорение $\mathbf{w} = d\mathbf{v}/dt$, — вид

$$\frac{d\mathcal{E}}{dt} = \frac{e^2}{6\pi\epsilon_0 c^3} \frac{w^2 - [\mathbf{v} \mathbf{w}]^2/c^2}{(1 - v^2/c^2)^3}.$$

Полная излученная энергия получается интегрированием этого выражения по времени:

$$\Delta\mathcal{E} = \frac{e^2}{6\pi\epsilon_0 c^3} \int_{-\infty}^{\infty} \frac{w^2 - [\mathbf{v} \mathbf{w}]^2/c^2}{(1 - v^2/c^2)^3} dt.$$

Если частица движется в электромагнитном поле, то как мы видели в гл. 7,

$$\mathbf{w} = \frac{e}{m} \sqrt{1 - \frac{v^2}{c^2}} \left\{ \mathbf{E} + [\mathbf{v} \mathbf{B}] - \frac{1}{c^2} \mathbf{v} (\mathbf{v} \mathbf{E}) \right\}$$

и поэтому

$$\Delta\mathcal{E} = \frac{e^2}{6\pi\epsilon_0 m^2 c^3} \int_{-\infty}^{\infty} \left\{ \left(\mathbf{E} + [\mathbf{v} \mathbf{B}] - \frac{1}{c^2} (\mathbf{E} \mathbf{v}) \right)^2 \frac{dt}{1 - v^2/c^2} \right\}.$$

Используя эти формулы, можно найти излучение электрона, ускоряемого в линейном и циклическом ускорителях.

При движении электрона в однородном магнитном поле с индукцией \mathbf{B} радиус орбиты и частота вращения определяются формулами:

$$r = \frac{m_e v}{eB \sqrt{1 - v^2/c^2}}, \quad \omega_B = \frac{v}{r} = \frac{eB}{m_e} \sqrt{1 - \frac{v^2}{c^2}},$$

а интенсивность излучения — формулой

$$\frac{d\mathcal{E}}{dt} = \frac{e^4 B^2 v^2}{6\pi\epsilon_0 m_e^2 c^3 (1 - v^2/c^2)}.$$

В ультрарелятивистском случае эту формулу удобно переписать в виде

$$\frac{d\mathcal{E}}{dt} = \frac{e^4 B^2}{6\pi\epsilon_0 m_e^2 c} \left(\frac{\mathcal{E}}{m_e c^2} \right)^2.$$

Интенсивность излучения в этом случае пропорциональна квадрату энергии частицы \mathcal{E} . Излучение сосредоточено в основном вблизи плоскости орбиты электрона в области углов $\Delta\theta \sim mc^2/\mathcal{E}$. Спектраль-

ное распределение излучения имеет резкий максимум, причем излучение сосредоточено в основном в области частот

$$\omega \sim \omega_B \left(\frac{\mathcal{E}}{m_e c^2} \right) = \frac{eB}{m_e} \left(\frac{\mathcal{E}}{m_e c^2} \right)^2.$$

23.6. Сила лучистого трения

Если рассматривать излучение электромагнитной энергии как непрерывный процесс (а такое рассмотрение вытекает из уравнений Максвелла), то потерю энергии частицы на излучение электромагнитных волн можно феноменологически трактовать как результат действия некоторой силы, аналогичной силе трения и называемой *силой лучистого трения*. Обозначая эту силу \mathbf{f}_r , мы можем сказать, что работа $\int_{t_1}^{t_2} \mathbf{f}_r \mathbf{v} dt$, совершенная этой силой в интервале времени (t_1, t_2) , должна быть равна разности энергий $W_2 - W_1$ частицы в эти моменты времени (\mathbf{v} — скорость частицы). Замечая, что

$$W_2 - W_1 = \int_{t_1}^{t_2} \mathcal{J} dt,$$

где $\mathcal{J} = e^2 \mathbf{w}^2 / (6\pi\epsilon_0 c^3)$ — интенсивность излучения, получим следующее уравнение для определения \mathbf{f}_r :

$$\int_{t_1}^{t_2} \mathbf{f}_r \mathbf{v} dt = - \frac{e}{6\pi\epsilon_0 c^3} \int_{t_1}^{t_2} \mathbf{w}^2 dt.$$

Чтобы определить \mathbf{f}_r , сделаем предположение, что излучение оказывает слабое влияние на движение частицы, т. е. может рассматриваться как малое возмущение. В случае периодического (в пренебрежении лучистым трением) движения с частотой ω_0 это значит, что для $t_2 = t_1 + 2\pi/\omega_0$ должны с большей точностью выполняться равенства $\mathbf{v}(t_2) = \mathbf{v}(t_1)$, $\mathbf{w}(t_2) = \mathbf{w}(t_1)$. Поэтому если мы преобразуем по частям интеграл

$$\int_{t_1}^{t_2} \mathbf{w}^2 dt = \int_{\mathbf{v}(t_1)}^{\mathbf{v}(t_2)} \mathbf{w} d\mathbf{v} = \mathbf{w} \mathbf{v} \Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} \frac{d\mathbf{w}}{dt} \mathbf{v} dt,$$

то внеинтегральное слагаемое исчезает, так что уравнение для \mathbf{f}_r примет вид

$$\int_{t_1}^{t_2} \left(\mathbf{f}_r - \frac{e^2}{6\pi\epsilon_0 c^3} \frac{d\mathbf{w}}{dt} \right) \mathbf{v} dt = 0.$$

Этот интеграл должен тождественно обращаться в нуль для любых $\mathbf{v}(t)$, так что

$$\mathbf{f}_r = \frac{e^2}{6\pi\epsilon_0 c^3} \frac{d\mathbf{w}}{dt}.$$

Это и есть окончательное выражение для силы лучистого трения. Обратим внимание на то, что

эта сила пропорциональна производной от ускорения по времени.

Подчеркнем, что этим выражением можно пользоваться только в том случае, когда сила $\mathbf{f}(r)$ мала по сравнению с другими силами, действующими на частицу. В этом можно убедиться на простом примере, рассмотрев действие силы лучистого трения на свободную частицу. Уравнение движения в этом случае имело бы вид

$$m\mathbf{w} = \frac{e^2}{6\pi\epsilon_0 c^3} \frac{d\mathbf{w}}{dt},$$

откуда

$$\mathbf{w} = \mathbf{w}_0 e^{\frac{3c}{2r_0} t}.$$

где \mathbf{w}_0 — ускорение в начальный момент времени и $r_0 = e^2/(4\pi\epsilon_0 m_e^2 c^2)$ — классический радиус электрона. Мы видим, что находящаяся под действием одной только силы лучистого трения частица должна была бы неограниченно самоускоряться, несмотря на потери энергии на излучение. Абсурдность этого вывода указывает на то, что выражением для силы лучистого трения следует пользоваться с осторожностью и только в том случае, когда эта сила мала по сравнению с другими силами, действующими на частицу.

В следующем параграфе при рассмотрении действия силы лучистого трения на гармонический осциллятор будет получен

один из критериев применимости понятия лучистого трения: $\lambda \gg r_0$, где λ — длина излучаемой волны и r_0 — классический радиус электрона. При длинах волн, меньших r_0 (или сравнимых с r_0), классическая электродинамика приводит к противоречию и, следовательно, неприменима.

Еще один критерий применимости классической электродинамики можно получить, рассматривая действие силы лучистого трения на частоту в электромагнитном поле. Уравнение движения такой частицы имеет вид

$$m \dot{\mathbf{v}} = e\mathbf{E} + e[\mathbf{v}\mathbf{B}] + \frac{e^2}{6\pi\epsilon_0 c^3} \ddot{\mathbf{v}}.$$

В системе отсчета, в которой заряд в данный момент покоится, имеем (в пренебрежении силой лучистого трения)

$$\ddot{\mathbf{v}} = \frac{e}{m} \dot{\mathbf{E}} + \frac{e}{m} [\dot{\mathbf{v}} \mathbf{B}] = \frac{e}{m} \left(\dot{\mathbf{E}} + \frac{e}{m} [\mathbf{E}\mathbf{B}] \right),$$

что позволяет переписать силу лучистого трения в виде

$$\mathbf{f}_r = \frac{e^3}{6\pi\epsilon_0 m c^3} \left(\dot{\mathbf{E}} + \frac{e}{m} [\mathbf{E}\mathbf{B}] \right).$$

Сравним второе слагаемое в круглых скобках с силой ϵE . Требуется, чтобы сила лучистого трения была малой поправкой, сведется при этом к неравенству

$$cB \ll e/(\epsilon_0 r_0^2).$$

Мы приходим к выводу, что

слишком интенсивные поля также не могут рассматриваться в классической электродинамике.

Рассмотрим теперь задачу, в которой сила f , мала по сравнению с другими силами, действующими на частицу, и использование понятия силы лучистого трения приводит к физически разумным результатам, — задачу о заряженном гармоническом осцилляторе. Если m — масса и ω_0 — собственная частота колебаний осциллятора, то радиус-вектор осциллятора при учете лучистого трения удовлетворяет уравнению

$$\ddot{\mathbf{r}} + \omega_0^2 \mathbf{r} = \frac{2r_0}{3c} \ddot{\mathbf{r}}.$$

Будем искать его решение в виде $\mathbf{r} = \mathbf{a}e^{-i\omega t}$, где \mathbf{a} и ω — начальная амплитуда и частота колебаний (вообще говоря, комплексные). Мы получим кубичное уравнение для ω :

$$\omega^3 = \omega_0^3 - \frac{2r_0}{3c} i\omega^3.$$

Из трех решений этого уравнения только два должны иметь физический смысл, а именно те, для которых ω^2 близко к ω_0^2 . Пренебрегая квадратичными по $\omega - \omega_0$ членами, найдем

$$\omega = \pm \omega_0 - i \frac{r_0}{3c} \omega_0^3 \equiv \omega_0 - i\gamma/2; \quad \gamma \equiv \frac{4\pi r_0}{3\lambda} \omega_0.$$

Таким образом, $\mathbf{r} = \mathbf{a} \exp\{-i\omega_0 t - \gamma t/2\}$, т. е. колебания медленно затухают по экспоненциальному закону. Мы видим, что в задаче о заряженном гармоническом осцилляторе лучистое трение проявляет себя полностью аналогично обычному трению.

23.7. Ширина спектральной линии

При наличии любого трения — не обязательно лучистого — осциллятор перестает совершать строго гармонические колебания и, следовательно, излучать строго монохроматические волны. Возникает вопрос: каким же будет спектр излучаемых таким осциллятором частот?

Чтобы ответить на этот вопрос, выразим полную излученную энергию $\mathcal{E} = \int_{-\infty}^{\infty} \mathcal{J} dt$, $\mathcal{J} = \dot{\mathbf{d}}^2 / (6\pi \epsilon_0 c^3)$ через фурье-компоненту дипольного момента осциллятора:

$$\ddot{d}_\omega = \int_{-\infty}^{\infty} e^{i\omega t} \ddot{d}(t) \frac{dt}{2\pi}.$$

Согласно формуле обращения Фурье,

$$\ddot{d}(t) = \int_{-\infty}^{\infty} e^{-i\omega t} \ddot{d}_\omega d\omega,$$

поэтому

$$\mathcal{E} \equiv \int_{-\infty}^{\infty} \frac{\ddot{d}^2}{6\pi\epsilon_0 c^3} dt = (6\pi\epsilon_0 c^3)^{-1} \int_{-\infty}^{\infty} dt d\omega d\omega' e^{-i(\omega-\omega')t} \ddot{d}_\omega \ddot{d}_{\omega'},$$

Вспомним, что

$$\int_{-\infty}^{\infty} e^{-i(\omega-\omega')t} dt = 2\pi\delta(\omega - \omega'),$$

где δ — уже неоднократно использованная нами δ -функция, равная нулю при $\omega \neq \omega'$ и бесконечности при $\omega = \omega'$ и нормированная так, что интеграл от нее равен единице. Появление этой функции позволяет выполнить интегрирование по ω' :

$$\mathcal{E} = \int d\mathcal{E}_\omega, \quad d\mathcal{E}_\omega = (3\epsilon_0 c^3)^{-1} |\ddot{d}_\omega|^2 d\omega.$$

Мы видим, что излучаемая энергия может быть представлена как интеграл по частотам от величины $|\ddot{d}_\omega|^2 / (3\epsilon_0 c^3)$, имеющей, таким образом, смысл *спектральной плотности излучаемой энергии*.

Спектральное распределение излучения обычно характеризуют так называемой *формой спектральной линии* $f(\omega)$ — отношением энергии, излучаемой в интервале частот $(\omega, \omega + d\omega)$, ко всей излучаемой энергии: $f(\omega)d\omega = d\mathcal{E}_\omega / \mathcal{E}$. Мы видим, что при дипольном излучении $f(\omega) = \text{const} |\ddot{d}_\omega|^2$, где const — нормировочная постоянная.

Найдем теперь фурье-компоненту \ddot{d}_ω для осциллятора, совершающего затухающие колебания. Если осциллятор начинает колебаться в момент $t = 0$, то согласно результатам предыдущего параграфа

$$\ddot{d} = -e a \omega_0^2 e^{-i\omega_0 t - \gamma t/2}.$$

Поэтому

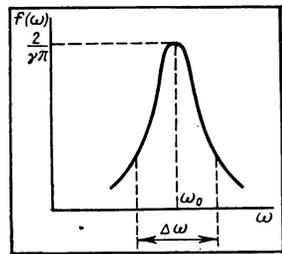
$$\ddot{d}_\omega = \frac{-e a \omega_0^2}{2\pi} \int_{-\infty}^{\infty} e^{-i(\omega_0 - \omega)t - \frac{\gamma}{2} t} dt = \frac{e a \omega_0^2 / (2\pi)}{-i(\omega_0 - \omega) - \gamma/2}.$$

Нормировочная постоянная в $f(\omega)$ находится без труда. Окончательно мы получим

$$f(\omega) = \frac{\gamma/(2\pi)}{(\omega - \omega_0)^2 + (\gamma/2)^2} \cdot$$

Мы видим, что слабозатухающий осциллятор излучает в узком интервале частот вблизи своей собственной частоты ω_0 ; ширина этого интервала $\Delta\omega$ равна коэффициенту затухания γ (рис. 23.5).

Если осциллятор тратит энергию только на излучение, так что из тормозящих его сил достаточно учитывать только лучистое трение, то возникает минимально возможная ширина спектральной линии $\gamma = 4\pi r_0 \omega_0 / (3\lambda)$. Эту величину называют *естественной шириной спектральной линии*.



23.5. Форма спектральной линии осциллятора

Естественная ширина спектральной линии обусловлена самим фактом излучения, т. е. взаимодействием излучающей системы с полем излучения.

Если излучающая система взаимодействует еще и с другими физическими объектами, то эти взаимодействия также приводят к уширению спектральных линий. Уширения, вызываемые различными механизмами взаимодействий, складываются так, что если $\Delta\omega_i$ — эффективная ширина линии, обусловленная i -м механизмом, то полная ширина линии $\Delta\omega = \sum_i \Delta\omega_i$.

Каждая из частичных или парциальных ширин $\Delta\omega_i$ просто связана с эффективным временем жизни τ_i излучающей системы, соответствующим i -му механизму взаимодействия. Так, время жизни осциллятора, затухание колебаний которого обусловлено самим фактом излучения, $\tau_r = 1/\gamma$, а естественная ширина линии $\Delta\omega = \gamma = 1/\tau_r$. Аналогичные соотношения справедливы и для других механизмов уширения, т. е. $\Delta\omega_i = 1/\tau_i$, поэтому

$$\Delta\omega = \sum (1/\tau_i).$$

Различным механизмам уширения соответствуют различные формы спектральной линии, не совпадающие, вообще говоря, с полученной выше естественной формой линии.

Рассмотрим подробнее уширение линии, обусловленное столкновениями излучающих систем, например осцилляторов друг с другом. Если бы столкновений не было, то осциллятор излучал бы монохроматическую волну частоты ω_0 , т. е. $E(t) = E_0 \exp(-i\omega t)$. Вследствие столкновений поле будет таким только в некотором интервале времени $(0, \tau)$, после чего последует столкновение, причем интервалы τ подчинены статистическому закону. Мы предположим, что вероятность $\omega(\tau)d\tau$ того, что интервал времени до соударения лежит между τ и $\tau + d\tau$, определяется формулой

$$\omega(\tau) d\tau = \frac{1}{\tau_c} \exp(-\tau/\tau_c),$$

где τ_0 — среднее значение τ (напомним, что $\int \omega(\tau) d\tau = 1$).
 Определим теперь фурье-компоненту поля излучения \mathbf{E}_ω :

$$\mathbf{E}_\omega = \int_{-\infty}^{\infty} \mathbf{E}(t) e^{i\omega t} \frac{dt}{2\pi} = \mathbf{E}_0 \int_0^{\tau} e^{i(\omega - \omega_0)t} \frac{dt}{2\pi} = \frac{\mathbf{E}_0}{2\pi} \frac{e^{i(\omega - \omega_0)\tau} - 1}{i(\omega - \omega_0)}.$$

Образовав $|\mathbf{E}_\omega|^2$, найдем интенсивность излучения. Нам остается усреднить последнюю величину по τ . В результате получим

$$\langle |\mathbf{E}_\omega|^2 \rangle = \int_0^{\infty} |\mathbf{E}_\omega|^2 \omega(\tau) d\tau = \left(\frac{E_0}{2\pi}\right)^2 \frac{1}{(\omega - \omega_0)^2 + 1/\tau_c^2}.$$

Мы видим, что форма линии, обусловленная столкновениями, совпадает с естественной формой; шириной линии оказывается величина $\gamma_c = 1/\tau_c$.

ОСНОВНЫЕ ФОРМУЛЫ

Исходные уравнения с источниками

$$\Delta \varphi - \frac{1}{c^2} \frac{\partial^2 \varphi}{\partial t^2} = -\rho/\epsilon_0,$$

$$\Delta \mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu_0 \mathbf{j}$$

Запаздывающие потенциалы

$$\varphi(\mathbf{r}, t) = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\mathbf{r}', t - \frac{1}{c}|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|} dV',$$

$$\mathbf{A}(\mathbf{r}, t) = \frac{\mu_0}{4\pi} \int \frac{\mathbf{j}(\mathbf{r}', t - \frac{1}{c}|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|} dV'$$

Дифференциальная интенсивность дипольного излучения

$$d\mathcal{J} = \frac{1}{(4\pi)^2 \epsilon_0 c^3} \left[\ddot{\mathbf{d}} \cdot \mathbf{n} \right]^2 d\Omega$$

Интегральная интенсивность дипольного излучения

$$\mathcal{J} = \frac{\ddot{\mathbf{d}}^2}{6\pi\epsilon_0 c^3}$$

Интенсивность квадрупольного излучения

$$\mathcal{J}_Q = \frac{1}{4\pi\epsilon_0} \frac{\ddot{D}_{ij}^2}{180c^5}, \quad D_{ij} = \Sigma e(3x_i x_j - r^2 \delta_{ij})$$

Интенсивность магнитно-дипольного излучения

$$\mathcal{J}_m = \frac{\mu_0 \dot{\mathbf{m}}^2}{6\pi c^3}, \quad \mathbf{m} = \frac{1}{2} \Sigma e[\mathbf{r}\mathbf{v}]$$

Интенсивность излучения антенны ($\lambda \gg d$)

$$\mathcal{J} = \frac{I_m^2}{48\pi\epsilon_0 c} (kd)^2$$

Энергия, излучаемая за 1 с частицей, движущейся в магнитном поле

$$\frac{d\delta}{dt} = \frac{e^4 B^2 v^2}{6\pi\epsilon_0 m_e^2 c^3 (1 - v^2/c^2)}$$

Частота максимума спектрального распределения излучения

$$\omega \sim \frac{eB}{m_e} \left(\frac{\delta}{m_e c^2} \right)^2$$

Естественная форма спектральной линии

$$f(\omega) = \frac{\gamma/(2\pi)}{(\omega - \omega_0)^2 + (\gamma/2)^2}, \quad \gamma = \frac{4\pi}{3} \frac{r_0 \omega_0}{\lambda}$$



Глава 24. ВЗАИМОДЕЙСТВИЕ ЗАРЯЖЕННЫХ ЧАСТИЦ И ИЗЛУЧЕНИЯ С ВЕЩЕСТВОМ

24.1. Рассеяние частиц в кулоновском поле ядер

При прохождении заряженных частиц и фотонов через вещество возникает целый ряд эффектов, обусловленных электромагнитным взаимодействием. Рассмотрение их мы начнем с упругого рассеяния заряженных частиц, вызываемого кулоновским полем ядер вещества*.

Если пренебречь экранирующим действием атомных электронов, расположенных вокруг ядра, то сила, действующая на частицу, равна Zze^2/r^2 , где Ze и ze' — заряды соответственно ядра и частицы и r — расстояние между ними. Массу ядра будем предполагать значительно большей массы частицы M . В этом случае ядро можно считать покоящимся, и частица движется по гиперболе, в одном из фокусов которой находится ядро, причем угол рассеяния θ связан с прицельным параметром b соотношением

$$\operatorname{tg} \frac{\theta}{2} = \frac{Zze^2}{Mbv^2},$$

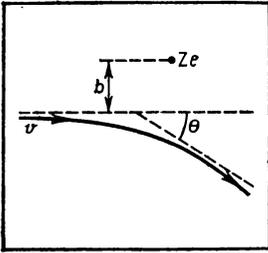
где v — скорость частицы (рис. 24.1).

Эта формула справедлива в нерелятивистской механике, когда $v \ll c$. При малых углах рассеяния она приобретает вид

$$\theta = \frac{2Zze^2}{Mbv^2}.$$

Физический смысл последней формулы легко понять, если учесть, что $\theta \sim \Delta p/p$, т. е. угол рассеяния пропорционален отношению импульса Δp , теряемого частицей (и направленного перпендикулярно первоначальному импульсу частицы p), к значению p . Очевидно,

* В этой главе мы используем ГС.



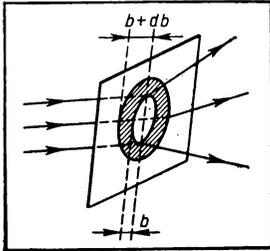
24.1.
Траектория частицы в поле ядра

$\Delta p = \int zeE_{\perp} dt$, где $E_{\perp} \sim Ze^2/b^2$ — перпендикулярная составляющая напряженности поля, создаваемого ядром, а эффективный интервал времени, который существует в интеграле, — время столкновения — по порядку величины составляет $\tau \sim b/v$; поэтому

$$\Delta p \sim zeE_{\perp}\tau \sim Zze^2/(bv).$$

Точный результат $\Delta p = 2Zze^2/(bv)$ отличается от этой оценки множителем 2. Разделив Δp на p , найдем угол рассеяния:

$$\theta = 2Zze^2/(pbv).$$



24.2.
К выводу формулы для дифференциального сечения рассеяния

Эта формула справедлива при всех значениях скорости частицы, в том числе и в релятивистском случае ($v \sim c$). При нерелятивистских скоростях $p = Mv$ и мы приходим к приведенной выше формуле для угла рассеяния.

Рассеяние характеризуют обычно *дифференциальным сечением* $d\sigma$, которое определяется как

$$d\sigma = dN/I,$$

где dN — число частиц, рассеянных за 1 с в интервал углов $(\theta, \theta + d\theta)$, и I — плотность потока падающих частиц, т. е. число частиц, проходящих за 1 с на единичную площадку, перпендикулярную их первоначальной скорости. Так как приличный параметр b однозначно связан с углом рассеяния, то, очевидно, все частицы, падающие на кольцо с радиусами $b, b + db$, рассеются в интересующий нас интервал углов (рис. 24.2). Поэтому

$$d\sigma = 2\pi b db.$$

Используя приведенную связь между b и θ , получим отсюда

$$d\sigma = \left(\frac{Zze^2}{2Mv^2} \right)^2 \frac{d\theta}{\sin^4(\theta/2)},$$

где $d\theta = 2\pi \sin\theta d\theta$ — элемент телесного угла, в который происходит рассеяние. Эта формула носит название *формулы Резерфорда*. Обратим внимание на то, что

при малых углах рассеяния $d\sigma \sim d\theta/\theta^3$ это значит, что рассеяние в кулоновском поле происходит главным образом на малые углы.

Сечение рассеяния $d\sigma$ характеризует индивидуальный акт рассеяния. Между тем при прохождении частицы через вещество, например через фольгу, происходит много актов рассеяния. Угол результирующего отклонения при многократном рассеянии является статистической суммой малых углов отклонения при индивидуальных

актах рассеяния, так как эти углы можно считать случайными величинами. Поэтому распределение результирующего отклонения по углам является гауссовым, т. е. вероятность рассеяния в интервал углов, $(\theta, \theta + d\theta)$ определяется формулой

$$P(\theta) d\theta = C \exp\left\{-\frac{\theta^2}{2 \langle \theta^2 \rangle}\right\} d\theta,$$

где $\langle \theta^2 \rangle$ — средний квадрат угла многократного рассеяния и C — постоянная, определяемая из условия нормировки $\int P(\theta) d\theta = 1$.

Найдем теперь $\langle \theta^2 \rangle$. Так как отдельные столкновения статистически независимы, то эта величина совпадает со средним значением квадрата угла рассеяния при индивидуальном столкновении — средним по различным траекториям, т. е. по различным прицельным параметрам. Если N — плотность ядер и l — длина пути частицы (в случае тонкой фольги — ее толщина), то число столкновений в объеме полого цилиндра высоты l , основание которого — кольцо с радиусами b и $b + db$, равно $N l 2\pi b db$, потому $\langle \theta^2 \rangle^2$ определяется формулой

$$\langle \theta^2 \rangle = N l \int_{b_{\min}}^{b_{\max}} \theta^2(b) 2\pi b db = \frac{8\pi N l Z^2 z^2 e^4}{\rho^2 v^2} \ln \frac{b_{\max}}{b_{\min}}.$$

Пределы интегрирования по b находятся из следующих соображений. Если бы заряд ядра полностью экранировался при $b > b_{\max}$ и вовсе не экранировался при $b < b_{\max}$, то величина b_{\max} определялась бы абсолютно строго; такой резкой границы, однако, нет, поэтому в качестве b_{\max} берут радиус атома $a_0 Z^{-1/3}$, где $a_0 = \hbar^2 / (m_e e^2)$ — борковский радиус и m_e — масса электрона.

Величина b_{\min} определяется максимальным углом отклонения при индивидуальном акте рассеяния. При рассеянии в кулоновском поле существенны только малые углы рассеяния, а так как формула для $\langle \theta^2 \rangle$ мало чувствительна к выбору b_{\min} (входящего только под знаком логарифма), то при оценке можно считать, что b_{\min} отвечает углу рассеяния $\theta \sim 1$, т. е. $b_{\min} = 2Zze^2 / (pv)$. Подставляя эти значения b_{\max} и b_{\min} в формулу для $\langle \theta^2 \rangle$, получим

$$\langle \theta^2 \rangle = \frac{8\pi N l Z^2 z^2 e^4}{\rho^2 v^2} \ln \frac{a_0 p v}{2Z^{4/3} z e^2}.$$

24.2. Ионизационные потери энергии

При прохождении через вещество заряженная частица не только испытывает упругое рассеяние, но и теряет энергию. При этом играют роль два основных механизма: во-первых, частица передает энергию атомным электронам, которые либо переходят в возбужденные состояния, либо покидают атомы, и, во-вторых, двигаясь с ускорением, частица испускает электромагнитные волны, уносящие ее энергию. О первом механизме говорят как об ионизационных потерях энергии (хотя может происходить только возбуждение, а не ионизация атома), второй же механизм называют тормозным излучением.

Начнем с рассмотрения ионизационных потерь. Пусть на расстоянии b от траектории частицы, обладающей зарядом ze , массой M и движущейся со скоростью v , находится электрон, скорость которого v_e будем предполагать малой по сравнению с v во все время взаимодействия частиц. Для того чтобы найти энергию, передаваемую частицей электрону, нужно знать, какой импульс передается при столкновении, а с такой задачей мы уже сталкивались в предыдущем параграфе. Вспоминая, что передаваемый электрону импульс $\Delta p = 2ze^2/(bv)$, получим для передаваемой электрону энергии выражение

$$\frac{(\Delta p)^2}{2m_e} = \frac{2z^2 e^4}{m_e b^2 v^2}.$$

Умножим энергию, передаваемую при отдельном столкновении, на число столкновений частицы с электронами. На единичном пути частица испытывает $2\pi n b db$ столкновений, при которых прицельный параметр заключен в интервале $(b, b + db)$, поэтому потери энергии частицы на единичном пути составят

$$-\frac{dE}{dx} = 2\pi n \int_{b_{\max}}^{b_{\min}} \frac{(\Delta p)^2}{2m_e} b db = \frac{4\pi z^2 e^4 n}{m_e v^2} \ln \frac{b_{\max}}{b_{\min}},$$

где n — плотность электронов, а b_{\max} и b_{\min} — пределы интегрирования, которые выбираются теперь из следующих соображений.

Как мы видели в предыдущем параграфе, эффективное время, в течение которого поле частицы действует на электрон, составляет $\tau \sim b/v$. Эта оценка относится к нерелятивистскому случаю; если же $v \sim c$, то продолжительность столкновения уменьшается до $\tau \sim (b/v)\sqrt{1-v^2/c^2}$. Для того чтобы наш расчет потерь был правильным, электрон среды должен рассматриваться как свободный, а для этого должно выполняться неравенство $\tau \ll T$, где T — период колебаний электрона в атоме. Точный расчет показывает, что если это неравенство не выполняется, то энергия, приобретаемая электроном, пренебрежимо мала — она пропорциональна экспоненциальному множителю $\exp(-\omega\tau)$, где $\omega = 2\pi/T$ — частота колебаний электрона. Таким образом,

$$b_{\max} = \frac{v}{\langle \omega \rangle \sqrt{1 - v^2/c^2}},$$

где $\langle \omega \rangle$ — средняя частота колебаний электронов в атомах вещества.

Нижний предел интегрирования b_{\min} следует считать равным длине волны де Бройля $\lambda = h/p$, где $p = m_e v / \sqrt{1 - v^2/c^2}$ — импульс электрона, так как введение прицельного параметра, рассматриваемого как классическая величина, имеет смысл только при выполнении неравенства $b > \lambda$. (Строго говоря, в качестве v в эти формулы входит скорость электрона в системе центра масс сталкивающихся частиц, а в качестве m_e — их приведенная масса. Интересуясь тяжелой и быст-

рой пролетающей частицей, мы отождествляем v со скоростью последней и m с массой электрона). Таким образом,

$$b_{\min} = \frac{\hbar}{mv} \sqrt{1 - v^2/c^2}$$

и формула для потерь энергии приобретает окончательно вид

$$-\frac{dE}{dx} = \frac{4\pi z^2 e^4 n}{mv^2} \ln \frac{mv^2}{\hbar \langle \omega \rangle (1 - v^2/c^2)}.$$

Эту формулу называют формулой Бора. Более точный расчет приводит к формуле Бете:

$$-\frac{dE}{dx} = \frac{4\pi z^2 e^4 n}{m_e v^2} \left(\ln \frac{2m_e v^2}{\langle I \rangle \left(1 - \frac{v^2}{c^2}\right)} - \frac{v^2}{c^2} \right),$$

где $\langle I \rangle$ — средний ионизационный потенциал атомов вещества (равный приблизительно $I \sim 13,5 Z$ эВ; Z — атомный номер).

Приведенные формулы справедливы для тяжелых частиц, но не для электронов, так как при их выводе считалось, что частица, проходящая через вещество, практически не отклоняется. Если проходящей частицей является электрон, то расчет потерь энергии должен проводиться на основе квантовой механики. Мы приведем только результат, принадлежащий Бете:

$$-\frac{dE}{dx} = \frac{2\pi e^4 n}{m_e v^2} \left[\ln \frac{mv^2 K}{2 \langle I^2 \rangle \left(1 - \frac{v^2}{c^2}\right)} - \ln 2 \left(2 \sqrt{1 - \frac{v^2}{c^2}} - 1 + \frac{v^2}{c^2} \right) + 1 - \frac{v^2}{c^2} \right],$$

где

$$K = \frac{m_e c^2}{\sqrt{1 - v^2/c^2}} - m_e c^2.$$

Определив потери энергии частицы, можно найти ее длину пробега R , т. е. путь, проходимый ею до полной остановки:

$$R = \int_{E_0}^0 dE \left(\frac{dE}{dx} \right)^{-1},$$

здесь E_0 — начальная энергия частицы. Приблизительно R ведет себя как $R \sim E^{3/2}$. Для α -частиц, движущихся в воздухе при 15°C , R приблизительно определяется формулой $R_\alpha \approx 0,32 E^{3/2}$ см, где E выражено в МэВ.

Приведем значения $-dE/d\xi \equiv -\rho^{-1} dE/dx$ (ρ — плотность вещества) для электронов, движущихся в воздухе:

$E, \text{эВ}$	$-\frac{dE}{d\xi}, \text{МэВ}/(\text{г}\cdot\text{см}^{-2})$	$E, \text{эВ}$	$-\frac{dE}{d\xi}, \text{МэВ}/(\text{г}\cdot\text{см}^{-2})$
10^4	19,5	10^8	2,47
10^5	3,67	10^9	2,79
10^6	1,69	10^{10}	3,48
10^7	1,95		

Данные для тяжелых частиц, движущихся в воздухе и свинце, приведены в следующей таблице:

Частицы	$E, \text{МэВ}$	$-\frac{dE}{d\xi}, \text{МэВ}/(\text{г}\cdot\text{см}^{-2})$		Частицы	$E, \text{МэВ}$	$-\frac{dE}{d\xi}, \text{МэВ}/(\text{г}\cdot\text{см}^{-2})$		
		в воз-духе	в свин-це			в воз-духе	в свин-це	
Протоны	1	300	150	μ -Мезоны	0,112	300		
	10	47	27,5			1,12	47	
	100	7,6	5,0			11,2	7,6	
	1000	2,3	1,6			112	2,3	
	10000	2,3	1,6			1120	2,3	
Дейтроны	2	300		α -Частицы	4	1200		
	20	47				40	188	
	200	7,6				400	30	
	2000	2,3				4000	9,2	
	20000	2,3				40000	9,2	

Приведем также значения длин пробегов электронов в зависимости от энергий электронов (выраженных в единицах mc^2 ; напомним, что $1m_e c^2 = 0,5 \text{ МэВ}$):

Вещество	Длина пробега электрона, см, в зависимости от его энергии, ед. mc^2				
	0,1	1	10	100	1000
Воздух	4,4	160	10^8	$1,7 \cdot 10^6$	$6,3 \cdot 10^4$
Вода	$4,7 \cdot 10^{-3}$	0,19	2,6	19	78
Свинец	$9 \cdot 10^{-4}$	$3,7 \cdot 10^{-2}$	0,3	1,25	2,5

24.3. Тормозное излучение

При столкновении электрона с какой-либо частицей или системой частиц наряду с его рассеянием происходит излучение фотона. Такой процесс называют тормозным излучением. В частности, электрон излучает фотоны, двигаясь в кулоновском поле, создаваемом ядром (или ядром вместе с окружающими его атомными

электронами). Именно в этом состоит второй механизм потерь энергии при движении заряда в веществе, о котором говорилось в предыдущем параграфе.

Подчеркнем еще раз,

что электрон обязательно должен взаимодействовать с другими частицами, чтобы мог излучиться фотон; свободный электрон, двигаясь в вакууме, излучать не может.

Это непосредственно вытекает из законов сохранения импульса и энергии, которые должны выполняться в любом процессе. Действительно, пусть \mathbf{p}_1 , \mathbf{p}_2 — импульсы и $E_1 = c \sqrt{m^2c^2 + p_1^2}$, $E_2 = c \sqrt{m^2c^2 + p_2^2}$ — энергии электрона до и после процесса излучения, а $\hbar\mathbf{k}$ и $\varepsilon = \hbar ck$ — импульс и энергия испускаемого фотона. Тогда $E_1 = E_2 + \hbar ck$, $\mathbf{p}_1 = \mathbf{p}_2 + \hbar\mathbf{k}$, откуда для угла ϑ между векторами \mathbf{k} и \mathbf{p}_2 получается соотношение $\cos \vartheta = E_2/(cp_2)$, которое не может выполняться, так как всегда $E > cp$.

Перейдем к определению сечения тормозного излучения электрона:

$$d\sigma_{\Gamma} = dN_{\Gamma} / I,$$

где dN_{Γ} — число фотонов, излучаемых за 1 с в телесный угол $d\omega$ с частотами в интервале $(\omega, \omega + d\omega)$, и I — плотность потока электронов (число электронов, падающих за 1 с на площадку площадью 1 м^2 , перпендикулярную скорости электронов \mathbf{v}).

Рассмотрим сперва область малых энергий электрона ($v \ll c$). В этом случае излучение носит дипольный характер и интенсивность его определяется известной нам формулой

$$d\mathcal{J} = \frac{1}{4\pi c^3} \left[\ddot{\mathbf{d}} \cdot \mathbf{n} \right]^2 d\omega,$$

где \mathbf{n} — единичный вектор в направлении излучения и \mathbf{d} — дипольный момент для одного электрона $\mathbf{d} = e\mathbf{r}$, поэтому

$$d\mathcal{J} = \frac{e^2}{4\pi c^3} \left[\ddot{\mathbf{r}} \cdot \mathbf{n} \right]^2 d\omega.$$

Нас интересует спектральное распределение излучения, поэтому мы должны разложить $\ddot{\mathbf{d}}$ в интеграл Фурье:

$$\ddot{\mathbf{d}} = \frac{1}{2\pi} \int \ddot{\mathbf{d}}_{\omega} e^{-i\omega t} d\omega,$$

подставить это выражение в формулу для $d\mathcal{J}$ и проинтегрировать $d\mathcal{J}$ по времени. Мы найдем тогда полную энергию $d\mathcal{E}$, излученную в элемент телесного угла:

$$d\mathcal{E} = \int \mathcal{J} dt = \int \mathcal{E}_{\omega} d\omega,$$

где \mathcal{E}_{ω} — спектральная плотность излучения. Такой расчет мы уже делали в § 27.3 и показали, что

$$d\mathcal{E}_\omega = \frac{d\omega}{2(2\pi)^2 c^3} \left| \left[\ddot{\mathbf{d}}_\omega \mathbf{n} \right] \right|^2.$$

Полученная формула для $d\mathcal{E}_\omega$ справедлива в дипольном приближении, когда $v \ll c$ и когда энергия $\hbar\omega$ излученного фотона мала по сравнению с энергией электрона, т. е. в области малых частот. В этом случае легко найти $\ddot{\mathbf{d}}_\omega$; в самом деле, полагая $\omega \rightarrow 0$ в выражении для фурье-компоненты $\ddot{\mathbf{d}}_\omega = \int \ddot{\mathbf{d}}(t) \exp(i\omega t) dt$, получим

$$\ddot{\mathbf{d}}_\omega = \int \ddot{\mathbf{d}}(t) dt = \dot{\mathbf{d}}_2 - \dot{\mathbf{d}}_1,$$

где $\dot{\mathbf{d}}_1$ и $\dot{\mathbf{d}}_2$ — значения дипольного момента до и после излучения. Учитывая, что в нашем случае $\mathbf{d} = e\mathbf{r}$, получим окончательно

$$d\mathcal{E}_\omega = \frac{e^2}{2(2\pi)^2 c^3} \left[(\mathbf{v}_2 - \mathbf{v}_1) \mathbf{n} \right]^2 d\omega d\omega_0 \quad (\hbar\omega \ll E_{1,2}),$$

где \mathbf{v}_1 , \mathbf{v}_2 и E_1 , E_2 — значения скоростей и энергий электрона до и после излучения. Разделив это выражение на энергию фотона $\hbar\omega$, найдем вероятность $d\omega_\gamma$ излучения фотона в направлении $\mathbf{n} = \mathbf{k}/k$ в интервале телесных углов $d\omega_0$ и в интервале частот $(\omega, \omega + d\omega)$:

$$d\omega_\gamma = \frac{e^2}{2(2\pi)^2 c^3 \hbar\omega} \left[\frac{\mathbf{k}}{k} (\mathbf{v}_1 - \mathbf{v}_2) \right]^2 d\omega d\omega_0$$

($\hbar k$ — импульс фотона).

Теперь мы в состоянии определить сечение тормозного излучения в нерелятивистском случае и в области малых частот. В этом случае процесс излучения не оказывает влияния на движение электрона. Иными словами, электрон рассеивается так, как если бы никакого излучения не было (само рассеяние необходимо для излучения, так как с ним связано наличие ускорения электрона и, следовательно, тот факт, что $\mathbf{v}_2 \neq \mathbf{v}_1$), так что достаточно умножить выведенную в § 24.1 формулу Резерфорда для сечения упругого рассеяния на вероятность излучения фотона $d\omega_\gamma$, чтобы получить сечение тормозного излучения $d\sigma_r$. Замечая, что импульсы электрона \mathbf{p}_1 и \mathbf{p}_2 до и после рассеяния почти не отличаются по модулю и, следовательно, $|\mathbf{p}_1 - \mathbf{p}_2| = 2p_1 \sin(\theta/2)$, можно записать формулу Резерфорда в виде

$$d\sigma = \frac{(2mZe^2)^2}{(\mathbf{p}_1 - \mathbf{p}_2)^4} d\omega_2,$$

где $d\omega_2$ — элемент телесного угла вектора \mathbf{p}_2 . Таким образом, окончательно мы получим

$$d\sigma_r = \frac{Z^2 e^6 d\omega d\omega_0 d\omega_2}{2\pi^2 c (\hbar\omega)^3 (\mathbf{p}_1 - \mathbf{p}_2)^4} [\hbar k (\mathbf{p}_1 - \mathbf{p}_2)]^2.$$

Эта формула относится к нерелятивистской области. Чтобы найти сечение $d\sigma_r$ в релятивистском случае, когда $v_1 \sim c$, нужно использовать квантовую механику. Мы ограничимся тем, что приведем только некоторые результаты.

Спектр излучения, т. е. сечение $d\sigma_r$, проинтегрированное по углу рассеяния электрона и углу вылета фотона, в кулоновском поле ядра в ультрарелятивистской области ($E_{1,2} \gg mc^2$) определяется формулой

$$d\sigma_r = 4Z^2 r_0^2 \alpha \left[1 + \left(1 - \frac{\hbar\omega}{E_1} \right)^2 - \frac{2}{3} \left(1 - \frac{\hbar\omega}{E_1} \right) \right] \left[\ln \frac{2E_1}{mc^2} + \ln \left(\frac{E_1}{\hbar\omega} - 1 \right) - \frac{1}{2} \right] \frac{d\omega}{\omega},$$

где $r_0 = e^2/(m_e c^2)$ — классический радиус электрона и $\alpha = e^2/(\hbar c) \sim 1/137$ — постоянная, характеризующая интенсивность электромагнитного взаимодействия (ее называют *постоянной тонкой структуры*). Мы видим, что при фиксированном отношении $\hbar\omega/E_1$ сечение тормозного излучения растет примерно как $\ln[E_1/(mc^2)]$.

Тормозное излучение используется, в частности, для получения фотонов в электронных усилителях; в этом случае $E_1 = \text{const}$, так что спектр тормозного излучения оказывается, грубо говоря, $d\omega/\omega$ (разумеется, согласно закону сохранения энергии, обязательно $\hbar\omega < (E_1 - mc^2)$).

Определив $d\sigma_r$, легко найти тормозные потери энергии электрона на единичном пути:

$$\left(-\frac{dE}{dx} \right)_r = N \int_0^{E_1 - mc^2} \hbar\omega \frac{d\sigma_r}{d\omega} d\omega$$

(N — плотность ядер). Если записать эту величину в виде

$$\left(-\frac{dE}{dx} \right)_r = N E_1 \Phi_r,$$

то для Φ_r в нерелятивистском ($E_1/(mc^2) - 1 \ll 1$) и в ультрарелятивистском ($E_1 \gg mc^2$) случаях получаются следующие формулы:

$$\Phi_r = \frac{16}{3} Z^2 r_0^2 \alpha \left(\frac{E_1}{mc^2} - 1 \ll 1 \right),$$

$$\Phi_r = 4Z^2 r_0^2 \alpha \left(\ln \frac{2E_1}{mc^2} - \frac{1}{3} \right) \left(\frac{E_1}{mc^2} \gg 1 \right).$$

Мы видим, что в нерелятивистской области отношение излученной энергии к начальной не зависит от начальной энергии, а в ультрарелятивистском — логарифмически растет с энергией. Последний вывод, однако, правилен лишь в пренебрежении экранированием заряда ядра атомными электронами; при учете экранирования выражение в скобках должно быть заменено на $(\ln(183/Z^{1/3}) + 1/18)$ и Φ_r — тоже не будет зависеть от энергии E_1 .

Сравним теперь радиационные потери электрона с его ионизационными потерями, рассмотренными в предыдущем параграфе. Предполагая, что $E_1 \gg mc^2$, и замечая, что $n = ZN$, перепишем формулу Бете для ионизационных потерь в виде

$$\left(-\frac{dE}{dx}\right)_i = 8\pi N Z r_0^2 m_e c^2 \ln \frac{E_1}{\sqrt{\langle l \rangle mc^2}}.$$

Таким образом, в ультрарелятивистском случае

$$\frac{(-dE/dx)_r}{(-dE/dx)_i} = \frac{ZE_1 \alpha}{4\pi m_e c^2} \cong \frac{ZE_1}{1600 m_e c^2}.$$

Потери энергии электрона на ионизацию и на излучение сравниваются при энергии электрона $E_c \sim 800$ МэВ/З

(что составляет для свинца ~ 10 МэВ, для меди ~ 25 МэВ и для воздуха ~ 100 МэВ).

24.4. Влияние поляризации среды на ионизационные потери

Рассматривая (см. § 24.2), ионизационные потери энергии, мы видели, что они определяются полем, создаваемым движущейся заряженной частицей в месте нахождения атомного электрона. Мы решали задачу, считая, что есть только пролетающая частица и один атомный электрон. Между тем электрон находится в среде, которая поляризуется движущейся частицей; поэтому поле в месте нахождения электрона ослабляется и он получает меньшую энергию, чем если бы он был изолированным. Таким образом, поляризация среды должна приводить к уменьшению ионизационных потерь (э ф ф е к т Ф е р м и).

Чтобы дать количественную теорию этого эффекта, обратимся к уравнениям Максвелла для полей в среде:

$$\operatorname{rot} \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{H}}{\partial t}, \quad \operatorname{div} (\mathbf{E} + 4\pi \mathbf{P}) = 4\pi \rho_e,$$

$$\operatorname{rot} \mathbf{H} = \frac{1}{c} \frac{\partial}{\partial t} (\mathbf{E} + 4\pi \mathbf{P}) + \frac{4\pi}{c} \mathbf{j}, \quad \operatorname{div} \mathbf{H} = 0,$$

где \mathbf{P} — вектор поляризации среды и ρ_e , \mathbf{j} — плотности заряда и тока, связанные с движущейся частицей. Если частица движется равномерно со скоростью \mathbf{v} , то ее траектория имеет вид $\mathbf{r} = \mathbf{v}t$ и поэтому, используя трехмерную δ -функцию $\delta(\mathbf{r}) = \delta(x)\delta(y)\delta(z)$, величины ρ_e , \mathbf{j} можно представить в виде

$$\rho_e = q\delta(\mathbf{r} - \mathbf{v}t), \quad \mathbf{j} = q\mathbf{v}\delta(\mathbf{r} - \mathbf{v}t),$$

где q — заряд частицы. Что касается вектора поляризации среды, то в простейшем случае его можно определить (как это уже делалось в §§ 16.2, 16.3), исходя из представления об атомных электронах как некоторых осцилляторах. Если ω_0 — частота и γ — коэффициент затухания осциллятора, то поляризуемость среды определяется формулой

$$\alpha(\omega) = \frac{e^2 n / m}{-\omega^2 - 2i\gamma\omega + \omega_0^2}$$

(n — плотность электронов).

Нас интересует задача с цилиндрической симметрией — избранным направлением выберем направление движения частицы (вдоль которого мы напомним ось z). Поэтому все величины можно разложить в интеграл Фурье по частоте ω и z -проекции волнового вектора k_z :

$$f(\mathbf{r}, t) = \int f_{\omega, k_z}(\rho) \exp\{-i\omega t + ik_z z\} d\omega dk_z,$$

где $f_{\omega, k_z}(\rho)$ — некоторая функция $\rho \equiv \sqrt{x^2 + y^2}$. Фурье-компоненту δ -функции легко найти, вспоминая, что $\int dt \exp(i\xi t) = 2\pi\delta(\xi)$; обозначив $\Delta_{\omega, k_z}(\rho)$ эту фурье-компоненту, имеем

$$\begin{aligned} \Delta_{\omega, k_z}(\rho) &= \frac{\delta(x)\delta(y)}{(2\pi)^2} \int \exp\{i\omega t - ik_z z\} \delta(z - vt) dz dt = \\ &= \frac{\delta(x)\delta(y)}{(2\pi)^2} \int \exp\{it(\omega - vk_z)\} dt = \frac{\delta(x)\delta(y)}{2\pi} \delta(\omega - vk_z) \end{aligned}$$

(входящее сюда произведение $\delta(x)\delta(y)$ зависит от x и y не независимо, а в виде комбинации $\rho = \sqrt{x^2 + y^2}$).

Выразим теперь напряженности полей через векторный потенциал \mathbf{A} :

$$\mathbf{H} = \text{rot } \mathbf{A}, \quad \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{A}}{\partial t}.$$

Для фурье-компоненты векторного потенциала мы получим уравнение

$$\left[k_z^2 - \Delta_\rho - \frac{\omega^2}{c^2} \varepsilon(\omega) \right] \mathbf{A}_{\omega, k_z}(\rho) = \frac{4\pi q}{c} \mathbf{v} \frac{\delta(x)\delta(y)}{2\pi} \delta(\omega - vk_z),$$

где Δ_ρ — лапласиан по переменной ρ , т. е.

$$\Delta_\rho f = \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial f}{\partial \rho} \right).$$

Решение этого уравнения можно, очевидно, искать в форме

$$\mathbf{A}_{\omega, k_z}(\rho) = \frac{4\pi q}{c} \mathbf{v} \delta(\omega - vk_z) u(\rho),$$

где $u(\rho)$ определяется уравнением

$$\left[\omega^2 \left(\frac{1}{v^2} - \frac{\varepsilon(\omega)}{c^2} \right) - \Delta_\rho \right] u(\rho) = \frac{\delta(x)\delta(y)}{2\pi}.$$

Мы видим, что функция $u(\rho)$ удовлетворяет однородному волновому уравнению (двумерному) везде, кроме точки $x = y = 0$, т. е.

$$\Delta_\rho u(\rho) + s^2 u(\rho) = 0, \quad s^2 \equiv \omega^2 \left(-\frac{1}{v^2} + \frac{\varepsilon(\omega)}{c^2} \right) \quad (\rho \neq 0).$$

Умножим уравнение для u с δ -функциями в правой части на ρ и проинтегрируем по $d\rho$. Учитывая, что для цилиндрически симметричной величины интегрирование по $dx dy$ эквивалентно интегрированию с $2\pi\rho d\rho$, получим

$$\left(\rho \frac{\partial u}{\partial \rho} \right)_{\rho \rightarrow 0} = 1.$$

Отсюда можно сделать следующий вывод: чтобы найти векторный потенциал, нужно решить однородное уравнение для $u(\rho)$ с выписанным граничным условием на производную от этой функции в точке $\rho = 0$.

Такому уравнению удовлетворяет цилиндрическая функция $K_0(is\rho)$. Производная от K_0 выражается, как известно, через цилиндрическую функцию K_1 , так что через эти две функции можно без труда выразить все компоненты электрического и магнитного полей.

При известных электрических и магнитных полях мы можем найти плотность потока энергии S_b через поверхность цилиндра некоторого радиуса b , окружающего траекторию частицы. Разделив S_b на v , мы получим энергию W_b , теряемую частицей на единичном пути и соответствующую прицельным параметрам, превышающим выбранное значение b :

$$W_b = \frac{c}{4\pi v} \int [\mathbf{E}\mathbf{H}] ds$$

(интегрирование производится по поверхности цилиндра радиуса b). Величина W_b выражается через цилиндрические функции:

$$W_b = \frac{e^2 b}{\pi} \int_{-\infty}^{\infty} \left(\frac{-1}{\varepsilon(\omega) v^2} + \frac{1}{c^2} \right) \omega s^* K_1(-is^*b) K_0(isb) d\omega.$$

Этот интеграл легко вычислить при достаточно малых b . В случае прозрачной среды ($\gamma \rightarrow 0$) ответ имеет вид

$$W_b = \frac{2\pi e^4 n}{mv^3} \left\{ \ln \frac{mv^2}{3,17\pi e^2 nb^2} + \ln \frac{\varepsilon_0 - 1}{\varepsilon_0 \left(1 - \frac{v^2}{c^2}\right)} - \frac{v^2}{c^2} \right\} \quad \left(v < \frac{c}{\sqrt{\varepsilon_0}} \right),$$

$$W_b = \frac{2\pi e^4 n}{mv^3} \left\{ \ln \frac{mv^2}{3,17\pi e^2 nb^2} - \frac{1 - \frac{v^2}{c^2}}{\varepsilon_0 - 1} \right\} \quad \left(v > \frac{c}{\sqrt{\varepsilon_0}} \right),$$

где $\varepsilon_0 = 1 + 4\pi e^2 n / (m\omega_0^2)$ — диэлектрическая постоянная в области низких частот.

Представляет интерес сравнить полученные формулы с выражением для ионизационных потерь, не учитывающим поляризацию среды. Последние мы уже рассматривали в § 24.2. Мы не можем, однако, непосредственно применить результаты § 24.2, а должны заново провести расчет — на этот раз не для бесконечной среды, а для цилиндра малого радиуса b . Приведем лишь окончательные выводы. Оказывается, что

поляризация среды приводит к уменьшению ионизационных потерь

(этот результат виден уже при непосредственном сравнении формул § 24.2 и настоящего параграфа: величина W_b при $E \gg mc^2$ стремится к постоянному пределу, в то время как найденные в § 24.2 потери растут как $\ln [E/(mc^2)]$). Точный расчет показывает, что учет поляризации среды приводит к уменьшению потерь для $E = 10^3 mc^2$ в два раза для воды и на $1/4$ для воздуха (при выборе $b = 10^{-8}$ см).

24.5. Излучение Вавилова — Черенкова

Полученная в предыдущем параграфе величина W_b определяет потери энергии частицы, если прицельный параметр превосходит заданное значение b . Мы привели общее выражение для W_b и проана-

лизировали случай малых b , порядка межатомных расстояний (физически отвечающий ионизационным потерям энергии при парных столкновениях с учетом поляризации среды).

Возникает вопрос: как ведет себя W_b при больших b ? Вспомним, что цилиндрические функции быстро осциллируют или экспоненциально убывают при увеличении аргумента, так что содержащие их интегралы убывают. Поэтому естественным кажется, что потери энергии убывают при увеличении прицельного параметра в соответствии с тем, что обуславливающие их электромагнитные силы на больших расстояниях почти полностью заэкранированы.

Ситуация, однако, резко меняется, если скорость частицы больше фазовой скорости электромагнитных волн в среде; $v > v_\Phi = c/\sqrt{\epsilon(\omega)}$. В этом случае в область интегрирования в формуле для W_b попадает точка $s = 0$, вблизи которой аргумент цилиндрической функции остается конечным при сколь угодно больших b (среда предполагается прозрачной ($\text{Im}\epsilon(\omega) = 0$), так что величина s^2 вещественна). При этом W_b при $b \rightarrow \infty$ оказывается отличным от нуля и определяется формулами

$$W_\infty = \frac{2\pi e^4 n}{mv^2} \left\{ -\frac{v^2}{c^2} - \ln\left(1 - \frac{v^2}{c^2}\right) \right\} \quad \left(v < \frac{c}{\sqrt{\epsilon_0}} \right),$$

$$W_\infty = \frac{2\pi e^4 n}{mv^2} \left\{ -\frac{1 - \frac{v^2}{c^2}}{\epsilon_0 - 1} + \ln \frac{\epsilon_0}{\epsilon_0 - 1} \right\} \quad \left(v > \frac{c}{\sqrt{\epsilon_0}} \right).$$

Каков же смысл энергии, уходящей от частицы на очень большие расстояния? Очевидно, что это энергия излучаемых частицей электромагнитных волн (а не энергия возбуждения или ионизации отдельных атомов). Такое излучение называется *излучением Вавилова — Черенкова* (или для краткости — *черенковским излучением*). (Оно было открыто С. И. Вавиловым и П. А. Черенковым в 1934 г.; теория эффекта была развита И. Е. Таммом и И. М. Франком.)

Возможность излучения Вавилова — Черенкова особенно наглядна при квантовом подходе. Мы уже говорили, что в вакууме равномерно движущаяся частица не может излучить фотона; при этом не могли бы одновременно выполняться законы сохранения энергии и импульса:

$$E_1 = E_2 + \hbar\omega \quad \text{и} \quad \mathbf{p}_1 = \mathbf{p}_2 + \hbar\mathbf{k},$$

где $E_{1,2} = c\sqrt{m^2c^2 + p_{1,2}^2}$ энергии электрона до и после излучения, $\hbar\omega = \hbar ck$ и $\hbar\mathbf{k}$ — энергия и импульс фотона. Иначе обстоит дело в среде с показателем преломления $N \equiv \sqrt{\epsilon(\omega)} > 1$: теперь энергия фотона с импульсом $\hbar\mathbf{k}$ есть $\hbar\omega = \hbar ck/N$, т. е. меньше, чем в случае вакуума, и законы сохранения энергии и импульса смогут удовлетвориться. В самом деле, используя законы сохранения энергии и импульса, легко получить следующее выражение для косинуса угла ϑ между \mathbf{k} и \mathbf{p}_2 :

$$\cos \vartheta = \frac{v_{\Phi}}{v_2} \left[\frac{E_1}{E_2} - \frac{E_1 - E_2}{2E_2} \left(\frac{c^2}{v_{\Phi}^2} - 1 \right) \right],$$

где $v_2 = p_2 c^2 / E_2$ и v_{Φ} — фазовая скорость волны. Таким образом, при $v > v_{\Phi}$ фотон может излучаться, причем под строго определенным углом к направлению движения частицы.

Если фотон уносит лишь небольшую долю импульса электрона: $\hbar k \ll p_1$ (что имеет место, в частности, при близком к единице показателе преломления), то выражение в квадратных скобках равно единице и

$$\cos \vartheta = v_{\Phi} / v \quad (\hbar k \ll p).$$

Не приводя выражений для векторного потенциала и полей при черенковском излучении, заметим лишь, что согласно результатам предыдущего параграфа их зависимость от ρ при $\rho \rightarrow \infty$ имеет вид $f \sim \rho^{-1/2} \exp(i s \rho)$, где по-прежнему

$$s^2 = \omega^2 \left(\frac{\epsilon(\omega)}{c^2} - \frac{1}{v^2} \right).$$

При мнимых значениях s эта величина экспоненциально убывает (что соответствует экранированию полей). Если же $v > v_{\Phi}$, то s делается вещественным и компоненты поля осциллируют по характерному для расходящейся цилиндрической волны закону $f \sim \rho^{-1/2} \exp(i s \rho)$. Энергия, излучаемая частицей на длине пути l , как можно показать, определяется формулой

$$\mathcal{E} = \frac{e^2 l}{c^2} \int_{(\beta N > 1)} \left(1 - \frac{1}{\beta^2 N} \right) \omega d\omega,$$

где $\beta = v/c$, N — показатель преломления и интегрирование производится по области частот, в которой $\beta N > 1$ (Ф о р м у л а Т а м м а — Ф р а н к а).

Черенковское излучение можно пояснить еще как результат суперпозиции волн, непрерывно излучаемых частицей. Если $\omega \neq kv$, то энергия, теряемая частицей на одном полупериоде, в точности равна энергии, приобретаемой ею на другом полупериоде волны, так что энергия от частицы волне в среднем не передается. Если же $\omega = kv$, т. е. $\cos \vartheta = v_{\Phi} / v$, то частица движется вместе с волной, находясь все время в точке с одинаковой фазой; в этом случае энергия частицы непрерывно переходит к волне.

Обратим внимание на то, что

черенковские потери энергии частицы не зависят от ее массы в отличие от тормозных потерь, интенсивность которых падает с увеличением массы частицы.

Подчеркнем, что рассмотренные в предыдущем параграфе потери энергии частицы с учетом поляризации среды автоматически содержат как ионизационные, так и черенковские потери (физический смысл последних имеет при этом величина W_{∞}).

Важным техническим приложением излучения Вавилова — Черенкова являются так называемые *черенковские счетчики* заряженных частиц.

24.6. Рассеяние фотона электроном

Перейдем к рассмотрению прохождения фотонов через вещество.

При этом могут происходить три процесса: поглощение фотонов с возбуждением или ионизацией атомов, рассеяние фотонов атомными электронами и превращение фотонов в электрон-позитронные пары.

Начнем с рассеяния фотона электроном (э ф ф е к т К о м п т о н а). Будем сперва предполагать, что энергия фотона значительно больше энергии связи электрона в атоме. В этом случае электроны можно рассматривать как свободные, так что должны выполняться законы сохранения энергии и импульса:

$$E_1 + \hbar\omega_1 = E_2 + \hbar\omega_2, \quad \mathbf{p}_1 + \hbar\mathbf{k}_1 = \mathbf{p}_2 + \hbar\mathbf{k}_2,$$

где индексы 1 и 2 служат для обозначения величин до и после рассеяния. Учитывая, что $E = \sqrt{m_e c^2 + p^2}$ и $\omega = ck$, отсюда легко выразить угол рассеяния фотона ϑ (угол между \mathbf{k}_1 и \mathbf{k}_2) через углы ϑ_1 , ϑ_2 между скоростью \mathbf{v}_1 электрона до рассеяния и \mathbf{k}_1 , \mathbf{k}_2 :

$$1 - \cos \vartheta = \frac{E_1}{\hbar\omega_2} \left(1 - \frac{v_1}{c} \cos \vartheta_1\right) - \frac{E_1}{\hbar\omega_1} \left(1 - \frac{v_1}{c} \cos \vartheta_2\right).$$

Если электрон вначале покоился ($v_1 = 0$, $E_1 = m_e c^2$; фактически это означает лишь, что $v_1 \ll v_2$, т. е. что скорость рассеянных электронов велика по сравнению со скоростью атомных электронов), то из этого соотношения вытекает формула

$$\omega_2 = \frac{\omega_1}{1 + \frac{\hbar\omega_1}{m_e c^2} (1 - \cos \vartheta)}.$$

Ее называют формулой Комптона.

Мы видим, что $\omega_2 \leq \omega_1$ (равенство имеет место при $\vartheta = 0$); это естественно: если электрон первоначально покоился, то фотон может только передать ему (но не забрать) энергию. В предельном случае малых энергий фотона ($\hbar\omega_1 \ll mc^2$) частота фотона при рассеянии почти не меняется; для очень жестких фотонов ($\hbar\omega_1 \gg mc^2$) частота $\hbar\omega_2 = m_e c^2 / (1 - \cos \vartheta)$ рассеянного фотона перестает зависеть от частоты падающего фотона (последняя формула становится несправедливой при очень малых углах рассеяния).

Из соотношения между ω_2 и ω_1 легко установить связь между соответствующими длинами волн:

$$\lambda_2 - \lambda_1 = \lambda_C (1 - \cos \vartheta),$$

где $\lambda_C \equiv 2\pi\hbar/(m_e c) \sim 2\pi \cdot 3,86 \cdot 10^{-11}$ см. Эту величину называют *комptonовской длиной волны электрона*.

Частота рассеянного фотона, как мы видели, однозначно определяется углом рассеяния (при заданной частоте падающего фотона), сам же угол рассеяния может быть любым. Иными словами, комptonов-

ское рассеяние (как и любой квантово-механический процесс) описывается статистическими законами и характеризуется дифференциальным сечением рассеяния

$$d\sigma = dN/I,$$

где dN — число фотонов, рассеянных под углами в интервале $(\vartheta, \vartheta + d\vartheta)$, и I — плотность потока падающих фотонов (т. е. число их, попадающих за 1 с на площадку площадью 1 см², перпендикулярную вектору \mathbf{k}_1).

Дифференциальное сечение рассеяния $d\sigma$ можно просто определить для очень мягких фотонов ($\lambda \gg \lambda_c$); при этом можно вообще не пользоваться понятием фотона и провести расчет, не выходя за рамки классической электродинамики. Действительно, пусть на электрон падает плоская волна частоты ω , электрическое поле которой имеет вид $\mathbf{E} = \mathbf{E}_0 \cos(\mathbf{k}\mathbf{r} - \omega t)$. Под влиянием поля электрон начинает двигаться, а так как $\lambda \gg \lambda_c$, т. е. $\hbar\omega \ll mc^2$, то энергия, приобретаемая электроном, невелика, а следовательно, мала и скорость электрона. Поэтому ускорение электрона определяется только электрической силой: $\ddot{\mathbf{r}} = (e\mathbf{E}_0/m_e)\cos\omega t$ (электрон покоится в начале координат). Движущийся с ускорением электрон начнет излучать электромагнитные волны; их и следует рассматривать как рассеянное излучение. Интенсивность этого излучения можно определить по общей формуле для дипольного излучения:

$$d\mathcal{J} = \frac{1}{4\pi c^3} (\ddot{\mathbf{d}})^2 \sin^2\theta d\theta,$$

где $\mathbf{d} = e\mathbf{r}$, θ — угол между \mathbf{k}' и $\ddot{\mathbf{r}}$, а $d\theta$ — элемент телесного угла, в котором лежит волновой вектор \mathbf{k}' рассеянного излучения. Разделив $d\mathcal{J}$ на плотность потока энергии в падающей волне $S = cE_0^2/(4\pi)$, получим дифференциальное сечение рассеяния

$$d\sigma = r_0^2 \sin^2\theta d\theta, \quad \hbar\omega \ll m_e c^2,$$

где $r_0 = e^2/(m_e c^2)$ — классический радиус электрона.

Проинтегрировав полученное выражение по углам, найдем полное сечение рассеяния мягкого фотона:

$$\sigma = \frac{8\pi}{3} r_0^2.$$

Эту формулу называют формулой Томсона, а сечение $\sigma = 0,66 \times 10^{-24}$ см² — *томсоновским сечением*.

До сих пор мы считали электрон свободным; но если величина $\hbar\omega$ становится сравнимой с энергией связи электрона в атоме, то электрон нельзя уже рассматривать как свободную частицу. Чтобы понять характер рассеяния в этом случае, будем трактовать электрон как классический осциллятор. Под влиянием электрического поля такой осциллятор совершает вынужденные колебания, описываемые уравнением

$$\ddot{\mathbf{r}} + \omega_0^2 \mathbf{r} = \frac{e}{m_e} \mathbf{E}_0 \cos \omega t$$

(ω_0 — собственная частота осциллятора). Отсюда легко определить $\mathbf{r}(t)$ и, следовательно, входящую в формулу для дипольного излучения величину $\ddot{\mathbf{d}} = e\ddot{\mathbf{r}}$. Представляя далее интенсивность дипольного излучения в виде $d\mathcal{J} = S d\sigma$ (S — по-прежнему плотность потока энергии падающей волны), найдем дифференциальное сечение рассеяния:

$$d\sigma = r_0^2 \frac{\omega^4}{(\omega_0^2 - \omega^2)^2} \sin^2 \theta d\omega.$$

В отличие от томсоновского сечения это сечение зависит от частоты фотона; в области очень малых частот ($\omega \ll \omega_0$) оно пропорционально ω^4 . Эта зависимость, в частности, объясняет голубой цвет неба: интенсивно рассеивается и потому преимущественно попадает в глаза наблюдателя коротковолновая часть солнечного света.

При $\omega = \omega_0$, т. е. в случае резонанса, написанная формула приводит к бесконечному сечению; в этом случае она непригодна, так как не учитывает эффекта затухания. Учет затухания, как мы неоднократно убеждались, приводит к замене резонансного знаменателя $(\omega_0^2 - \omega^2)^2$ на $[(\omega_0^2 - \omega^2)^2 + \gamma^2 \omega^2]$, где γ — коэффициент затухания, так что в резонансе

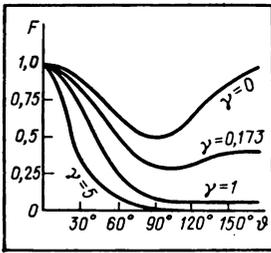
$$d\sigma = r_0 (\omega_0 / \gamma)^2 \sin^2 \theta d\omega.$$

Подчеркнем, что атомный электрон очень приближенно можно трактовать как классический осциллятор, так как атом подчиняется не классической, а квантовой механике. Поэтому сечение рассеяния электромагнитной волны атомом определяется другой, более сложной формулой, чем формула для сечения рассеяния классическим осциллятором. При этом существенно, что в процессе рассеяния атом может перейти из одного энергетического состояния в другое. Поэтому частота рассеянной волны может отличаться от частоты падающей волны. Если частоты этих волн равны ω_1 и ω_2 , то общая формула для дифференциального сечения рассеяния в интервале телесного угла $d\omega_2$ имеет вид

$$d\sigma = \omega_1 \omega_2^3 d\omega_2 \left| \sum_s \frac{(\mathbf{e}_2 \mathbf{d})_{2s} (\mathbf{e}_1 \mathbf{d})_{s1}}{E_1 - E_s + \hbar \omega_1} + \frac{(\mathbf{e}_1 \mathbf{d})_{2s} (\mathbf{e}_2 \mathbf{d})_{s1}}{E_1 - E_s - \hbar \omega_2} \right|^2,$$

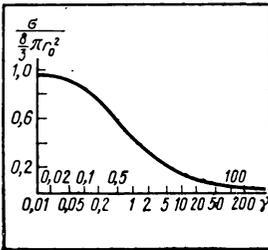
где \mathbf{e}_1 и \mathbf{e}_2 — векторы поляризации обеих волн, $(\mathbf{e} \mathbf{d})_{ij}$ — матричные элементы скалярного произведения векторов \mathbf{d} и \mathbf{e} , E_1 и E_2 — энергии атома до и после рассеяния, связанные законом сохранения энергии $E_1 + \hbar \omega_1 = E_2 + \hbar \omega_2$, и суммирование производится по всем «промежуточным» состояниям s атома с энергией E_s . При $\omega_1 = \omega_2$ (этот случай называют *когерентным рассеянием*) сечение рассеяния пропорционально ω^4 , так же как и в классической теории.

Возвратимся теперь к эффекту Комптона при произвольных энергиях фотона. Формула для дипольного излучения при этом уже не-



24.3.

Функция $P = \frac{d\sigma}{d\omega} \frac{1}{r_0^2} \frac{\omega_2}{\omega_1}$, характеризующая угловое распределение рассеянных фотонов при эффекте Комптона



24.4.

Функция $\sigma / (8/3 \pi r_0^2)$, описывающая зависимость полного сечения эффекта Комптона от энергии падающего фотона

ли, частота рассеянного фотона является однозначной функцией ϑ ; подставив в эту формулу найденную в начале параграфа зависимость $\omega_2(\vartheta)$, мы получим угловое распределение рассеянных фотонов. Величина $d\sigma/d\omega$ как функция угла ϑ и параметра $\gamma = \hbar\omega_1/(m_e c^2)$ представлена на рис. 24.3.

Если проинтегрировать дифференциальное сечение по углам, то мы получим полное сечение σ эффекта Комптона. Зависимость σ от γ изображена на рис. 24.4. Мы приведем формулы для σ только в ультрарелятивистском ($\hbar\omega_1 \gg m_e c^2$) и в нерелятивистском ($\hbar\omega_1 \ll m_e c^2$) случаях:

$$\sigma = \pi r_0^2 \frac{m_e c^2}{\hbar\omega_1} \left(\ln \frac{2\hbar\omega_1}{m_e c^2} + \frac{1}{2} \right) \quad (\hbar\omega_1 \gg m_e c^2),$$

$$\sigma = \frac{8}{3} \pi r_0^2 \left(1 - 2\gamma \frac{\hbar\omega_1}{m_e c^2} \right) \quad (\hbar\omega_1 \ll m_e c^2)$$

применима; более того, исследование эффекта Комптона при произвольном $\hbar\omega$ вообще невозможно в рамках классической физики, т. е. с использованием только механики Ньютона и электродинамики Максвелла. Дело в том, что если энергия фотона сравнима или превосходит энергию покоя электрона, то последний при столкновении с фотоном может приобретать значительную энергию, а так как распределение рассеянных фотонов по направлениям (а следовательно, и по энергиям) подчиняется статистическим закономерностям, то таким же закономерностям будет подчиняться и электрон после рассеяния. Статистические же закономерности в поведении электрона лежат вне рамок классической физики и составляют предмет квантовой механики. Иными словами, определение сечения рассеяния фотона электроном при $\hbar\omega \geq m_e c^2$ может быть произведено только в квантовой механике, причем с учетом требований релятивистской теории. Поэтому мы приведем только некоторые результаты теории эффекта Комптона.

Дифференциальное сечение рассеяния (в случае неполяризованных частиц) определяется формулой

$$d\sigma = \frac{1}{2} r_0^2 \left(\frac{\omega_2}{\omega_1} \right)^2 \left(\frac{\omega_2}{\omega_1} + \frac{\omega_1}{\omega_2} - \sin^2 \vartheta \right) d\omega,$$

где ϑ — по-прежнему угол рассеяния [фотона (угол между \mathbf{k}_2 и \mathbf{k}_1) и $d\omega$ — элемент телесного угла вектора \mathbf{k}_2 (электрон до рассеяния предполагается покоящимся)]. Как мы видели,

(последняя формула представляет собой, очевидно, формулу Томсона и первую поправку к ней по параметру γ).

24.7. Фотоэффект и образование электрон-позитронных пар

Если энергия фотона $\hbar\omega$ превосходит энергию связи электрона в атоме, т. е. энергию ионизации атома I , то фотон может быть поглощен атомом, а атомный электрон покинет атом, обладая кинетической энергией $T = \hbar\omega - I$. Это явление называют **фотоэффектом** (оно было открыто Герцем в 1887 г. и объяснено Эйнштейном в 1905 г. на основе квантовых представлений о природе света).

Электроны в атоме расположены в различных оболочках, каждой из которых отвечает своя энергия связи. Самую внутреннюю оболочку называют *K*-оболочкой; на ней находятся два электрона с энергией связи, приближенно равной RZ^2 , где $R = 13,5$ эВ — *постоянная Ридберга* и Z — атомный номер. Затем идут *L*- и *M*-оболочки с меньшими энергиями связи, равными примерно $\frac{1}{4} R(Z-5)^2$ и $\frac{1}{9} R(Z-13)^2$. Наименьшей энергией связи обладают электроны на оптических орбитах, поэтому для их отрыва нужна наименьшая частота фотона.

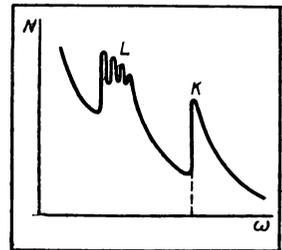
С ростом частоты в поглощении фотонов начинают участвовать электроны из все более глубоких оболочек. Когда энергия фотона достигает значения энергии связи электрона в какой-либо оболочке, происходит резкое увеличение числа поглощаемых фотонов. Схематическая зависимость числа N поглощенных фотонов от частоты показана на рис. 24.5. Точку, в которой поглощение испытывает скачок, обусловленный выбиванием *K*-электронов, называют *краем K-полосы* поглощения. Соответствующая частота равна $\omega_K = I_K/\hbar$, где I_K — энергия связи *K*-электрона.

Фотоэффект (подобно любому квантово-механическому эффекту) характеризуется *эффективным сечением*, определяемым как отношение числа выбитых за 1 с электронов к плотности потока падающих фотонов. Для *K*-электронов сечение фотоэффекта определяется формулой

$$\sigma_K = \frac{128 \pi e^2 \omega^3}{3 m_e \omega^4} \exp \left\{ -4\xi \operatorname{arc} \operatorname{ctg} \xi \right\} \times \\ \times (1 - \exp(-2\pi\xi))^{-1},$$

где $\xi = \omega_K^{1/2} (\omega - \omega_K)^{-1/2}$ и принято $\hbar = c = 1$.

Приведенное выражение относится к отдельному *K*-электрону (всего в атоме их два). Если просуммировать аналогичные (хотя и значительно более сложные) выражения по всем электронам атома, то мы получим отнесенное к одному атому сечение фотоэффекта σ_ϕ .



24.5. Зависимость числа поглощенных фотонов от частоты

По известному σ_{Φ} легко выяснить, как происходит поглощение фотонов в веществе. Если N — плотность атомов, то длина пробега фотона по отношению к фотоэффекту составит

$$l_{\Phi} = \frac{1}{N\sigma_{\Phi}}$$

и, следовательно, относительное уменьшение числа фотонов на пути dx есть dx/l_{Φ} . Пусть теперь $n \equiv n(x)$ — число фотонов в поглотителе на расстоянии x от его границы. Тогда $dn = -n dx/l_{\Phi}$, откуда

$$n(x) = n_0 e^{-x/l_{\Phi}} = n_0 e^{-N\sigma_{\Phi}x}.$$

где n_0 — число фотонов на входе в поглотитель.

Сечение фотоэффекта, испытывая скачки, убывает с ростом частоты. При $\omega > \omega_K$ оно уже мало и более существенным становится эффект Комптона. Следует, однако, иметь в виду, что при эффекте Комптона происходит, собственно, не поглощение фотонов, а их рассеяние, так что этот эффект вносит вклад в потери энергии (и направленного импульса) исходного фотонного пучка, но не в суммарную плотность фотонов.

Если энергия фотона превосходит $2m_e c^2 \sim 1$ МэВ, то в игру вступает новый процесс — образование электрон-позитронных пар. Позитрон e^+ отличается от электрона e^- только знаком заряда, так что пара e^+e^- электрически нейтральна, так же как и фотон. Поэтому с точки зрения закона сохранения электрического заряда фотон может «родить» электрон-позитронную пару, но так как энергия покоя пары равна $2m_e c^2$, то энергия фотона должна при этом превосходить $2m_e c^2$. Однако

один фотон в вакууме не может образовать пары, так как при этом не могли бы одновременно выполняться законы сохранения энергии и импульса.

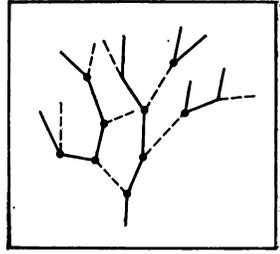
Для образования пары одним фотоном необходимо присутствие посторонней частицы, например ядра. Полное сечение образования пары в поле ядра (с учетом экранирования)

$$\sigma_{\pi} = \frac{2Z^2 e^6}{9m_e^2 c^5 \hbar} \left\{ 14 \ln \left(183 Z^{-1/3} \right) - \frac{1}{3} \right\}$$

(Z — атомный номер). Длина пробега фотона в веществе по отношению к этому процессу составит $l_n = 1/(N\sigma_{\pi})$, где N — плотность ядер, так что изменение интенсивности фотонного пучка происходит по закону $n = n_0 e^{-x/l_n}$. Заметим, что длина пробега быстрого электрона по отношению к процессу тормозного излучения имеет тот же порядок величины, что и l_n .

С процессами тормозного излучения и образования пар связано явление электромагнитных ливней, заключающееся в следующем. Пусть быстрый электрон с энергией порядка гигаэлектрон-вольт влетает в слой вещества. Тормозясь в кулоновском поле ядер, электрон излучает фотоны с энергией того же порядка, фотоны рожают пары, после чего процесс начинается снова: каждая из частиц

пары, тормозясь, излучает фотоны, которые, в свою очередь, рожают пары, и т. д. (рис. 24.6).



24.6. Электромагнитный ливень

Сплошные линии соответствуют электронам и позитронам, пунктирные — фотонам

Таким образом, из первоначально одной быстрой частицы — электрона возникает много вторичных частиц: электронов, позитронов и фотонов с более низкими энергиями. Об этих частицах говорят, что они образуют электромагнитный ливень. Ливень прекращается, когда энергия частиц опускается до таких значений, при которых в торможении фотонов начинает доминировать эффект Комптона, а в торможении электронов — ионизационные потери.

Ясно, что первичной частицей ливня не обязательно должен быть электрон; с тем же успехом ливень может породиться и энергичным фотоном.

Легко оценить число частиц в ливне на глубине x (считая от места влета первичной частицы). Так как длина пробега фотона по отношению к процессу образования пар e^+e^- и длина пробега электрона (и позитрона) по отношению к процессу тормозного излучения по порядку величины одинаковы, то можно ввести некоторую среднюю длину пробега частицы ливня $\langle l \rangle$:

$$\langle l \rangle = \text{const}/(N \langle \sigma \rangle),$$

где $\langle \sigma \rangle = Z^2 r_0^2 e^2 / (\hbar c)$ и const обозначает числовую постоянную порядка единицы (r_0 — классический радиус электрона). Число частиц ливня на глубине x

$$n(x) \sim 2^{x/\langle l \rangle}.$$

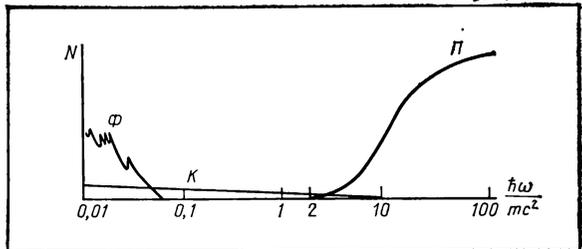
Если E_0 — энергия начальной частицы, то на глубине x , согласно закону сохранения энергии, средняя энергия частиц $E_x \sim E_0 2^{-x/\langle l \rangle}$. Ливень прекратится, когда эта величина достигнет критического значения E_c , при котором ионизационные потери энергии электрона становятся больше потерь на излучение ($E(x) \sim E_c$). При соответствующем значении x , которое мы обозначим l_m , число частиц ливня будет максимальным. Очевидно,

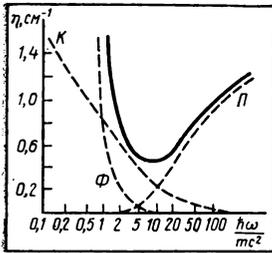
$$l_m \sim \langle l \rangle \left(\ln \frac{E_0}{E_c} \right) / \ln 2.$$

24.7.

Схематическая зависимость числа N поглощенных фотонов от их первоначальной энергии:

Π — образование пар, K — эффект Комптона, Φ — фотоэффект





24.8.

Зависимость коэффициента поглощения η фотонов в свинце от их энергии

Пунктирные линии — вклады отдельных эффектов, сплошная линия — полное поглощение (обозначения те же, что на рис. 24.7)

бужденное состояние с меньшей энергией, чем у исходного, и при этом будет излучен фотон с энергией

$$\hbar\omega_{i \rightarrow f} = E_i - E_f,$$

где i, f обозначают начальное и конечное состояния и E_i, E_f — уровни энергии атома. Напротив, если на атом в состоянии f падает фотон с энергией $\hbar\omega_{i \rightarrow f}$, то атом может перейти в состояние i с энергией $E_i = E_f + \hbar\omega_{i \rightarrow f}$, фотон при этом будет поглощен атомом.

Эти энергетические соотношения выражают закон сохранения энергии при испускании и поглощении фотона.

Они сыграли важнейшую роль при становлении квантовой физики и входят в число известных постулатов Бора (1913), с помощью которых Бору удалось впервые объяснить стабильность атома и дискретность атомных спектров.

Оба процесса — и испускание, и поглощение фотона (подобно фотоэффекту и различным процессам рассеяния) — подчиняются статистическим закономерностям, т. е. характеризуются определенными вероятностями. Для процессов с участием фотонов эти вероятности обладают замечательным свойством — они зависят от числа фотонов n_ω , присутствовавших до рассматриваемого процесса, а именно: вероятность ω_- поглощения фотона частоты ω за 1 с пропорциональна числу таких фотонов n_ω , а вероятность ω_+ испускания — числу таких фотонов плюс единица, т. е. $(n_\omega + 1)$, причем коэффициенты пропорциональности одинаковы для обоих процессов. Иными словами,

$$\frac{\omega_-}{\omega_+} = \frac{n_\omega}{n_\omega + 1}.$$

Перепишем это равенство в виде

После того как число частиц в ливне достигнет максимума, ливень постепенно расосется.

Вернемся в заключение кратко к основным процессам поглощения фотонов. Общий характер поглощения в зависимости от энергии фотона иллюстрируется рис. 24.7; на рис. 24.8 дана такая зависимость для свинца. Обратим внимание на резкий минимум поглощения — *окно прозрачности* для жестких фотонов.

24.8. Излучение и поглощение фотонов атомами

Если атом находится в возбужденном состоянии, то он может перейти в основное состояние или в другое воз-

$$\omega_+ = (n_\omega + 1)\omega_0, \quad \omega_- = n_\omega \omega_0,$$

где ω_0 не зависит от n_ω , мы видим, что вероятность испускания фотона состоит из двух слагаемых: ω_0 и $n_\omega \omega_0$. Первое слагаемое представляет собой вероятность испускания в том случае, когда до излучения отсутствовали фотоны испускаемой частоты; его называют вероятностью *спонтанного излучения*. Второе слагаемое пропорционально числу n_ω фотонов испускаемой частоты, присутствовавших до излучения: оно описывает так называемое *индуцированное, или вынужденное, излучение*. Что касается вероятности ω_- поглощения фотона, то она совпадает с вероятностью вынужденного излучения. Понятия вынужденного и спонтанного излучения ввел Эйнштейн (1917), установивший связь между вероятностями излучения и поглощения, поэтому коэффициент ω_0 называют иногда *коэффициентом Эйнштейна*.

Вычисление коэффициента ω_0 входит в компетенцию квантовой механики, так как процессы испускания и поглощения фотонов подчиняются статистическим закономерностям. Тем не менее в одном очень важном случае, когда длина волны λ фотона велика по сравнению с размерами излучающей системы, в рассматриваемом случае с размерами a атома вероятность ω_0 может быть связана с интенсивностью классического эффекта — дипольного излучения.

Мы уже говорили о такой связи в § 24.3, приравнивая интенсивность дипольного излучения

$$d\mathcal{J} = (4\pi c^3)^{-1} e^2 <\mathbf{r}^2> \omega_0^4 \sin^2 \vartheta d\omega$$

(ω_0 — частота осциллятора, ϑ — угол между дипольным моментом $\boldsymbol{\epsilon}\mathbf{r}$ и \mathbf{k} ; усреднение величины \mathbf{r}^2 производится по времени) энергии $\hbar\omega d\omega$, уносимой фотонами, где $d\omega$ — вероятность излучения фотона за 1 с (величины $d\mathcal{J}$ и $d\omega$ относятся к одному и тому же элементу телесного угла излученного фотона). Таким образом, $d\omega = d\mathcal{J}/(\hbar\omega)$.

Точный квантово-механический расчет показывает, что при $\lambda \gg a$ в этой формуле достаточно заменить частоту ω_0 диполя (осциллятора) частотой $\omega_{i \rightarrow f}$ излучаемого фотона и понимать под усредненным по периоду колебаний квадратом дипольного момента $e^2 <\mathbf{r}^2>$ величину $2e^2 |\mathbf{r}_{i \rightarrow f}|^2$, где $\mathbf{r}_{i \rightarrow f}$ — так называемый матричный элемент \mathbf{r} между состояниями i и f :

$$\mathbf{r}_{i \rightarrow f} = \int \psi_f^*(\mathbf{r}) \mathbf{r} \psi_i(\mathbf{r}) d^3r,$$

и ψ_i, ψ_f — атомные волновые функции для состояний i, f . Итак,

$$d\omega = \frac{e^2 \omega_{i \rightarrow f}^3}{2\pi \hbar c^3} |\mathbf{r}_{i \rightarrow f}|^2 \sin^2 \vartheta d\omega.$$

Проинтегрировав $d\omega$ по телесному углу $d\omega$, получим полную вероятность спонтанного излучения фотона частоты $\omega_{i \rightarrow f}$, т. е.

$$\omega_0 = (3\hbar c^3)^{-1} 4e^2 \omega_{i \rightarrow f}^3 |\mathbf{r}_{i \rightarrow f}|^2.$$

По порядку величины $|\mathbf{r}_{i \rightarrow f}|^2$ (как и $<\mathbf{r}^2>$) есть a^2 , поэтому

$$\omega_0 \sim \frac{e^2}{\hbar c} \left(\frac{a}{\lambda} \right)^2 \omega \sim \left(\frac{a}{\lambda} \right)^2 \frac{\omega}{137}.$$

Мы видим, что для рентгеновских лучей (когда λ не намного превосходит a) вероятность излучения на одном периоде волны составляет $\sim 1/100$; для видимого света она равна 10^{-10} — 10^{-11} .

Зная соотношение Эйнштейна между вероятностями поглощения и испускания света, мы можем решить следующую задачу. Пусть в некотором замкнутом объеме находится термодинамически равновесная среда, атомы которой излучают и поглощают фотоны. Ясно, что возникающий газ фотонов тоже в конце концов придет в состояние теплового равновесия. Каким будет при этом распределение фотонов по энергиям?

Для ответа на этот вопрос напомним, что распределение атомов по энергиям в состоянии теплового равновесия определяется формулой Больцмана $N = C \exp[-E/(k_B T)]$, где T — температура, k_B — постоянная Больцмана и C — нормировочная постоянная. Поэтому числа атома в состояниях с энергиями E_i и E_f соответственно составляют

$$N_i = C \exp\left(-\frac{E_i}{k_B T}\right), \quad N_f = C \exp\left(-\frac{E_f}{k_B T}\right).$$

Рассмотрим теперь переходы атомов $i \rightarrow f$ с излучением фотона частоты $\omega \equiv \omega_{i \rightarrow f}$ и обратные переходы $f \rightarrow i$ с поглощением такого же фотона. Как мы видели, вероятности таких переходов, отнесенные к одному атому, связаны соотношением

$$\omega_+/\omega_- = (n_\omega + 1)/n_\omega,$$

где n_ω — число фотонов в рассматриваемой замкнутой полости. Но у нас есть N_i атомов с энергиями E_i и N_f атомов с энергиями E_f ; значит, полные числа излучаемых и поглощаемых фотонов относятся как

$$\frac{\omega_+ N_i}{\omega_- N_f} = e^{\frac{E_f - E_i}{k_B T}} \frac{n_\omega + 1}{n_\omega} = \frac{n_\omega + 1}{e^{\frac{\hbar\omega}{k_B T}} n_\omega}.$$

В состоянии равновесия числа излучаемых и поглощаемых фотонов должны быть равны, т. е. $\omega_+ N_i = \omega_- N_f$, поэтому $n_\omega + 1 = n_\omega \exp[\hbar\omega/(k_B T)]$. Отсюда для n_ω получим

$$n_\omega = \frac{1}{e^{\hbar\omega/(k_B T)} - 1}.$$

Эту формулу называют формулой или распределением Планка. (Мы уже неоднократно использовали ее не только для фотонов, но и для фононов или магнонов. Она была установлена М. Планком в 1900 г.)

Представим себе теперь, что из полости удалены все атомы. Атомы стенок полости тоже излучают и поглощают фотоны, и если стенки находятся в состоянии термодинамического равновесия, то в равновесие с ними придет и фотонный газ в полости. В этом случае фотоны излучаются и поглощаются макроскопическими стенками, энергетический спектр которых практически непрерывен, поэтому непрерывен и фонный спектр. Мы приходим к важному выводу:

находящийся в равновесии при температуре T фотонный газ (так называемое равновесное, или черное, излучение) независимо от его происхождения всегда характеризуется плановским распределением фотонов по энергиям.

Формулу Планка можно получить и иначе, если учесть, что число фотонов обязательно дискретно, поэтому энергия излучения данной частоты ω может принимать только дискретные значения $\varepsilon_n = n\hbar\omega$ ($n = 0, 1, \dots$). Будем рассматривать такое излучение как некоторое тело или физическую систему; вероятность того, что в состоянии равновесия эта система обладает энергией ε_n , пропорциональна $\exp[-\varepsilon_n/(k_B T)]$. Отсюда легко найти среднее значение энергии фотонов ε_n :

$$\langle \varepsilon_n \rangle = \frac{\sum_{n=0}^{\infty} \varepsilon_n \exp[-\varepsilon_n/(k_B T)]}{\sum_{n=0}^{\infty} \exp[-\varepsilon_n/(k_B T)]} .$$

или (среднее статистическое) число фотонов частоты ω :

$$n_\omega = \frac{\sum_{n=0}^{\infty} n \exp[-n\hbar\omega/(k_B T)]}{\sum_{n=0}^{\infty} \exp[-n\hbar\omega/(k_B T)]} .$$

В самом деле, обозначим $Z(\xi)$ сумму геометрической прогрессии:

$$Z(\xi) = \sum_{n=0}^{\infty} e^{-n\xi} = \frac{1}{1 - e^{-\xi}}, \quad \xi \equiv \frac{\hbar\omega}{k_B T} .$$

Тогда $\langle \varepsilon_n \rangle = \hbar\omega n_\omega$ и

$$n_\omega = -\frac{1}{Z(\xi)} \frac{dZ(\xi)}{d\xi} = \frac{e^{-\xi}}{1 - e^{-\xi}} = \frac{1}{e^{\hbar\omega/(k_B T)} - 1} .$$

Мы снова пришли к распределению Планка.

Найдем интенсивность теплового излучения. Рассмотрим с этой целью полость в виде прямоугольного параллелепипеда со сторонами A, B, C . На границе полости должны выполняться определенные

граничные условия для полей, но, чтобы этот вопрос здесь детально не рассматривать, можно поступить несколько иначе, а именно считать, что поля периодичны с периодами A , B и C вдоль осей x , y и z . Так как поля в полости рассматриваемой симметрии имеют вид плоских волн $\exp(\mathbf{ikr} - i\omega t)$, где \mathbf{k} — волновой вектор и $\omega = ck$, то зависимость от координаты x определяется множителем $\exp(ik_x x)$. Чтобы эта функция была периодической с периодом A , должно выполняться условие $k_x A = 2\pi n_1$, где n_1 — целое число. Аналогичные условия должны выполняться для проекций \mathbf{k} на оси y и z . Отсюда следует, что

$$k_x = 2\pi n_1/A, \quad k_y = 2\pi n_2/B, \quad k_z = 2\pi n_3/C.$$

Эти формулы позволяют пронумеровать волны в полости. Мы видим, что каждой тройке целых чисел n_1, n_2, n_3 соответствует один волновой вектор и, следовательно, две волны, отличающиеся поляризацией \mathbf{e} (напомним, что $\mathbf{ek} = 0$). Ясно, что наше условие периодичности при больших A, B, C практически не нарушает непрерывности спектра, так как проекции волновых векторов двух волн с соседними n_1, n_2, n_3 отличаются на $2\pi/A, 2\pi/B, 2\pi/C$, а эти величины стремятся к нулю при $A, B, C \rightarrow \infty$.

Мы можем теперь определить число волн в интервалах $(k_x, k_x + \Delta k_x)$, $(k_y, k_y + \Delta k_y)$, $(k_z, k_z + \Delta k_z)$. Число волн равно, очевидно, $\Delta N = 2\Delta n_1 \Delta n_2 \Delta n_3$, а так как $\Delta n_1 = A \Delta k_x / (2\pi)$ и т. д., то $\Delta N = 2(2\pi)^{-3} V \Delta k_x \Delta k_y \Delta k_z$, где $V = ABC$ — объем полости.

Переходя к дифференциалам, получим

$$dN = 2(2\pi)^{-3} V d^3 k \quad (d^3 k \equiv dk_x dk_y dk_z).$$

В числителе здесь стоит элемент объема в \mathbf{k} -пространстве. Удобно использовать не декартовы, а сферические координаты и перейти от k к частоте $\omega = ck$; мы получим тогда

$$dN = V \omega^2 d\omega / (\pi^2 c^3).$$

Энергия этих волн, т. е. волн в интервале частот $(\omega, \omega + d\omega)$, составляет

$$d\mathcal{E}_\omega = \hbar \omega n_\omega dN = \frac{V \hbar \omega^3 d\omega}{\pi^2 c^3 \left(\exp \frac{\hbar \omega}{k_B T} - 1 \right)}.$$

Проинтегрировав это выражение по частоте, получим полную энергию равновесного излучения:

$$\mathcal{E} = \frac{V}{\pi^2 c^3} \int_0^\infty \frac{\hbar \omega^3 d\omega}{\exp \frac{\hbar \omega}{k_B T} - 1}.$$

Вводя вместо ω новую переменную интегрирования $\xi = \hbar \omega / (k_B T)$, перепишем это выражение в виде

$$\mathcal{E} = \frac{V (k_B T)^4}{\pi^2 c^3 \hbar^3} \zeta, \quad \zeta \equiv \int_0^\infty \frac{\xi^3 d\xi}{e^\xi - 1}.$$

Входящий сюда интеграл ζ представляет собой просто числовую константу — она равна $\pi^4/15$. Таким образом,

энергия черного излучения пропорциональна четвертой степени температуры.

Что касается зависимости от формы и размеров полости, то в выражение для \mathcal{E} входит только объем полости; этот результат относится не только к полости в форме параллелепипеда, но и к полости произвольной формы.

Как мы видим, интенсивность равновесного излучения зависит только от температуры источника и не зависит от химического состава источника и других факторов. Поэтому такое излучение характеризуется большой стабильностью интенсивности и используется как эталон.

Любопытным фактом является то, что вся наша Вселенная наполнена не имеющим источника равновесным радиоизлучением, соответствующим температуре $T = 2,7$ К. Это излучение находилось, по видимому, в равновесии с веществом горячей Вселенной несколько миллиардов лет тому назад. Тогда его температура была $\sim 3 \cdot 10^8$ К. Происходящее все время расширение Вселенной сопровождается увеличением всех длин (в том числе длины волны излучения) и, следовательно, охлаждением излучения, средняя частота которого $\langle \omega \rangle \sim \sim k_B T / \hbar$, как и определяющая ее температура T , обратно пропорциональна средней длине волны. Таким образом, расширение Вселенной привело к понижению температуры излучения до нынешнего уровня. Сам факт существования этого излучения связан с ранними этапами эволюции Вселенной, поэтому это излучение называют *реликтовым*

24.9. Микроволновые и оптические лазеры

Существование наряду со спонтанным вынужденного излучения лежит в основе работы генераторов и усилителей электромагнитных волн, создающих пучки волн колоссальной интенсивности и когерентности, — *мазеров* и *лазеров* (мазер — это сокращение из начальных букв английских слов microwave amplification by stimulated emission radiation — микроволновое усиление, обусловленное стимулированным излучением). Оптический мазер называют лазером. Этот метод генерирования был независимо открыт Н. Г. Басовым и А. М. Прохоровым, а также Гордоном, Цайгером и Таунсом в 1954—1955 гг. Возможность применения метода для усиления радиоволн была указана ранее (1951) В. А. Фабрикантом, М. М. Вудинским и Р. А. Бутаевой.

Изучая в гл. 16 общие свойства электрической и магнитной восприимчивостей, мы пришли к выводу, что мнимые части этих величин должны быть положительными, в соответствии с тем что среда может поглощать, но не излучать электромагнитную энергию. Этот вывод

относится к средам, находящимся в состоянии полного равновесия. Но в гл. 17 мы видели, что в неравновесной среде (системе плазма — пучок) возможно нарастание (а не затухание) колебаний, а при таком нарастании пучковой неустойчивости мнимая часть электрической восприимчивости делается (в некотором интервале частот) отрицательной.

Среды с отрицательной мнимой частью восприимчивостей (иногда их называют активными) могут быть использованы для генерирования и усиления колебаний.

Разъясним принципы создания такой среды и ее использования в мазерах и лазерах. Среда неактивна, если вероятность w_i того, что она находится в состоянии с энергией ε_i , убывает с ростом ε_i (в таких случаях говорят о *нормальной населенности уровней* энергии). В частности, нормальной населенностью характеризуется среда в состоянии теплового равновесия: для такой среды, как мы знаем, w_i определяется формулой Больцмана $w_i \sim \exp[-\varepsilon_i/(k_B T)]$.

Для того чтобы среда стала активной, нужно повысить вероятность возбужденных состояний атомов, или, как говорят иначе, добиться *инверсии населенности уровней*. (Разумеется, достаточно, чтобы *инверсная*, или *обратная*, населенность реализовывалась хотя бы для двух уровней.)

В случае мазеров с этой целью используются два метода. Первый из них (возможный только для газов) заключается в том, чтобы с помощью внешних полей пространственно отделить атомы в возбужденном состоянии от атомов в основном состоянии; совокупность возбужденных атомов и будет тогда активной средой.

Второй метод (возможный как для газов, так и для жидкостей и твердых тел) сводится к «перебрасыванию» атомов из основного состояния в возбужденное, без их пространственного разделения.

Допустим теперь, что тем или иным методом получена активная среда. Как она может быть использована для генерирования электромагнитных волн? Если активную среду предоставить самой себе, то ее возбужденные атомы «высветятся», т. е. перейдут в основное состояние и излучат при этом электромагнитные волны с частотой $\omega_{ji} = (\varepsilon_j - \varepsilon_i)/\hbar$. Но такой спонтанный переход будет беспорядочным процессом, так как разные атомы излучают независимо друг от друга и фазы излученных волн никак не скоррелированы между собой.

Есть, однако, возможность сделать так, чтобы отдельные атомы излучали синфазно, т. е. с одной и той же фазой (и, естественно, с одной и той же частотой). Для этого атомные излучатели нужно «подстегнуть», т. е. не допуская (правильнее, не принимая во внимание) спонтанности излучения, иметь «спусковые крючки» излучения, действующие синхронно. Этого достигают, используя эффект *вынужденного излучения*. Эффект состоит в том, что атом в возбужденном состоянии можно заставить излучать вынужденно под влиянием падающего на него излучения даже в том случае, когда атом и «не думает» спонтанно излучать. При этом фаза вынужденного излучения атома не случайная, а строго зафиксированная величина — она совпадает

с фазой вынуждающего излучения. Особенно велик этот эффект при резонансе, когда частота вынуждающего излучения совпадает с частотой атомного перехода. Используя это явление, мы достигаем того, что все возбужденные атомы излучают когерентно, т. е. появляется возможность создания генератора электромагнитных волн, состоящего из колоссального числа атомных излучателей, излучающих синхронно и синфазно.

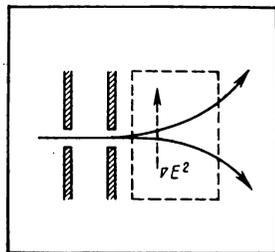
Разъясним подробнее действие газового мазера, использующего пучок молекул аммиака NH_3 . Молекула NH_3 имеет много энергетических уровней, среди которых есть два уровня 1 и 2 с близкими энергиями ε_1 и ε_2 (так, что $\varepsilon_2 - \varepsilon_1 \equiv 2A = 10^{-4}$ эВ, что соответствует частоте $\omega/(2\pi) = (\varepsilon_2 - \varepsilon_1)/(2\pi\hbar) = 2,4 \cdot 10^{10}$ с $^{-1}$ или длине волны $\lambda = 1,25$ см), причем существенно то, что молекула NH_3 в состояниях 1 и 2 обладает электрическим дипольным моментом \mathbf{d} . Поэтому если внести молекулу NH_3 в электрическое поле E , то энергии состояний 1 и 2 изменятся и примут значения

$$\begin{aligned} \varepsilon'_1 &= \varepsilon_0 - \sqrt{A^2 + E^2 d^2}, \\ \varepsilon'_2 &= \varepsilon_0 + \sqrt{A^2 + E^2 d^2}, \end{aligned}$$

где $\varepsilon_0 = 1/2(\varepsilon_1 + \varepsilon_2)$.

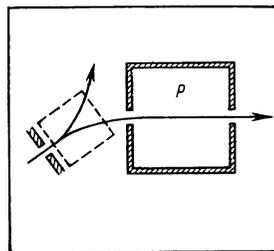
Представим себе теперь, что узкий пучок газа NH_3 пропускают через область с сильным неоднородным электрическим полем, меняющимся в направлении, перпендикулярном движению пучка. Вследствие неоднородности поля на молекулы действуют силы: молекулы с энергией ε'_1 отклоняются в сторону больших полей (скажем, вверх), а молекулы с энергией ε'_2 — в сторону меньших полей (вниз). В результате пучок молекул NH_3 расщепляется на два пучка молекул с энергиями ε'_1 и ε'_2 (рис. 24.9, 24.10). После разделения пучков уже нет нужды в том, чтобы молекулы двигались в постоянном электрическом поле. При выходе из электрического поля избыточная энергия $\varepsilon'_1 - \varepsilon_2$ превращается в кинетическую энергию. Величина $\varepsilon'_1 - \varepsilon_2$ мала (пропорциональна E^2), и мы не будем ее в дальнейшем учитывать.

Теперь мы должны заставить все молекулы с энергией ε_2 одновременно и когерентно излучать фотоны. С этой целью пучок этих молекул пропускают через резонатор P (рис. 24.10) с собственной частотой $\omega_0 = (\varepsilon_2 - \varepsilon_1)/\hbar$. Задача заключается в том, чтобы молекулы из возбужденного состояния с энергией ε_2 перешли, находясь в резонаторе, в основное состояние с энергией ε_1 , причем переход должен проис-



24.9.

Разделение пучка молекул газа неоднородным электрическим полем



24.10.

Схематическое изображение газового мазера

ходить синхронно и синфазно. С этой целью в резонатор вводят переменное поле $\tilde{E} = \tilde{E}_0 \cos \omega t$, которое вызывает переход молекул $2 \rightarrow 1$, а вместе с ним вынужденное излучение на частоте ω_0 .

Вероятность перехода особенно велика при выполнении резонансного условия $\omega = \omega_0$. В этом случае вероятность того, что молекула с энергией ε_2 , попав в резонатор в момент времени $t = 0$, перейдет к моменту времени t в состояние с энергией ε_1 , определяется формулой $\omega_{2 \rightarrow 1}(t) = \sin^2(\tilde{E}td/\hbar)$. Из этой формулы следует, что время прохождения молекулой резонатора $T = L/v$ (L — длина резонатора, v — скорость молекулы) нужно сделать таким, чтобы выполнялось равенство $\tilde{E}Td/\hbar = \pi/2$. Тогда вероятность того, что молекула, войдя в резонатор с энергией ε_2 , выйдет из него с энергией ε_1 , равна единице. Иными словами, в этом случае все молекулы, вошедшие в резонатор в возбужденном состоянии 2, выйдут из него в основном состоянии 1. При этом каждая из них отдает резонатору энергию $\hbar\omega_0 = \varepsilon_2 - \varepsilon_1$ в виде энергии электромагнитной волны.

Остановимся на характере поля \tilde{E} , вызывающего вынужденное излучение молекул. Мы его считали заданным. В этом случае наш мазер, т. е. система, состоящая из разделителя пучка молекул и резонатора, через который проходит пучок возбужденных молекул, является усилителем поля \tilde{E} , но не генератором.

Однако в этом же приборе заложена возможность генерирования волн на частоте ω_0 . Действительно, в резонаторе всегда есть поле частоты ω_0 , хотя его амплитуда может быть ничтожно малой (например, поле, обусловленное тепловыми флуктуациями). Это поле вызовет вынужденное излучение, которое усилит исходное поле. В более сильном поле будет происходить более интенсивное вынужденное излучение, которое еще больше усилит исходное поле, и т. д. Таким образом, в принципе здесь возможно лавинообразное усиление сколь угодно малого поля частоты ω_0 .

Но для того, чтобы был возможен такой эффект самовозбуждения, должны выполняться некоторые условия. Основное из них сводится к тому, чтобы возникающая энергия электромагнитного поля превосходила потери энергии в резонаторе. Для этого число молекул N_2 в состоянии 2, влетающих в резонатор за 1 с, должно превосходить некоторое критическое значение. В самом деле, мощность, излучаемая пучком,

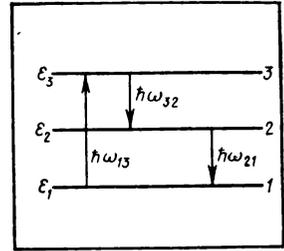
$$P = \hbar\omega_0 N_2 \omega_{2 \rightarrow 1}(t),$$

где $\omega_{2 \rightarrow 1}(t)$ — вероятность перехода $2 \rightarrow 1$ к времени t . Предполагая t достаточно малым, мы положим $\omega_{2 \rightarrow 1}(t) = (\pi\tilde{E}_0 Td/\hbar)^2$. Потери мощности в резонаторе равны $\omega_0 W/Q$, где Q — добротность, $W = V\tilde{E}_0^2/(16\pi)$ — энергия, запасенная в резонаторе, и V — объем резонатора. Полагая в выражении для $\omega_{2 \rightarrow 1}(t)$ время t равным L/v , получим из равенства $P = \omega_0 W/Q$ критическое значение $N^{(c)}$ величины N_2 :

$$N^{(C)} = \frac{\hbar \omega^2 S}{4\pi^2 L Q d^2},$$

где S — площадь поперечного сечения резонатора. Оценка показывает, что критическое значение $N^{(C)}$ составляет по порядку величины 10^{13} молекул в секунду.

Разъясним теперь метод получения активной среды с помощью инверсии населенности уровней без пространственного разделения атомов. Рассмотрим трехуровневый (рис. 24.11) твердотельный мазер, в котором используется система с тремя уровнями ε_1 , ε_2 , ε_3 (он был предложен Н. Бломбергеном и А. Прохоровым).



24.11. Переходы в трехуровневом мазере

Система поглощает фотон частоты ω_{31} и переходит из состояния 1 в состояние 3. Время жизни этого состояния очень невелико, и система быстро, испустив фотон частоты ω_{32} , переходит на «рабочий» уровень 2, время жизни которого велико, благодаря чему населенность этого уровня может быть сделана сравнимой или даже выше населенности основного состояния 1. На этих двух уровнях и работает мазер, т. е. он усиливает и генерирует частоту $\omega_{21} = (\varepsilon_2 - \varepsilon_1)/\hbar$; для твердотельных трехуровневых мазеров она обычно лежит в оптическом диапазоне. Объемный резонатор сделать при этом невозможно (из-за его слишком малых размеров), поэтому в оптических мазерах — лазерах применяются два зеркала, между которыми генерируются стоячие световые волны.

В трехуровневом мазере может быть достигнута стационарная инверсия населенности уровней. Чтобы убедиться в этом, напишем уравнение баланса для населенностей N_1 , N_2 , N_3 уровней 1, 2, 3. Обозначим w_{ij} вероятность перехода $i \rightarrow j$, обусловленную колебаниями кристаллической решетки и другими релаксационными процессами. Пусть далее W_{ij} — вероятность вынужденного перехода $i \rightarrow j$, обусловленного приложенным полем частоты ω_{ij} (предполагается, что присутствуют сильное насыщающее поле частоты ω_{31} и слабое поле частоты ω_{32}). Тогда изменение во времени населенности N_3 уровня 3 определится уравнением

$$\frac{dN_3}{dt} = w_{13} N_1 - w_{31} N_3 + w_{23} N_2 - w_{32} N_3 - W_{31} (N_3 - N_1) - W_{32} (N_3 - N_2).$$

Аналогичные уравнения справедливы для изменений населенности уровней 1 и 2.

Заметим, что вероятности w_{ij} и w_{ji} не независимы, а связаны между собой в силу принципа детального равновесия. В самом деле, равновесные населенности уровней \bar{N} удовлетворяют соотношениям $\bar{N}_i w_{ij} = \bar{N}_j w_{ji}$, но в тепловом равновесии, согласно распределению Больцмана, $\bar{N}_i \sim \exp[-\varepsilon_i/(k_B T)]$, поэтому

$$\frac{\omega_{ji}}{\omega_{ij}} = \exp\left(-\frac{\hbar\omega_{ij}}{k_B T}\right) \approx 1 - \frac{\hbar\omega_{ij}}{k_B T}$$

(предполагается, что $\hbar\omega \ll k_B T$). В нулевом приближении по малому параметру $\hbar\omega_{ij}/(k_B T)$ населенности всех трех уровней равны, поэтому можно там, где $N_{1,2,3}$ множатся на этот параметр, положить $N_1 = N_2 = N_3 = 1/3 N$, где $N = N_1 + N_2 + N_3$. Таким образом, мы получим

$$\begin{aligned} \frac{dN_3}{dt} = & \omega_{13} \left(N_1 - N_3 - N \frac{\hbar\omega_{31}}{3k_B T} \right) + \\ & + \omega_{23} \left(N_2 - N_3 - N \frac{\hbar\omega_{32}}{3k_B T} \right) - W_{31} (N_3 - N_1) - W_{32} (N_3 - N_2). \end{aligned}$$

Аналогичные уравнения справедливы для dN_1/dt и dN_2/dt .

Нас интересует случай стационарных населенностей, когда $N_{1,2,3}$ не зависит от времени. Обычно вероятность W_{31} значительно больше вероятности W_{32} и всех вероятностей тепловых переходов (т. е. всех ω). При этом для стационарных населенностей справедливо соотношение

$$N_3 - N_2 = \frac{\hbar N}{3k_B T} \frac{\omega_{32}\omega_{32} - \omega_{21}\omega_{21}}{\omega_{32} + \omega_{12} + W_{32}}.$$

Отсюда следует условие инверсии: $\omega_{32}\omega_{32} > \omega_{21}\omega_{21}$. Умножив $N_3 - N_2$ на $\hbar\omega_{32}W_{32}$, найдем выходную мощность на частоте ω_{32} :

$$P = \hbar\omega_{32}(N_3 - N_2)W_{32}.$$

Примером трехуровневой системы, широко используемой в качестве лазера, является кристалл рубина — корунда Al_2O_3 с небольшой примесью оксида хрома Cr_2O_3 (с которым и связан характерный цвет рубина). В решетке рубина хром присутствует в виде трехзарядных ионов Cr^{3+} ; в спектре этих ионов и имеются три используемых уровня. Длина волны накачки у рубина 556 нм, длина излучаемых волн 692,9 и 694,3 нм (красный цвет). Рубиновый лазер может работать как в импульсном, так и в непрерывном режиме. В принципе, лазеры позволяют получать излучение в спектральном диапазоне от далекой инфракрасной области до ультрафиолетовой с пиковой мощностью в импульсе до гигаватта или непрерывной мощностью, превышающей сотни ватт. Лазеры позволяют получать сигналы очень высокой спектральной чистоты, со стабильностью частоты порядка 10^{-14} .

ОСНОВНЫЕ ФОРМУЛЫ

Формула Резерфорда	$d\sigma = \left(\frac{Zze^2}{2Mv^2}\right)^2 \frac{d\omega}{\sin^4(\theta/2)}$
--------------------	---

Средний квадрат угла рассеяния	$\langle \theta^2 \rangle = \frac{8\pi NIZ^2z^2e^4}{p^2v^2} \ln \frac{\alpha_0 pv}{2Z^{1/3}ze^2}$
--------------------------------	---

Ионизационные потери (формула Бете)

$$-\frac{dE}{dx} = \frac{4\pi z^2 e^4 n}{m_e v^2} \left(\ln \frac{2m_e v^2}{\langle I \rangle \left(1 - \frac{v^2}{c^2}\right)} - \frac{v^2}{c^2} \right)$$

**Энергия излучения Вавилова — Черенкова (формула Тамма—Франка)
Формула Комптона**

$$\mathcal{E} = \frac{e^2 l}{c^2} \int_{(\beta N > 1)} \left(1 - \frac{1}{\beta^2 N}\right) \omega d\omega$$

$$\omega_2 = \frac{\omega_1}{1 + \frac{\hbar \omega_1}{m_e c^2} (1 - \cos \theta)}$$

Сечение комptonовского рассеяния

$$\sigma = \pi r_0^2 \frac{m_e c^2}{\hbar \omega_1} \left(\ln \frac{2\hbar \omega_1}{m_e c^2} + \frac{1}{2} \right) (\hbar \omega_1 \gg m_e c^2),$$

$$\sigma = \frac{8\pi}{3} r_0^2 \left(1 - 2 \frac{\hbar \omega_1}{m_e c^2}\right) (\hbar \omega_1 \ll m_e c^2)$$

Сечение образования пар

$$\sigma_{\Pi} = \frac{2Z^2 e^6}{9m_e^2 c^5 \hbar} \left\{ 14 \ln(183Z^{-1/3}) - \frac{1}{3} \right\}$$

Вероятности излучения и поглощения фотона

$$\omega_+ = (n_\omega + 1) \omega_0, \quad \omega_- = n_\omega \omega_0$$

Вероятность спонтанного излучения

$$\omega_0 = \frac{4e^2}{3\hbar c^3} \omega^3_{i \rightarrow f} |\mathbf{r}_{i \rightarrow f}|^2$$

Интенсивность излучения черного тела

$$\mathcal{E} = \frac{\pi^4}{15} \frac{V (k_B T)^4}{\pi^2 c^3 \hbar^3}$$

Изучив конкретные электрические и магнитные явления, мы обратимся снова к становлению теории электромагнетизма Максвелла — на этот раз главным образом с целью разъяснения взаимосвязи теории Максвелла с другими фундаментальными физическими теориями и обусловленного этой взаимосвязью дальнейшего развития учения об электромагнетизме.

Мы исходили из закона Кулона для электростатического взаимодействия зарядов и закона Ампера для магнитостатического взаимодействия токов. Сами по себе эти законы, несмотря на их новизну и важность, по существу, не вносили ничего принципиально нового в физическую картину мира: она оставалась ньютоновской, т. е. целиком базирующейся на универсальной ньютоновской механике. Новые законы, казалось бы, только подчеркивали универсальность механики Ньютона, так как они описывали два новых типа возможных сил, силы же, по Ньютону, могут быть самыми различными. Иными словами, электростатика и электродинамика постоянных токов целиком вписывались в общую динамику Ньютона и могли, по существу, рассматриваться как ее конкретные приложения. Дело в том, что, несмотря на возможность формального введения понятий электрического и магнитного полей, эти понятия еще не имели подлинно физического смысла.

Ситуация коренным образом изменилась после экспериментального открытия Фарадеем явления и закона электромагнитной индукции и теоретического предсказания Максвеллом возможности создания магнитного поля переменным электрическим полем (идея тока смещения). Получалось так, что переменные поля могут, поддерживая друг друга, существовать независимо от зарядов и токов. Иными словами, электрическое и магнитное поля стали сами по себе физической реальностью, и перестали быть чисто формальным понятием, атрибутом зарядов и токов, понятием, которым можно пользоваться, но можно и не пользоваться.

Реальность полей нашла свое отражение в дифференциальных уравнениях Максвелла, которые представляют собой законы движения полей, так же как законы Ньютона являются законами механического движения вещества.

Электромагнитное поле заняло свое место рядом с веществом.

Физическая картина мира стала базироваться на двух фундаментальных теориях — механике Ньютона и электродинамике Максвелла.

На основе учения об электромагнетизме сперва была создана электротехника и в дальнейшем радиотехника, что привело к новой научно-технической революции.

Золотой век в физике продолжался недолго — всего около 50 лет. В начале нашего столетия стало ясно, что оба столпа физической науки — механика Ньютона и электродинамика Максвелла — находятся в резком противоречии между собой. Противоречие это связано прежде всего с тем, что согласно максвелловской электродинамике в вакууме могут распространяться свободные электромагнитные волны и, в частности, свет, причем с некоторой определенной, универсальной скоростью. Но в ньютоновской механике скорость материальных тел не может быть абсолютной — она всегда зависит от системы отсчета и подчиняется классическому закону сложения скоростей, связанному с классическими преобразованиями Галилея. По отношению к преобразованиям Галилея законы Ньютона инвариантны; этим выражается невозможность выделения избранной инерциальной системы отсчета — все эти системы эквивалентны и неотличимы друг от друга. Что же касается уравнений Максвелла, то они не инвариантны по отношению к преобразованиям Галилея и поэтому, казалось бы, допускают существование избранной системы отсчета, в которой именно и распространяется свет. Однако опыт Майкельсона показал, что такой системы нет, так как скорость света не зависит ни от движения источника света, ни от движения наблюдателя. Отсюда напрашивался вывод, что преобразования Галилея, с которыми отождествлялась эквивалентность всех инерциальных систем отсчета, не могут считаться универсальными.

К этому следует добавить, что из электродинамики Максвелла вытекал вывод, что третий закон Ньютона — закон равенства действия и противодействия — не может рассматриваться как универсальный. Действительно, если два тела связаны между собой световым лучом, то, так как свет несет с собой импульс, но распространяется с конечной скоростью, оба тела не могут действовать друг на друга в один и тот же момент времени с одинаковыми силами.

Противоречие между механикой Ньютона и электродинамикой Максвелла было преодолено специальной теорией относительности Эйнштейна, но преодолено большой ценой: пришлось признать, что время не может считаться абсолютным, как это предполагается в механике Ньютона, и является относительным и текущим по-разному в разных системах отсчета. Связь между моментами времени и точками пространства, отвечающими одному и тому же событию в разных инерциальных системах отсчета, определяется преобразованиями Лоренца. По отношению к этим преобразованиям инвариантны уравнения Максвелла, но не уравнения Ньютона. Поэтому уравнения механики были переформулированы Эйнштейном так, чтобы и они стали инвариантными по отношению к преобразованиям Лоренца.

Таким образом, противоречие между классической механикой Ньютона и электродинамикой Максвелла было преодолено теорией относительности Эйнштейна, причем разрешение противоречия затронуло только механику, электродинамика же Максвелла осталась неизменной. Замена классической механики Ньютона релятивистской механикой Эйнштейна имела реальный смысл только в области больших скоростей движения частиц, порядка скорости света, в области же нерелятивистских движений обе классические теории практически эквивалентны.

Между тем и в нерелятивистской области возникал новый кризис, связанный на сей раз с существованием атомов.

Атом состоит из ядра, около которого движутся электроны. Скорости их нерелятивистские, поэтому, казалось бы, мы вправе пользоваться классической механикой Ньютона при изучении движения электрона. С другой стороны, движущийся с ускорением электрон должен согласно электродинамике Максвелла излучать электромагнитные волны, волны же уносят с собой энергию, т. е. атом должен терять энергию и его энергии хватит на излучение только на ничтожное время, порядка 10^{-8} с, после чего атом должен прекратить свое существование. Иными словами, в рамках классической физики — классической механики и классической электродинамики — невозможно понять устойчивость атома.

Чтобы найти выход из тупика, понадобилось, как и при создании теории относительности, пересмотреть основные понятия классической физики. Только на сей раз анализу подверглись не представления о пространстве и времени, а классические представления о характере движения.

Сущность этих представлений заключается в том, что движение частицы всегда происходит по определенной траектории. Понятие движения неотделимо от понятия траектории.

Но что, собственно, означает движение по траектории? Это значит, что в каждый момент времени имеют смысл понятия координаты частицы и ее скорости или импульса. Между тем одновременное использование этих понятий лишено всякого смысла. Для пояснения этого положения рассмотрим следующий мысленный эксперимент (эксперимент Гейзенберга).

Пусть вдоль прямой движется электрон со строго заданной скоростью. Нам же нужно определить его положение. Для этого осветим электрон, т. е. бросим на него сходящийся пучок света. Изображение, создаваемое таким пучком, не будет точечным, а из-за волновой природы света должно иметь некоторые размеры порядка длины волны света. Поэтому точность локализации электрона равна по порядку величины длине волны падающего на электрон света. Уменьшая длину волны, можно достичь сколь угодно точной локализации электрона.

Однако «освещение» электрона не есть «безобидный» процесс. Так как свет несет с собой импульс, а при освещении электрона происходит взаимодействие света с электроном, то электрон приобретет некоторый импульс, и, естественно, изменится его скорость. Измене-

ние это тем больше, чем меньше длина волны света, т. е. чем точнее локализовано положение электрона.

Итак, чем точнее мы определяем положение электрона, тем больше возмущаем его движение, т. е. изменяем его скорость. Изменение импульса электрона при освещении светом и неточность в его локализации таковы, что их произведение не может быть меньше некоторой универсальной константы — так называемой постоянной Планка.

Это важнейшее соотношение, установленное впервые Гейзенбергом (его называют соотношением неопределенностей Гейзенберга), оказывается справедливым не только в нашем мысленном опыте, но и в любом другом эксперименте, ставящем себе целью одновременное точное определение положения и значения импульса частицы. Иными словами, соотношение неопределенностей Гейзенберга носит универсальный характер.

Но не означает ли оно, что наши возможности познания внешнего мира ограничены? Такой вывод был бы неправильным. Соотношение неопределенностей Гейзенберга означает лишь, что понятие траектории, строго говоря, лишено смысла и им не всегда можно пользоваться. Но это ни в коем случае не означает, что наши возможности познания мира ограничены. Отказ от классического понятия траектории нельзя отождествлять с отказом от возможностей познания внешнего мира. Не ограничение возможностей познания мира, а ограничение применимости определенных классических понятий — вот что означает это соотношение.

Если формально принять постоянную Планка (ее называют также квантовой постоянной) равной нулю, то исчезнет корреляция в определении координаты и импульса и мы придем к классическому понятию траектории.

Таким образом, классическая механика должна быть предельным случаем более общей физической теории, соответствующим обращению в нуль квантовой постоянной. Такой более общей физической теорией является квантовая механика, следующая после механики Ньютона, электродинамики Максвелла, специальной и общей теории относительности Эйнштейна, великая физическая теория, сбъяснившая и описавшая с единой точки зрения огромный круг явлений, относящихся к квантовой физике, и в частности к атомной.

Квантовая механика сопоставляет каждой физической величине — импульсу частицы, ее энергии, моменту и т. д. — не просто числовые значения, а некоторый оператор (эрмитовский), собственные значения которого образуют спектр возможных, наблюдаемых на опыте значений рассматриваемой величины.

Каждое состояние любой динамической системы, например атома, описывается в квантовой механике волновой функцией, удовлетворяющей определенному уравнению — уравнению Шредингера. Волновая функция частицы, например электрона, в каком-либо состоянии непосредственно связана с вероятностной концепцией квантовой механики — квадрат модуля волновой функции определяет распределение вероятностей нахождения частицы в том или ином месте пространства.

Квантовая механика должна быть последовательной и непротиворечивой физической теорией. Для этого необходимо, чтобы квантовыми свойствами обладали не одни электроны, а вся материя в целом, в частности и электромагнитное поле. Действительно, если бы свет не обладал квантовыми свойствами, то, сделав интенсивность пучка света в рассмотренном выше мысленном эксперименте Гейзенберга достаточно слабой, мы смогли бы определить точное положение электрона (выбрав длину волны света достаточно малой), не изменив при этом его импульса. Иными словами, соотношение неопределенностей, лежащее в основе квантовой механики, не выполнялось бы.

Но свет обладает квантовыми свойствами. Это значит, что его можно представить себе состоящим из отдельных частиц — фотонов, движущихся с определенной скоростью и обладающих определенной энергией и определенным импульсом. Для фотона, так же как и для электрона, может быть построена квантовая механика.

Учитывая корпускулярную «структуру» света, можно показать, что пучок света, падающий на электрон, в рассмотренном опыте Гейзенберга будет оказывать давление на электрон при любой интенсивности света и соотношение неопределенностей будет «спасено».

Итак, свет характеризуется как волновыми, так и корпускулярными свойствами, но эти свойства дополнительные. Это значит, что приборы, которые устанавливают одно из этих свойств, действуют на свет таким образом, что второе свойство в этих измерениях не проявляется. Аналогичная ситуация имеет место для электронов и любых других частиц, ибо все они обладают как корпускулярными, так и волновыми свойствами. Эти свойства электронов проявляются, например, в опытах по интерференции электронов в кристаллах (или в мысленных опытах по прохождению электронов через два отверстия в экране). Но волновые свойства электронов перестают проявляться, как только приборы устанавливают их корпускулярные свойства.

Теперь мы можем разъяснить устойчивость атома согласно квантовой механике.

Если атом находится в квантовом состоянии с наименьшей энергией, то он не может излучать электромагнитные волны, ибо излучение происходит отдельными фотонами, уносящими конечные порции энергии. Электрон же, находясь в основном состоянии (в состоянии с наименьшей энергией), не может передать своей энергии. Напротив, в возбужденных состояниях передача энергии возможна, и в результате излучения, т. е. вылета фотона, атом теряет часть своей энергии и переходит в основное состояние.

Квантовая механика (вместе с предположением о корпускулярной природе света) дает возможность определить среднее значение времени, в течение которого происходит эта внутренняя перестройка атома, т. е. определить среднее время жизни атома в возбужденном состоянии.

В рамках квантовой механики можно найти и значения интенсивности излучения, описать свойства поглощения света атомами, молекулами. При этом используется классическая электродинамика Максвелла, к которой присоединяется квантовое условие, учитывающее корпускулярные свойства света.

Таким образом, возникает новая физическая теория — к в а н т о в а я э л е к т р о д и н а м и к а, представляющая собой квантовую теорию электромагнитных взаимодействий.

Развитию квантовой электродинамики предшествовало создание Дираком релятивистской квантовой механики электрона, в которой объединялись квантовое описание электрона с требованием специальной теории относительности об инвариантности уравнений любой физической теории относительно преобразований Лоренца.

В релятивистской квантовой механике электрон характеризуется не одной волновой функцией, как в нерелятивистской квантовой механике, а четырьмя, образующими вместе единую математическую величину — биспинор, преобразующийся по определенному закону при преобразованиях Лоренца. Четыре компоненты биспинора подчиняются, как функции координат и времени, определенным дифференциальным уравнениям, установленным Дираком и носящим его имя. На основе этих уравнений, являющихся релятивистским обобщением уравнения Шредингера, сделано два важнейших физических вывода.

Во-первых, установлено, что электрон должен обладать внутренним моментом количества движения — спином, равным $1/2$ (в единицах квантовой постоянной). Таким образом, спин появляется в теории автоматически, а не вносится в нее искусственно.

Во-вторых, уравнения Дирака допускают решения, соответствующие состояниям электрона с отрицательной энергией. Такие состояния, казалось бы, бессмысленны, и поэтому их существование представляло вначале большие трудности для теории Дирака. Однако эти трудности были преодолены, для чего понадобилось сделать удивительное предположение. Оно заключается в том, что континуум состояний с отрицательной энергией предполагается заполненным электронами. Тогда в силу п р и н ц и п а з а п р е т а П а у л и, согласно которому в каждом квантовом состоянии не может находиться больше одного электрона, электроны не смогут переходить из одного состояния с отрицательной энергией в другое и такой бесконечный фон электронов, «сидящих» на уровнях с отрицательной энергией (так называемый электронный вакуум), не будет проявляться сам по себе. Проявляются дырки в этом бесконечном распределении — дираковском «море» электронов. Дырки ведут себя как частицы, обладающие положительными энергиями, массой и зарядом. Их масса равна массе электрона, а заряд только знаком отличается от знака заряда электрона. Эти частицы — их называют позитронами — представляют собой античастицы по отношению к электронам.

Можно представить себе, что из «фона» отрицательных электронных состояний с помощью электромагнитных полей «вытягивается» электрон в континуум состояний с положительной энергией. Тогда возникают реальные электрон и дырка, т. е. позитрон. Это значит, что электромагнитное поле «родило» электрон-позитронную пару. Представим себе теперь, что электрон «падает» в дырку. Освобождающаяся при этом энергия уносится фотонами. Итак, возможна аннигиляция электрон-позитронной пары в фотоны. Из законов сохране-

ния энергии и импульса следует, что при аннигиляции пары должно родиться не менее двух фотонов.

Оба эти явления — рождение пар и их аннигиляция — были сперва теоретически предсказаны Дираком. Впоследствии опыты подтвердили эти гениальные предсказания.

Рассмотренные эффекты означают, что электромагнитное поле, т. е. фотоны, столь же материально, как и электроны и позитроны. Электромагнитное поле и электроны (позитроны) выступают в квантовой электродинамике как две равноценные формы единой материи.

Можно сказать, что идея Фарадея и Максвелла о реальности поля получила удивительное подтверждение и развитие.

В основе квантовой электродинамики лежат уравнения Максвелла, описывающие классическое электромагнитное поле, и квантово-механические уравнения Дирака, описывающие релятивистский электрон.

Уравнения Дирака, как и уравнения Максвелла, имеют полевой характер и отражают вместе с последними волновую природу материи.

Замечательное предсказание Дираком возможности существования частицы, отличающейся от электрона только знаком заряда, — позитрона — привело к важному выводу о том, что уравнения Дирака описывают не только электроны, но и позитроны. Проводя аналогию между уравнениями Дирака и Максвелла, можно сказать, что уравнения Максвелла описывают электромагнитное поле, а Дирака — электрон-позитронное поле.

Понятия полей — электромагнитного и электрон-позитронного — являются основными в квантовой электродинамике. Эти понятия должны отражать не только волновые свойства материи. Как функции координат и времени электромагнитное и электрон-позитронное поля удовлетворяют определенным волновым уравнениям. Они должны также отражать корпускулярные свойства материи и математически описываться некоторыми операторами. Таким образом, мы приходим к понятиям квантовых полей, удовлетворяющих как функции координат и времени системе связанных дифференциальных уравнений Максвелла и Дирака и подчиняющихся как операторы определенным перестановочным соотношениям.

Поля и соответствующие им частицы взаимодействуют между собой. В результате этого взаимодействия одни частицы исчезают, а другие — возникают. Например, в процессе излучения возникает фотон, а в процессе образования фотоном электрон-позитронной пары фотон исчезает и появляются электрон и позитрон. Различные квантово-электродинамические процессы, как и все квантовые эффекты, характеризуются определенными вероятностями. Они определяются квадратами модулей элементов некоторой матрицы — так называемой матрицы расщепления. Чтобы найти ее (а это главная задача квантовой электродинамики), нужно решить связанную систему операторных дифференциальных уравнений квантовой электродинамики — уравнений Максвелла и Дирака. Точное общее решение этой задачи пока неизвестно. Удастся найти только приближенное решение уравнений квантовой электродинамики в рамках теории возмущений.

Однако высшие приближения теории возмущений содержат расходящиеся выражения, смысл которых был понят не сразу. Новый этап в развитии квантовой электродинамики, который можно с полным правом назвать вторым ее рождением, начался в 50-х годах нашего столетия, когда были вскрыты физические причины расходимостей в квантовой электродинамике и установлены методы их устранения.

Чтобы понять создавшееся в квантовой электродинамике положение, следует иметь в виду, что квантованные поля обладают определенными физическими свойствами даже тогда, когда числа частиц, связанных с полями, равны нулю. В этом случае говорят, что система полей находится в состоянии в а к у у м а. Последний обладает определенными физическими свойствами. Это не пустота, а «живой» вакуум, своеобразная физическая среда, обладающая целым рядом свойств.

Любой внешний заряд поляризует вакуум. Благодаря этому каждый электрон, образно выражаясь, покрывается поляризационной электрон-позитронной «шубой», которая воспринимается внешним наблюдателем как эффективное уменьшение заряда электрона. Иными словами, если e_0 — заряд «голого» электрона, находящегося в «шубе», то наблюдаемый заряд электрона $e = e_0 + \Delta e$, где Δe — изменение заряда, вызываемое поляризацией вакуума.

Но этим не исчерпывается взаимодействие вакуума с электроном. Последний все время испускает и вслед за тем поглощает фотоны, вследствие чего должна изменяться его энергия, или, что то же самое, масса электрона. Изменение массы электрона, обусловленное его взаимодействием с вакуумом, называют электромагнитной массой электрона. Если m_0 — масса «голого» гипотетического электрона, не взаимодействующего с вакуумом, то наблюдаемая масса реального электрона $m = m_0 + \Delta m$, где Δm — электромагнитная масса электрона.

Возникает задача — определить величины Δe и Δm . При ее решении мы сталкиваемся, однако, с принципиальной трудностью, заключающейся в том, что если буквально следовать теории, то для Δe и Δm получаются бесконечные выражения, имеющие вид расходящихся интегралов. Теория нуждается еще в определенной процедуре регуляризации, т. е. устранения бесконечностей из различных величин, имеющих непосредственный физический смысл. Эта процедура основана на простой физической идее перенормировки, согласно которой величины $e_0 + \Delta e$ и $m_0 + \Delta m$ должны отождествляться с наблюдаемыми конечными значениями заряда и массы электрона.

Таким образом, мы не принимаем во внимание бесконечные значения электромагнитных заряда и массы электрона. Ведь все равно в современную теорию и заряд и масса частиц вводятся чисто феноменологически! Важно то, что различных типов расходимостей, встречающихся в квантовой электродинамике, не бесконечно много, а всего два. Оба они сводятся только к расходящимся величинам Δe и Δm . Поэтому все расходимости в квантовой электродинамике могут быть одновременно устранены перенормировкой. На этом основании кван-

товую электродинамику называют перенормируемой теорией.

Идея перенормировки оказалась на редкость плодотворной и дала возможность предсказать целый ряд замечательных физических явлений и создать их количественную теорию. К числу этих явлений относится существование аномального магнитного момента электрона, радиационное смещение атомных уровней и, наконец, различные нелинейные электродинамические эффекты (например, рассеяние света светом в пустоте), в силу которых вакуум в действительности представляет собой физическую среду с некоторыми зависящими от напряженностей поля проницаемостями ϵ и μ .

Предсказания квантовой электродинамики подтверждаются в гигантском интервале энергий.

Новые эффекты, предсказываемые квантовой электродинамикой, помимо чисто физического представляют значительный интерес и с общепhilosophической точки зрения. Изучение их подтверждает известный тезис В. И. Ленина о неисчерпаемости свойств электрона, тезис, в котором выражена идея о неисчерпаемости свойств материи.

В настоящее время по образу и подобию квантовой электродинамики строится общая квантовая теория полей.

Приложение. ФОРМУЛЫ ВЕКТОРНОЙ АЛГЕБРЫ И ВЕКТОРНОГО АНАЛИЗА

Векторная алгебра

Сложение векторов

$$\mathbf{a} = i\alpha_x + j\alpha_y + k\alpha_z;$$
$$a_l = \alpha_x \cos \widehat{l x} + \alpha_y \cos \widehat{l y} + \alpha_z \cos \widehat{l z}$$

Скалярное произведение

$$\mathbf{ab} = \alpha_x b_x + \alpha_y b_y + \alpha_z b_z$$

Векторное произведение

$$[\mathbf{ab}] = -[\mathbf{ba}] = \begin{vmatrix} i & j & k \\ \alpha_x & \alpha_y & \alpha_z \\ b_x & b_y & b_z \end{vmatrix},$$

$$[\mathbf{ab}]_x = \alpha_y b_z - \alpha_z b_y, \dots,$$

$$\mathbf{a} [\mathbf{bc}] = \mathbf{b} [\mathbf{ca}] = \mathbf{c} [\mathbf{ab}] = \begin{vmatrix} \alpha_x & \alpha_y & \alpha_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix},$$

$$[\mathbf{a} [\mathbf{bc}]] = \mathbf{b} (\mathbf{ac}) - \mathbf{c} (\mathbf{ab})$$

Векторные поля

Оператор набла

$$\nabla = i \frac{\partial}{\partial x} + j \frac{\partial}{\partial y} + k \frac{\partial}{\partial z}$$

Градиент

$$\text{grad } \varphi \equiv \nabla \varphi = i \frac{\partial \varphi}{\partial x} + j \frac{\partial \varphi}{\partial y} + k \frac{\partial \varphi}{\partial z}$$

Дивергенция

$$\text{div } \mathbf{a} \equiv \nabla \mathbf{a} = \frac{\partial \alpha_x}{\partial x} + \frac{\partial \alpha_y}{\partial y} + \frac{\partial \alpha_z}{\partial z}$$

Ротор

$$\operatorname{rot} \mathbf{a} \equiv [\nabla \mathbf{a}] = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ a_x & a_y & a_z \end{vmatrix},$$

$$(\operatorname{rot} \mathbf{a})_x = \frac{\partial a_z}{\partial y} - \frac{\partial a_y}{\partial z}, \dots$$

Формулы с двумя ∇

$$\nabla \nabla \varphi = \Delta \varphi, \quad \Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2},$$

$$[\nabla [\nabla \mathbf{a}]] = \nabla (\nabla \mathbf{a}) - \Delta \mathbf{a},$$

$$[\nabla \nabla \varphi] = 0,$$

$$\nabla [\nabla \mathbf{a}] = 0$$

Формулы для комбинированных полей

$$\nabla (\varphi + \psi) = \nabla \varphi + \nabla \psi, \quad \nabla (\varphi \psi) = (\nabla \varphi) \psi + \varphi (\nabla \psi),$$

$$\nabla (\mathbf{a} + \mathbf{b}) = \nabla \mathbf{a} + \nabla \mathbf{b}, \quad [\nabla (\mathbf{a} + \mathbf{b})] = [\nabla \mathbf{a}] + [\nabla \mathbf{b}],$$

$$\nabla (\varphi \mathbf{a}) = \mathbf{a} \nabla \varphi + \varphi \nabla \mathbf{a}, \quad [\nabla (\varphi \mathbf{a})] = [\nabla \varphi \mathbf{a}] + \varphi [\nabla \mathbf{a}],$$

$$\nabla (\mathbf{a} \mathbf{b}) = (\mathbf{a} \nabla) \mathbf{b} + (\mathbf{b} \nabla) \mathbf{a} + [\mathbf{a} [\nabla \mathbf{b}]] + [\mathbf{b} [\nabla \mathbf{a}]],$$

$$\nabla [\mathbf{a} \mathbf{b}] = \mathbf{b} [\nabla \mathbf{a}] - \mathbf{a} [\nabla \mathbf{b}],$$

$$[\nabla [\mathbf{a} \mathbf{b}]] = \mathbf{a} (\nabla \mathbf{b}) - \mathbf{b} (\nabla \mathbf{a}) + (\mathbf{b} \nabla) \mathbf{a} - (\mathbf{a} \nabla) \mathbf{b}.$$

Интегральные соотношения

$$\int_V \nabla \varphi dV = \int_{\Sigma} \varphi \mathbf{n} ds,$$

$$\int_V \nabla \mathbf{a} dV = \int_{\Sigma} \mathbf{a} \mathbf{n} ds,$$

$$\int_V [\nabla \mathbf{a}] dV = \int_{\Sigma} [\mathbf{n} \mathbf{a}] ds,$$

$$\int_{\Sigma} [\nabla \mathbf{a}] \mathbf{n} ds = \int_C \mathbf{a} d\mathbf{l}$$

(объем V ограничен поверхностью Σ , \mathbf{n} — единичный вектор, нормальный к поверхности и направленный вне ее; поверхность Σ ограничена контуром C)

Литература

- Абрикосов А. А. Введение в теорию нормальных металлов. М., 1972.
Ахиезер А. И. и др. Электродинамика плазмы. М., 1974.
Ахиезер А. И., Барьяхтар В. Г., Пелетминский С. В. Спиновые волны. М., 1967.
Ахиезер А. И., Берестецкий В. Б. Квантовая электродинамика. М., 1969.
Бараш Ю. С., Гинзбург В. Л. — УФН, 1976, т. 118, вып. 3, с. 523.
Бломберген Н. Нелинейная оптика. М., 1966.
Блохинцев Д. И. Введение в квантовую механику. М., 1944.
Борн М., Вольф Э. Основы оптики. 2-е изд. М., 1973.
Владимиров В. В., Волков А. Ф., Мейлихов Е. З. Плазма полупроводников. М., 1979.
Гинзбург В. Л. Теоретическая физика и астрофизика. М., 1981.
Глаубер Р. Оптическая когерентность и статистика фотонов. — В кн.: Квантовая оптика и квантовая радиофизика. М., 1966.
Де Жен П. Сверхпроводимость металлов и сплавов. М., 1968.
Джексон Дж. Классическая электродинамика. М., 1965.
Зельдович Я. Б., Новиков И. Д. Релятивистская астрофизика. М., 1967.
Зингер Дж. Мазеры. М., 1961.
Кадомацев Б. Б. Коллективные явления в плазме. М., 1976.
Карпман В. И. Нелинейные волны в диспергирующих средах. М., 1973.
Киттель К. Физическая теория доменной структуры ферромагнетиков. — УФН, 1950, т. 151, вып. 4, с. 452—544.
Ландау Л. Д., Лифшиц Н. М. Статистическая физика. 2-е изд. М., 1964.
Ландау Л. Д., Лифшиц Е. М. Теория поля. 5-е изд. М., 1967.
Ландау Л. Д., Лифшиц Е. М. Электродинамика сплошных сред. М., 1957.
Ландсберг Г. С. Оптика. 2-е изд. М., 1947.
Лонгмайер К. Физика плазмы. М., 1966.
Мандельштам Л. И. Лекции по некоторым вопросам теории колебаний. — В кн.: Мандельштам Л. И. Полное собрание трудов/ Под ред. М. А. Леонтовича. М., 1950, т. 5, с. 428—467.
Островский Ю. И. Голография. М., 1970.
Пантел Р., Путхов Г. Основы квантовой электроники. М., 1972.
Паули В. Теория относительности. М., 1947.
Стрэттон Дж. А. Теория электромагнетизма. М., 1948.
Тамм И. Е. Основы теории электричества. М., 1966.
Тамм И. Е., Франк И. М. Излучение электрона при равномерном движении в преломляющей среде. — В кн.: Тамм И. Е. Собрание научных трудов. М., 1975, т. 1, с. 113—120.
Фейнман Р., Лейтон Р., Сэндс М. Фейнмановские лекции по физике. М., 1966, т. 8, с. 271, 1967, т. 9, с. 359.
Ферми Э. Квантовая механика (конспект лекций). М., 1965.
Ферми Э. Ядерная физика. М., 1951.
Шавлов А. Оптические мазеры. — УФН, 1961, т. 75, вып. 3, с. 569—582.
Энгель А., Штенбек М. Физика и техника электрического разряда в газах. М., 1935, т. 1, т. 2.
Яриш А. Квантовая электроника и нелинейная оптика. М., 1973.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- Аббе инвариант 395
Абберация 398
— света 123
Адрон 6
Активная среда 476
Альфеновская скорость 360
Альфеновские волны 359
Ампера закон 54
Амплитуда комплексная 290
— тока комплексная 78
Анггармонический осциллятор 416
— — во внешнем поле 417
Аппигиляция электрон-позитронной пары 487
Анод 211
Аномальная дисперсия 322
Аномальный скин-эффект 324
— магнитный момент электрона 490
Антисимметрическая волновая функция 243
Антипараллельная ориентация спинов 243
Антиферромагнетики 245
Античастица 487
Астоново темное пространство 218
- Баланс энергии в среде 106
Безстолкновительная плазма 344
Безстолкновительное затухание 348
Бетатрон 76
Биеение 328
Бинормаль 369
Био—Савара закон 58
Бирадиаль 369
Блоха закон 240, 380
— стенка 250
Бозе—Эйнштейна статистика 168
Бозон 168
Больцмана формула 158
Бора формула 453
Брегга—Вульфа условия 411
Брюстера угол 341
Быстрая магнитозвуковая волна 360
Бэте формула 453
- Вавилова—Черенкова эффект 460
Вакуум квантового поля 489
Вакуумный диод 211
Ван Циттера—Черникс теорема 307
Вектор обратной решетки 187
— Поинтинга 95
— поляризации 99
Векторы аксиальные и полярные 51
Верде постоянная 376
Вероятностная интерференция волновой функции 404
Вероятность поглощения фотона атомом 470
— разрушения электронной пары 256
— рассеяния 451
— спонтанного излучения 470
- Взаимодействие между токами в среде 221
— сильное 6
— слабое 7
— трех волн 425
— электромагнитных волн с акустическими 426
— электронов с фонами 255
Видность полос 304
Виртуальные фотоны 489
Вихревые нити 268
Власова уравнение 344
Влияние поляризации света на ионизационные потери 458
Вмороженность магнитных силовых линий 358
Внешняя рефракция 370
Возбуждение атома 216
— ионно-звуковых колебаний 351, 353
— ленгмюровских колебаний 351
Волна плоская монохроматическая 274
— частично-поляризованная 278
— электромагнитная в вакууме 92
Волновая функция 403
Волновод 282
— с двухсвязным сечением 285
Волновое уравнение 92
— — с источником 430
— поле в замкнутой полости 474
Волновой пакет 328
Волновые поверхности 387
Волны бегущие и стоячие 280
— поперечные электромагнитные 284
— собственные резонатора 286
— спиновые 378
Вольт-амперная характеристика 218
— — проводников 201
Вольты элемент 209
Восприимчивость диамагнитного газа 232
Вращение плоскости поляризации 374
Временная дисперсия 344
Время релаксации намагничивания 234
— — электрона 190
— свободного пробега электрона 190
— собственное 121
Вторичная волна 405
Вторичные электроны 211
Вынужденное рассеяние Манделштама—Бриллюэна 427
Выпрямитель 219
— полупроводниковый 183
Вырожденная плазма 344
Вырожденный электронный газ 166
Вырывание холодное 181
- Газовый лазер 477
— разряд 211

- Газонаполненные фотоэлементы 219
- Гальванический элемент 209
- Гальваномагнитные явления 236
- Гаусса теорема 14
- Гауссова система единиц 109
- Гауссово распределение 309
- Гейзенберга соотношение неопределенностей 404
- Генерация второй гармоники 417, 425
- третьей гармоники 419
- Гибридные резонансы 355
- Гидродинамика плазмы 356
- Гирации вектор 363
- Гирация плазмы в магнитном поле 375
- Гиротропная среда 363
- Гирочастота 354
- Главное сечение кристалла 367
- Главные диэлектрические проницаемости кристалла 363
- значения диэлектрической восприимчивости 162
- направления поляризуемости 162
- Глубина проникновения поля в сверхпроводник 265
- скин-слоя 323
- Голография 412
- «Гольд» электрон 489
- Гомополярная химическая связь 243
- Горячая Вселенная 475
- Граничная энергия Ферми 169
- Граничный импульс Ферми 169
- Гюйгенса принцип 404

- Два рода сверхпроводников 263
- Движение частицы в электромагнитном поле 132
- Двойное лучепреломление 368
- Двойной электрический слой 209
- Двухтемпературная гидродинамика плазмы 359
- плазма 346
- Дебая радиус 208
- температура 194
- Де Бройля волны 403
- Декремент затухания 321
- Детектирование 440
- Джозефсона эффект 271
- Джоульса теплота 47
- Диаграмма направленности 435
- Диамagnetизм 220
- Ландау 232
- сверхпроводников 256
- Диамagnetная восприимчивость 236
- Динамические голограммы 414
- Диполь электрический 17
- Дипольное излучение 433, 455
- приближение 456
- Дирака биспинор 487
- уравнение 487
- Дираковское «море» электронов 487
- Дисперсионное уравнение 355
- Дисперсия плазменных колебаний 346
- проницаемости 317
- Дифракционная решетка 409
- Дифракция Фраунгофера 407
- на бесконечно длинной щели 408
- двух щелях 410
- нейтронов в кристалле 411
- Дифференциальное сечение рассеяния 450
- Диэлектрическая проницаемость в газодинамических трубах 319
- для двухосных и одноосных кристаллов 163
- плазмы в магнитном поле 355
- плотных газов 159
- полярных молекул 159
- Диэлектрический шар в электрическом поле 37
- Длина когерентности 305
- пробега 215, 453
- фотона в веществе 468
- электрона 195, 196
- пути оптическая 296
- собственная 119
- фокусировки 421
- Добротность резонатора 286
- Домены 249
- Друде формула 184

- Емкость проводника 25
- Естественная оптическая активность 376
- ширина спектральной линии 447

- Зависимость спонтанной намагниченности от температуры 240
- Закон всемирного тяготения 5
- отсутствия магнитных зарядов 89
- сохранения электрического заряда 9
- электролиза Фарадея 206
- электромагнитной индукции Фарадея 69
- Западающий потенциал 431
- Заряд пробный 10
- электрический 8
- Затухание ионного звука 351
- Ландау 348
- ленгмюровских колебаний 350
- Зона проводимости 176, 238
- Зонная решетка Френеля 413
- Зоны энергетические 175

- Идеальный оптический прибор 388
- Излучение в кулоновском поле 454
- волн антенной 439
- Изменение хода времени 119
- — — в случае распада элементарной частицы 122
- Изоляторы 20
- Импеданс 78
- Импульс, передаваемый при столкновении 452
- поля 95
- частицы 127
- Инварь варианты поля 126
- Инвария 393
- Инверсия заселенность уровней 476
- Индуктивность 64
- соленоида 65
- Индукционные генераторы 73
- Индукция электростатическая 23
- Индукционное излучение 471
- Инкремент нарастания колебаний 351
- Интеграл действия 144
- столкновений 188
- Интенсивности закон в геометрической оптике 388
- Интенсивность излучения 435
- — релятивистской частицы 441
- — черного тела 475
- Интервал 121
- в неинерциальной системе 143
- Интерференция волн 289
- интенсивности 308
- при отражении 295
- Интерферометр Майкельсона 299
- Ионизационная камера 219
- Ионизационные потери 451
- Ионизация 215
- нейтрального газа 210
- Ионизованная плазма 342
- Искривление энергетических зон 183

- Катод 211
- Катодное темное пространство 218
- Катодный слой 218
- Каустики 388
- Квазиимпульс 187
- Квазилинейная теория плазмы 352
- Квадратичная нелинейность 419
- Квадрупольное излучение 438
- Кваптование момента 226
- магнитного тока, пронизывающего сверхпроводник 258
- Квантованное поле 488

- Квантово-механическая теория возмущений 489
 Керра постоянная 372
 — эффект 371
 — — магнитный 376
 Кинетическое уравнение в магнитном поле 237
 — — электронов 188
 Кирхгофа правило 47
 Классический радиус электрона 437
 Клаузиуса—Моссотти формула 161
 К-оболочка в атоме 461
 Ковариантность 138
 Когерентное излучение 477
 — рассеяние 465
 Когерентность полная 305
 Колебания собственные 82
 Колебательный контур 81
 Комбинационное рассеяние на спиновой волне 428
 Комплексная фазовая скорость 320
 Комптона формула 463
 Комптоновская длина волны электрона 463
 Комптон-эффект 463
 Конденсатор электрический 26
 Коническая рефракция 369
 Константа неоднородного обменного взаимодействия 247
 Контактная разность потенциалов 209
 Контрастность картины 304
 Конформные отображения 383
 Корреляционные функции напряжения и тока 85
 Коттона—Мутона эффект 372
 Коэрцитивная сила 241
 Коэффициент взаимной индукции 65
 — затухания 82, 321
 — ионизации 216
 — модуляции 440
 — отражения 341
 — самоиндукции 65
 Красное смещение 149
 Кривая намагничивания ферромагнетиков 240
 Кристоффеля символы 146
 Критическая температура 254
 Критическое магнитное поле 257
 Круговая поляризация волны 277
 Кубическая нелинейность 419
 Кулон 9
 Кулона закон 8
 Кюри закон 229
 — температура 240
 Кюри—Вейса закон 240
- Лагранжа—Эйлера уравнение 154
 Лагранжево описание жидкости 256
 Лазер 475
 — рубиновый 480
 Ламинарная структура сверхпроводника 267
 Ланде множители 230
 Ланжевена формула 229
 — функция 229
 Лапласа уравнение 16
 Лармора прецессия 233
 — теорема 233
 — частота 233
 Ларморов радиус 52
 Лиенар—Вихерта потенциал 432
 Линза тонкая 397
 — двояковыпуклая 398
 Линзы формула 399
 Линия силовая 10
 Лондонов уравнение 262
 Лоренца преобразование 117
 — — для электромагнитного поля 125
 — сила 50
 — условие 429
 — формула для локального поля 160
 Лоренца—Лорентца формула 161
 Лоренцева форма спектральной линии 302
 Люмен 389
- Магнетик в магнитном поле 223
 Магнетон Бора 226
 Магнитная анизотропия 248
 — восприимчивость вещества 220
 — — парамагнитного газа 225
 — — ферромагнетика 382
 — индукция в эллипсоиде 225
 — поверхность 60
 — проницаемость 220
 — энергия ферромагнетика 251
 Магнитное поле 51
 Магнитной анизотропии постоянная 380
 Магнитные подрешетки 245
 — силовые линии 51
 — электроны 242
 Магнитный заряд 224
 — момент тока 59
 — поток 66
 — — в соленоиде 222
 Магнитоактивная плазма 355
 Магнитогидродинамические волны 359
 Магнитодвижущая сила 222
Магнитодипольное излучение 438
 Магнитосопротивление 222, 239
 Магнитоспиновый резонанс 383
 Магнитостатический потенциал 224
 Магноны 379
 Мазер 475
 — твердотельный 479
 Майкельсона опыты 298
 Максвелла первое уравнение 88
 — теорема 392
 — уравнение для материальных сред 103
 — — , четырехмерная форма 138
 Максимальный угол отклонения 451
 Манделштама—Бриллюэна эффект 426
 Масса эффективная электрона 197
 Матрица рассеяния 488
 — поляризационная 278
 Матричный элемент перехода 271
 Медленная магнитозвуковая волна 360
Мейснера эффект 258
 Мениск 398
 Метод коррекции волновых фронтов 413
 — — электрического изображения 26
 Минимальность действия 402
 Минимизация энергии ферромагнетика 251
 Минковского псевдоевклидово пространство 120
 Мировая точка 120
 Мнимая часть диэлектрической проницаемости 335
 Многократное рассеяние 450
 М-оболочка в атоме 467
 Модуляционная неустойчивость 423
 Молекулы полярные и неполярные 157
 Молекулярное поле Вейса 245
 Мопертюи принцип 401
 Мягкие фотоны 464
- Найквиста формула 87
 Намагниченность 100
 Направление легчайшего намагничивания 241
 Нееля температура 245
 Нелинейное удвоение частоты
 Необыкновенная волна 367
 Неравновесная среда 476
 — функция распределения 351
 Нерелятивистское сечение рассеяния 456
 Несущая волна 328
 Низкочастотные колебания плазмы 347
 Николья призма 368
 N-образная вольт-амперная характеристика 201
 Нормальная дисперсия 322
 — заселенность уровней 478
 Нормальные металлы 251
 Ньютона кольца 297
- Обменная энергия 244**

- Обменное взаимодействие 443
 Обменный интеграл 244
 Образование пары в поле ядра 468
 — электрон-позитронных пар 467
 Общая квантовая теория полей 490
 Объемный заряд 212
 Обыкновенная волна 367
 Окно прозрачности для жестких фотонов 470
 Ома закон для магнитной цепи 222
 Онзагера формула 160
 Оперсжающий потенциал 432
 Опорная волна 412
 Определения основных единиц СИ 115
 Оптическая длина пути 389
 — изомерия 376
 — — кристалла 367
 — — первого рода 369
 — сила 393
 Оптическое изображение 367
 Орбитальный магнитный момент 226
 Отклонение света в гравитационном поле 153
 Относительная магнитная проницаемость 220
 Отображение параксиальными лучами 395
- Параллельная ориентация спинов 243
 Парамагнетизм 220
 — выроденного электронного газа 232
 — паулевский 232
 Парсевала формула 291
 Паули принцип 168, 487
 Перенормировка массы 489
 — заряда 489
 — теории 490
 Перестановочные соотношения 488
 Петля гистерезиса 241
 Плазма в магнитном поле 354
 — с током 353
 Плазменная частота 343
 Планка постоянная 168, 403
 — формула 472
 Плотность внутренней энергии среды 107
 — излучения спектральная 302
 — свободной энергии полей в среде 108
 — уршей электронов 237
 — токопроводимости 99
 — электрической энергии 28
 — электромагнитной энергии в волне 275
 Поверхностная волна 339
 — энергия ферромагнитного замена 251
 Поглощение фотонов в веществе 468
 — энергии в среде 335
 Подвижность иона 206
 Показатель преломления 332
 — — векторный 364
 Поле движущихся зарядов 429
 — заряженного шара 15
 — заряженной плоскости 15
 — соленоида 56
 — термодинамически равновесно 301
 Полная энергия излучения 415
 Полное сечение эффекта Комптона 466
 Полосы интерференционные 294
 — равного наклона 297
 Полупроводники диричные 178
 — примесные 117
 — электронные 178
 Поля усредненные 97
 — равномерно движущегося электрического заряда 126
 Поляризатор 369
 Поляризация вакуума 489
 — диэлектриков 33
 — левая 374
 — линейная 276
 — правая 374
 — электрическая 277
 — электромагнитной волны 276
 Поляризованность нелинейная 417
 — спонтанная 164
 Поляризуемость 34
 — молекулы 156
- Поперечные гальваномагнитные явления 239
 Потенциал векторный 57
 — зажигания 218
 — квадрупольного момента 19
 — спаривания 261
 — химический электрона 171
 — экранированный 208
 — электростатический 12
 Потенциальная функция тока 63
 Поток электрического поля 12
 — элементарный в сверхпроводящем кольце 266
 Преобразование аффинное 393
 — импульса и энергии 135
 — проективное 394
 Преломление силовых линий 33
 Приближение параксиальное 395
 Призма 340
 Принцип вариационный 145
 — равного оптического пути 392
 — суперпозиции 10
 — эквивалентности 140
 Прицельный параметр 450
 Проводимость удельная электрическая 44
 Проводник 20
 Проводники собственные 177
 Проводящая жидкость 356
 Продольный гальваномагнитный эффект 239
 Проекция орбитального магнитного момента 220
 — спина на направление магнитного поля 217
 Проницаемость диэлектрическая 31
 Пространство изображения 392
 Прохождение заряженных частиц через вещество 449
 Процедура регуляризации 490
 Пуассона уравнение 16
 Пучки когерентные 291
 — некогерентные 291
 Пучок в плазме 353
- Равновесное распределение фотонов 472
 Радиационное смещение уровней 490
 Радиус гравитационный 149
 — экранировки 208
 Разложение по плоским волнам 326
 — света 339
 Разность потенциалов контактная 182
 Разряд темный 418
 — тлеющий 418
 Рассеяние в кулоновском поле 450
 — на малые углы 449
 — света на свете 490
 Расходимости в квантовой электродинамике 489
 Резерфорда формула 450
 Резонанс парамагнитный 382
 — ферромагнитный 381
 Резонансное условие 348
 Резонаторы 285
 Реликтовое излучение 475
 Ридберга постоянная 417
 Ричардсона—Дешмана формула 180
 Рудгерса формула 260
- Самовозбуждения эффект 478
 Самомодуляция 422
 Самофокусировка и самоканалирование 420
 Свободная энергия поля в среде 333
 Сверхпроводимость 254
 Сверхпроводящая фаза 255
 Сверхпроводящее состояние 254
 Сверхтекучая жидкость 255
 Свет неполяризованный 278
 Связь между мировым и собственным временем 149
 — — энергией и импульсом для безмассовой частицы 130

— обратная 83
 Сегнето- и пирозлектрики 165
 Сечение рассеяния в ультрарелятивистской области 457
 — электрона молекулой 215
 — тормозного излучения 455
 Сила лучистого трения 443
 — осциллятора 319
 — тока 42
 Силы пондеромоторные 40
 Сильные и слабые электролиты 203
 Симметричная волновая функция 243
 Симметрия нелинейной восприимчивости 419
 Синхронизация часов 119
 Система единиц СИ 112
 Скалярное произведение 4-векторов 134
 Скачок температуры в ферромагнетике 246
 — электронной теплоемкости 254
 Скорость дрейфа 236
 — фазовая 315
 Слабые ферромагнетики 240
 Сложение скоростей 122
 Смещение электрическое 31
 Снелиуса закон 337
 S-образная вольт-амперная характеристика 201
 Собственный магнитный момент атома 225
 Соотношение тензоров 136
 Соотношение между величинами в СИ и ГС 115
 — неопределенностей в оптике 330
 — для волновых пакетов 330
 Сопrotивление переменного тока 323
 — проводника в магнитном поле 236
 Состояние промежуточное сверхпроводников 267
 — с минимумом энергии в ферромагнетике 244
 — отрицательной энергией 487
 Спаренные электроны 256
 Спектр излучения 457
 — энергетический кристалла 175
 — энергии дискретный 403
 — непрерывный 404
 Спектральная плотность излучения 443, 456
 Спектральное распределение излучения 455
 Спин электрона 167, 487
 Спиновая корреляция 244
 Спиновый магнитный момент 226
 Спиноры 137
 Средний ионизационный потенциал 453
 — квадрат угла многократного рассеяния 451
 Степень диссоциации 205
 — когерентности 303
 — высших порядков 310
 — монохроматичности 292
 Стокса параметры 278
 — формула 112
 Суммарный магнитный момент 226
 — орбитальный момент 227
 Сферические волны 431
 Счетчик фотонных совпадений 311

Таблица единиц
 Тамма—Франка формула 462
 Температура вырождения 232
 — сверхпроводящего перехода 255
 — электронная 200
 Тензор диэлектрической проницаемости анизотропной среды 362
 — квадрупольного момента 19
 — магнитной проницаемости 220
 — метрический 143
 — — в слабом гравитационном поле 147
 — размагничивающих коэффициентов 224
 — электромагнитного поля 137
 Тепловая скорость электронов 215
 Тепловые фононы 427
 Теплоемкость электронная 174
 Термодинамика сверхпроводников 259
 Типы взаимодействий 5

Ток насыщения 214
 — смещения 91
 — трехфазный 74
 Томсона формула 464
 Томсоновское сечение 464
 Тонкой структуры постоянная 457
 Тормозное излучение 451
 Точка стигматическая 392
 Точки сопряженные 392
 Транзистор 184
 Трехуровневая система 479

Угловое распределение фотонов 466
 Угол полного внутреннего отражения 339
 Уравнение движения в гравитационном поле 145
 — — осциллятора 318
 — непрерывности 43
 — определения формы луча 390
 Ускорение в релятивистской динамике 132
 Условие эргодичности 300
 Условия возникновения самостоятельного разряда 217
 — граничные для электрического поля 32
 — инверсии 480

Фаза волны 321
 Фазовая скорость 321
 Фазовый переход второго рода 254
 Фазы ферро- и параэлектрическая 165
 Фарад 25
 Фарадея эффект 376
 — постоянная 206
 Ферма принцип 389
 Ферми распределение 170
 — эффект 458
 Ферми—Дирака статистика 168
 Ферриты 245
 Ферромагнетики 220
 Флуктуации электромагнитные 83
 Фокальная плоскость пространства изображения 394
 — — предмета 394
 Фокус 388
 Фононы 193, 379
 Форма спектральной линии 436
 Фотоэффект 180
 Френеля дифракция 405
 — интеграл 407
 — уравнение 364
 — формула 341
 — эллипсоид 366
 Фронт волны 322
 Функции корреляционные высших порядков 309
 Функция корреляционного поля 299
 — нормированная корреляционная 304
 — распределения вырожденного электронного газа 231
 — — фононов 194
 — — электронов в электрическом поле 191

Характерная длина нелинейного взаимодействия 425
 Холла эффект 236
 «Холодная» плазма 346
 Хэнбери—Брауна и Твисса эффект 312

Центр инверсии 393

Частицы элементарные 5
 Частота циклотронная 53
 — — в релятивистском случае 133
 Черенковские потери энергии 462
 Черенковский счетчик заряженных частиц 463
 Черное излучение 473
 4-векторы и 4-тензоры 133
 Четырехмерное уравнение движения частицы в поле 139

Четырехмерный волновой вектор 140

Шварцшильда метрика 149

Ширина домена 251

— эффективная спектральной линии 302

Шкала волн 275

Шредингера уравнение 403

Э.д.с. взаимной индукции 75

Э.д.с. самоиндукции 75

Эйконал 387

Эйлера описание жидкости 356

Эйнштейна коэффициент 471

— соотношение 181

— уравнение гравитации 148

Эйнштейна—де-Гааза опыт 242

Электрическое сопротивление 46

Электродный потенциал 209

Электродвижущая сила 46

Электромагнитно-спиновые волны 384

Электролиз 205

Электромагнитные ливни 468

Электрон-позитронное поле 488

Электронная пара 256

Электроны проводимости 176

Электропроводность сильного электролита 208

Электрострикция 420

Электрофоретическая сила 207

Эмиссия термоэлектронная 178

Энергия диполя во внешнем поле 18

— ионизации 216

— магнитной анизотропии 248

— поверхностная сверхпроводника 268

— покоя 130

— поля в диэлектрике 38

Эффективное сечение фотоэффекта 467

Юнга опыт 292

ОГЛАВЛЕНИЕ

	ПРЕДИСЛОВИЕ	3
	Введение. Место электромагнетизма в современной физической картине мира	5
I. ОСНОВЫ ЭЛЕКТРОМАГНЕТИЗМА И ФИЗИКИ ПРОСТРАНСТВА — ВРЕМЕНИ	Глава 1. Электростатическое поле в вакууме	8
	1.1. Закон Кулона	8
	1.2. Теорема Гаусса и уравнение Пуассона	13
	1.3. Электрические диполь и квадруполь	17
	1.4. Проводник в электростатическом поле	20
	1.5. Электростатическая индукция	22
	1.6. Электрическая емкость проводника	25
	1.7. Энергия электрического поля	27
	Глава 2. Электростатика диэлектриков	31
	2.1. Электрическая индукция и поляризация	31
	2.2. Диэлектрик в однородном поле	36
	2.3. Энергия поля в диэлектрике	38
	Глава 3. Постоянный электрический ток	41
	3.1. Плотность электрического тока	41
	3.2. Закон Ома	43
	3.3. Линейная электрическая цепь	45
	3.4. Джоулева теплота	47
	Глава 4. Магнитное поле в вакууме	50
	4.1. Сила Лоренца	50
	4.2. Законы Ампера и Био—Савара	53
	4.3. Магнитный момент тока	59
	4.4. Взаимодействие токов	60
4.5. Индуктивность	64	
Глава 5. Электромагнитная индукция и переменные токи	68	
5.1. Закон Фарадея	68	
5.2. Относительность электрического и магнитного полей	70	
5.3. Индукционные генераторы тока	73	
5.4. Э. д. с. взаимной индукции и самоиндукции	74	
5.5. Бетатрон	75	
5.6. Индуктивность и емкость в цепи переменного тока	77	

5.7. Колебательный контур	81
5.8. Электромагнитные флуктуации	84
Глава 6. Законы электромагнетизма	87
6.1. Уравнения Максвелла	87
6.2. Свободное электромагнитное поле в вакууме	91
6.3. Усредненные микроскопические поля	96
6.4. Уравнения Максвелла для макроскопических полей в веществе	101
6.5. Баланс энергии в среде при наличии электромагнитного поля	106
6.6. Системы электромагнитных единиц	109
Глава 7. Физика пространства—времени	117
7.1. Преобразования Лоренца	117
7.2. Собственное время	121
7.3. Сложение скоростей	122
7.4. Преобразования Лоренца для электромагнитного поля	124
7.5. Импульс и энергия частицы	127
7.6. Движение заряженной частицы в электромагнитном поле	131
7.7. 4-векторы и 4-тензоры	133
7.8. Эффект Доплера	139
7.9. Принцип эквивалентности	140
7.10. Связь метрики пространства—времени с материей	143
7.11. Красное смещение	149
7.12. Отклонение луча света в гравитационном поле	152
Глава 8. Диэлектрики	156
8.1. Полярные и неполярные молекулы	156
8.2. Диэлектрическая проницаемость плотных газов и жидкостей	159
8.3. Диэлектрическая проницаемость кристаллов	161
8.4. Кристаллы со спонтанной поляризованностью	164
Глава 9. Металлы и полупроводники	166
9.1. Вырожденный электронный газ	166
9.2. Распределение Ферми	170
9.3. Металлы, диэлектрики и полупроводники	174
9.4. Термоэлектронная эмиссия	178
9.5. Контактная разность потенциалов	181
9.6. Формула Друде	184
9.7. Кинетическое уравнение для электронов	186
9.8. Электропроводность металлов	192
9.9. Вольт-амперные характеристики полупроводников	196
Глава 10. Электролиты	203
10.1. Сильные и слабые электролиты	203
10.2. Электролиз	205
10.3. Электропроводность растворов	206
10.4. Гальванические элементы	209

II. ЭЛЕКТРИЧЕСКИЕ И МАГНИТНЫЕ СВОЙСТВА ВЕЩЕСТВА

Глава 11. Газовый разряд	210
11.1. Анодный ток	210
11.2. Температура заряженных частиц	214
11.3. Самостоятельный разряд	215
Глава 12. Магнетики	219
12.1. Магнетик во внешнем поле	219
12.2. Магнитная восприимчивость парамагнитного газа	225
12.3. Парамагнетизм вырожденного электронного газа	231
12.4. Восприимчивость диамагнитного газа	232
12.5. Гальваномагнитные явления в металлах	236
12.6. Характерные особенности ферромагнетиков	239
12.7. Природа ферромагнетизма	242
12.8. Обменная энергия и энергия магнитной анизотропии	246
12.9. Доменная структура ферромагнетиков	249
Глава 13. Сверхпроводники	254
13.1. Явление сверхпроводимости	254
13.2. Диамагнетизм сверхпроводников	256
13.3. Термодинамика сверхпроводников	259
13.4. Глубина проникновения магнитного поля	260
13.5. Два рода сверхпроводников	263
13.6. Промежуточное состояние	266
13.7. Контакт между сверхпроводниками	269
III. ЭЛЕКТРОМАГНИТНЫЕ ВОЛНЫ И ОПТИКА	
Глава 14. Электромагнитные волны в вакууме	274
14.1. Плоская монохроматическая волна	274
14.2. Поляризация электромагнитной волны	276
14.3. Частично поляризованная волна	277
14.4. Волны в пространстве между металлическими поверхностями	279
14.5. Волноводы	281
14.6. Резонаторы	285
Глава 15. Интерференция и когерентность	288
15.1. Интенсивность суперпозиции монохроматических волн	288
15.2. Некогерентные и когерентные пучки	291
15.3. Опыт Юнга	292
15.4. Двухлучевая интерференция при отражении	295
15.5. Опыт Майкельсона	297
15.6. Корреляционная функция поля	299
15.7. Степень когерентности	303
15.8. Интерференция интенсивностей	308
Глава 16. Электромагнитные волны в веществе	314
16.1. Плоские волны в однородном изотропном диэлектрике	314
16.2. Дисперсия проницаемостей	316
16.3. Комплексная фазовая скорость	320
16.4. Скин-эффект	322
16.5. Фазовая и групповая скорости	325

16.6. Энергия поля в диспергирующих средах	333
16.7. Отражение и преломление волн	337
Глава 17. Колебания плазмы	342
17.1. Ленгмюровские колебания	342
17.2. Диэлектрическая проницаемость электронной плазмы	343
17.3. Ионный звук	346
17.4. Резонансное взаимодействие волн и частиц и затухание Ландау	348
17.5. Пучковая неустойчивость	351
17.6. Колебания плазмы в магнитном поле	353
17.7. Магнитогидродинамическое течение	356
17.8. Магнитогидродинамические волны	358
Глава 18. Кристаллооптика	361
18.1. Тензор диэлектрической проницаемости	361
18.2. Фазовая скорость и лучевой вектор	363
18.3. Одноосные и двухосные кристаллы	366
18.4. Искусственная анизотропия	370
Глава 19. Колебания гиротропных сред	373
19.1. Вектор гирации	373
19.2. Естественная и искусственная гиротропия	374
19.3. Спиновые волны и ферромагнитный резонанс	377
19.4. Связанные электромагнитно-спиновые волны	382
Глава 20. Геометрическая оптика	385
20.1. Уравнение эйконала	385
20.2. Интенсивность света	387
20.3. Принцип Ферма	389
20.4. Теория оптических отображений	391
20.5. Параксиальная оптика	394
Глава 21. Дифракция	399
21.1. Разрешающая сила оптических приборов	399
21.2. Оптико-механическая аналогия	401
21.3. Принцип Гюйгенса и дифракция Френеля	404
21.4. Дифракция Фраунгофера	407
21.5. Голография	412
Глава 22. Нелинейная оптика	416
22.1. Нелинейные поляризованности и восприимчивости	416
22.2. Самофокусировка и самомодуляция светового пучка	420
22.3. Нелинейное взаимодействие электромагнитных волн	424
22.4. Эффект Манделъштама—Бриллюэна	426
Глава 23. Излучение электромагнитных волн	429
23.1. Запаздывающие потенциалы	429
23.2. Потенциалы Лиенара—Вихерта	433
23.3. Дипольное излучение	432

**IV. ВЗАИМОДЕЙСТВИЕ
ЭЛЕКТРОМАГНИТНЫХ
ВОЛН С ЧАСТИЦАМИ
И ВОЛНАМИ**

23.4. Антенна	438
23.5. Излучение релятивистского заряда	441
23.6. Сила лучистого трения	443
23.7. Ширина спектральной линии	445
Глава 24. Взаимодействие заряженных частиц и излучения с веществом	449
24.1. Рассеяние частиц в кулоновском поле ядер	449
24.2. Ионизационные потери энергии	451
24.3. Тормозное излучение	454
24.4. Влияние поляризации среды на ионизационные потери	458
24.5. Излучение Вавилова—Черенкова	460
24.6. Рассеяние фотона электроном	463
24.7. Фотоэффект и образование электрон-позитронных пар	467
24.8. Излучение и поглощение фотонов атомами	470
24.9. Микроволновые и оптические мазеры	475
Развитие учения об электромагнетизме	482
Приложение. Формулы векторной алгебры и векторного анализа	491
Литература	493
Предметный указатель	494

ЗАКЛЮЧЕНИЕ

**Александр Ильич Ахиезер,
Илья Александрович Ахиезер**

ЭЛЕКТРОМАГНЕТИЗМ И ЭЛЕКТРОМАГНИТНЫЕ ВОЛНЫ

Зав. редакцией литературы по физике и математике Е. С. Гридасова. Редактор С. А. Крылов. Младшие редакторы С. А. Доровских, Н. П. Майкова. Художник В. Н. Хомяков. Художественный редактор В. И. Пономаренко. Технический редактор Н. В. Яшукова. Корректор Г. И. Кострикова.

ИБ № 3924

Изд. № ФМ-756. Сдано в набор 24.04.84. Подп. в печать 30.01.85. Формат 60×90/16. Бум. тип. № 2. Гарнитура литературная. Печать высокая. Объем 31,5 усл. печ. л.+0,25 усл. печ. л. форзац. 31,75 усл. кр.-отт. 32,41 уч.-изд. л.+0,43 уч.-изд. л. форзац. Тираж 8000 экз. Зак. № 402. Цена 1 р. 40 к.

Издательство «Высшая школа», 101430, Москва, ГСП-4, Неглинная ул., д. 29/14.

Ярославский полиграфкомбинат Союзполиграфпрома при Государственном комитете СССР по делам издательств, полиграфии и книжной торговли. 150014, Ярославль, ул. Свободы, 97.

1р. 40к.