

# ВОЛНОВАЯ ФИЗИКА

Курс лекций для ФМШ

Часть 1. ОПТИКА И ВОЛНЫ

1. ПЕРЕМЕННЫЙ ТОК

А. П. Ершов

12 февраля 2007 г.

В последней части курса нам осталось понять, как устроены окружающие нас предметы. Мы уже знаем, что все сделано из атомов, кроме света, сделанного из фотонов. Но что такое атомы, пока не ясно.

Когда-то существование атомов нужно было доказывать. Сейчас каждый слышит про атомы почти с рождения. Это словно бы и неплохо, но все имеет свои теневые стороны. Привыкнув к идее атома, труднее сформулировать для себя вопрос: почему атомы именно такие? В начале прошлого семестра мы уже отмечали загадку атомного масштаба  $1 \text{ \AA}$ , или  $10^{-8}$  см. Чем определяется этот размер? Почему бы, например, электрону не летать раз в сто ближе к ядру (или вообще во сколько угодно раз ближе)?

В англоязычных странах популярен отрывок из «Рапсодии» Дж. Свифта:

So, naturalists observe, a flea  
Has smaller fleas that on him prey;  
And these have smaller still to bite 'em;  
And so proceed *ad infinitum*.

т.е. на блохах живут более мелкие блошки, на тех свои паразиты, еще меньше, и так далее до бесконечности. Цитируют эти стихи обыкновенно в глубокомысленном тоне и без комментариев, как бы подразумевая, что умному достаточно. Должно быть, тут отражено восхищение сложностью и неисчерпаемостью мироздания. Эта-то идея бесконечного ряда матрешек и оказалась неверна<sup>1</sup>. Действительно, блохи страдают от паразитов: клещей, червей, бактерий. Современный микроскоп показывает душераздирающие картины, например, кишечник блохи, паразитирующей на грызунах, полностью забитый палочками чумы<sup>2</sup>. «Червь, терзающий блоху, сам может быть поражен клещом, в свою очередь зараженным жгутиковыми, в клетках которых живет болезнетворный вирус». Но это и все: никакой организм не может быть меньше, скажем, ста атомных размеров. Прогрессия масштабов быстро доходит до предела, раз всякий паразит значительно меньше хозяина. Всего есть шесть уровней, считая человека, а вовсе не бесконечность<sup>3</sup>.

Если подумать, то это выходит гораздо интереснее, чем примитивная («дурная») бесконечная иерархия уровней. Мы увидим, что с уменьшением масштаба возникает совершенно новая физика, которую мудрецы прошлого и вообразить не могли.

Это – волновая физика, которой мы и займемся. Атомы такие, как они есть, потому, что они построены из определенных типов волн и сами тоже являются волнами. По сравнению со всем, что у нас было раньше, это – страна чудес. Но начнем, как всегда, с самых простых вещей.

---

<sup>1</sup>На русском языке близким эквивалентом культурной ситуации является стихотворение Пушкина «Движенья нет, сказал мудрец брадатый...».

<sup>2</sup>Rotshild M., Shlein Y., Ito S. A colour atlas of insect tissues via the flea. Science Book, N.Y., 1986. См. резюме: В мире науки (Scientific American), 1987, №1.

<sup>3</sup>Свифт вообще живо интересовался масштабным фактором, вспомним его лилипутов и великанов. Ряд интересных подробностей см. в: Н. Винер. Я – математик. РХД, Москва – Ижевск, 2001, стр. 84.

# Глава 1

## ПЕРЕМЕННЫЙ ТОК

Применение электричества начиналось с постоянного тока. Однако основа нынешней электротехники – переменный ток. Особенно важен случай периодического синусоидального тока (на котором работают бытовые приборы и электрические машины в промышленности<sup>1</sup>). Генераторы переменного тока проще и удобнее: по существу, это вращающиеся в магнитном поле контуры, или рамки. От генераторов через систему передающих устройств электричество доставляется потребителям в виде синусоидального напряжения в бытовой розетке:  $U = U_0 \cos(\omega t)$ . В России частота  $\omega = 2\pi \cdot 50 \cong 314$  Гц.

Переменный ток, в особенности синусоидальный, очень подходит для начала знакомства с волновой физикой. В то же время эта тема еще электродинамическая, благодаря чему переход будет плавным.

### 1.1 Цепи переменного тока. Импеданс

Кроме резисторов (сопротивлений), в цепях переменного тока используются конденсаторы и индуктивности. Конденсатор на первый взгляд – это просто разрыв цепи; однако переменный ток он пропускает, в виде тока смещения. Разберем на примере, как это получается.

Возьмем  $RC$ -цепь, на которую подается переменное напряжение  $U(t)$  (рис. 1.1). Какой пойдет ток? Сумма напряжений  $RI + q/C = U$  или

$$RI + \frac{1}{C} \int I dt = U .$$

Дифференцируем и получаем уравнение

$$R \frac{dI}{dt} + \frac{I}{C} = \frac{dU}{dt} .$$

Для общей зависимости  $U(t)$  решение сложное и маловразумительное. Рассмотрим частный случай, чтобы понять суть. Пусть  $U(t) = 0$  при  $t < 0$ ,  $U(t) = U_0$  при

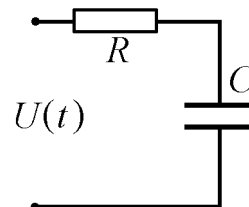


Рис. 1.1.

<sup>1</sup>Чаще под переменным током и понимается не любой непостоянный, а именно синусоидальный.

$t > 0$  (скачок напряжения). Тогда при  $t > 0$  правая часть  $dU/dt = 0$ . Решаем:  $dI/I = -dt/RC$ ,  $I = I_0 \cdot \exp(-t/RC)$ . Определяем  $I_0$ : в начальный момент  $C$  еще не успел зарядиться; тогда на нем нет напряжения и  $I_0 = U_0/R$ . Окончательно получаем ток и напряжение на конденсаторе:

$$I = (U_0/R) \cdot \exp\left(-\frac{t}{RC}\right), \quad U_C = U_0 - RI = U_0 \cdot \left(1 - \exp\left(-\frac{t}{RC}\right)\right).$$

Характерное время спада тока (или заряда конденсатора)  $\tau = RC$ . Если входное напряжение  $U(t)$  меняется за более длительное время, то можно считать, что  $C$  успевает заряжаться и на нем будет практически  $U(t)$ . Конденсатор тогда приблизительно соответствует разрыву в цепи. Если же  $U$  меняется быстро в масштабе  $RC$ , то напряжение на емкости малó по сравнению с  $U$  (почти все падает на сопротивлении; конденсатор работает как «закоротка»).

Перейдем к наиболее важному случаю синусоидального напряжения:  $U = U_0 \cos(\omega t)$ . Тогда

$$R \cdot \frac{dI}{dt} + \frac{I}{C} = -U_0 \omega \cdot \sin(\omega t).$$

Явно должно быть колебательное решение, но ни  $\sin(\omega t)$ , ни  $\cos(\omega t)$  не годятся. Пробуем  $I = I_0 \cos(\omega t + \varphi)$ , где  $\varphi$  – пока неизвестная постоянная величина, называемая **сдвигом фазы**. Подставляем:

$$\begin{aligned} -RI_0 \omega \sin(\omega t + \varphi) + (I_0/C) \cos(\omega t + \varphi) &= -U_0 \omega \sin((\omega t + \varphi) - \varphi) = \\ -U_0 \omega \cos \varphi \sin(\omega t + \varphi) + U_0 \omega \sin \varphi \cos(\omega t + \varphi). \end{aligned}$$

Переменные  $\sin(\omega t + \varphi)$  и  $\cos(\omega t + \varphi)$  должны уравниваться по отдельности. Отсюда

$$RI_0 = U_0 \cos \varphi, \quad I_0/\omega C = U_0 \sin \varphi.$$

Получаем  $\operatorname{tg} \varphi = 1/\omega RC$ ,  $I_0 = U_0/\sqrt{R^2 + (1/\omega C)^2}$ . Решение:

$$I = \frac{U_0}{\sqrt{R^2 + (1/\omega C)^2}} \cos(\omega t + \varphi), \quad \text{где } \operatorname{tg} \varphi = \frac{1}{\omega RC}.$$

Для наглядности полезны **векторные диаграммы**. Изобразим ток  $I_0$  горизонтальным вектором. Переменный ток  $I_0 \cos(\omega t)$  получим, закрутив исходный вектор против часовой стрелки с угловой скоростью  $\omega$ : горизонтальная проекция вектора именно и даст нужную величину. Напряжение на сопротивлении  $RI$  изобразится вектором, параллельным току. Этот вектор вращается точно так же, отбрасывая на горизонтальную ось проекцию – настоящее напряжение.

При токе  $I = I_0 \cos(\omega t)$  напряжение на конденсаторе (пропорциональное заряду, т.е. интегралу от тока) будет  $(I_0/\omega C) \sin(\omega t)$ . Его изображаем вектором длины  $I_0/\omega C$ , направленным под прямым углом к току вниз на рисунке 1.2. Как раз при повороте на угол  $\omega t$  его горизонтальная проекция будет  $(I_0/\omega C) \sin(\omega t)$ .

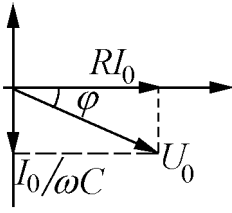


Рис. 1.2.

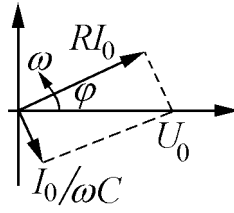


Рис. 1.3.

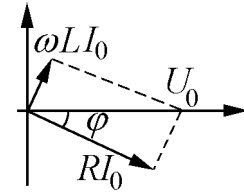


Рис. 1.4.

Сумма реальных напряжений – это сумма проекций на горизонтальную ось, равная проекции суммы векторов. Поэтому можно нарисовать вектор суммарного напряжения длиной  $U_0 = I_0 \cdot \sqrt{(1/\omega C)^2 + R^2}$ . Это и будет амплитуда входного напряжения. Напряжение отстает по фазе от тока на угол  $\varphi$ , причем  $\operatorname{tg}\varphi = 1/\omega RC$ .

Для полного соответствия с предыдущим решением можно повернуть картинку в исходном положении на угол  $\varphi$ , тогда горизонтальным вектором изобразится напряжение, а ток будет вектором длины  $U_0/\sqrt{R^2 + (1/\omega C)^2}$ , имеющим начальный наклон  $\varphi$ . Далее вся диаграмма крутится со скоростью  $\omega$  против часовой стрелки (рис. 1.3).

Попробуем метод векторных диаграмм на  $LR$  – цепи. В электротехнике в виде исключения удобнее система СИ. Для емкостей и сопротивлений формулы СИ и СГС не отличаются: можно считать, что мы уже перешли к СИ, измеряя ток в амперах, напряжение – в вольтах, емкость в фарадах, а сопротивление в омах. Эдс индукции в СИ выражается проще, чем в СГС (нет коэффициента  $c^2$ ):  $\mathcal{E} = -d\Phi/dt = -L \cdot dI/dt$ . Чтобы скомпенсировать эту эдс и прогонять ток через индуктивность (представляющую собой, попросту, кривой провод), нужно приложить напряжение обратного знака:  $U_L = +L \cdot dI/dt$ . При токе  $I_0 \cos(\omega t)$  напряжение на индуктивности  $U_L = -\omega LI_0 \sin(\omega t)$ . Это – проекция вектора длины  $\omega LI_0$ , опережающего ток на угол  $\pi/2$ . Сумма векторов напряжений будет  $I_0 \cdot \sqrt{(\omega L)^2 + R^2}$  и направлена под углом  $\varphi$  к току, тоже с опережением;  $\operatorname{tg}\varphi = \omega L/R$ . Если задать напряжение горизонтальным,  $U = U_0 \cos(\omega t)$ , придаем картинке начальный поворот на угол  $\varphi$  назад (рис. 1.4). Имеем  $I = (U_0/\sqrt{R^2 + (\omega L)^2}) \cos(\omega t - \varphi)$ . Результат получается быстрее и нагляднее, чем при прямом решении дифференциального уравнения.

Существует специальный математический аппарат, позволяющий еще ускорить решение таких задач. По формуле Эйлера  $\exp(ix) = \cos(x) + i \sin(x)$ . Запишем входное напряжение в виде  $U_0 \exp(i\omega t)$ , то есть прибавим к настоящему  $U$  мнимую часть. Разумеется, она вызовет добавочный мнимый ток. Суммарный ток тоже ищем в виде  $I = I_0 \exp(i\omega t)$ . Мнимые части мы отбросим в конце расчета.

Напряжение на индуктивности  $U_L = +L \cdot dI/dt$ . При токе  $I = I_0 \exp(i\omega t)$  напряжение  $U_L = i\omega LI_0 \exp(i\omega t) = i\omega L \cdot I$ . Решаем:

$$RI + i\omega LI = U; \quad I = \frac{U}{R + i\omega L},$$

так как во всех слагаемых одинаковая зависимость от времени:  $\exp(i\omega t)$ . Вопрос, как это понимать. Вспомним, что надо взять действительную часть тока. Полезна **показательная форма** комплексного числа:

$$z = x + iy = \sqrt{x^2 + y^2} \cdot \exp(i\varphi), \text{ где } \operatorname{tg}\varphi = y/x.$$

Знаменатель  $R + i\omega L = \sqrt{R^2 + (\omega L)^2} \exp(i\varphi)$ , где  $\operatorname{tg}\varphi = \omega L/R$ . Тогда

$$I = \frac{U_0}{\sqrt{R^2 + (\omega L)^2}} \cdot \exp(i(\omega t - \varphi)) \Rightarrow \frac{U_0}{\sqrt{R^2 + (\omega L)^2}} \cdot \cos(\omega t - \varphi).$$

Получен тот же результат, что при векторном способе, но заметно быстрее, даже при первой пробе. Аналогично для конденсатора току  $I = I_0 \exp(i\omega t)$  соответствует напряжение  $U_C = I_0 \exp(i\omega t)/(i\omega C) = I/(i\omega C)$ .

Итак, для каждого элемента при токе в виде мнимой экспоненты комплексное напряжение пропорционально току, т.е. существует аналог сопротивления:

$$U = ZI; \quad Z_R = R, \quad Z_C = \frac{1}{i\omega C}, \quad Z_L = i\omega L.$$

Эти комплексные «сопротивления» называются **импедансами**. Для них выполняются те же правила соединения и законы Кирхгофа, как для обычных сопротивлений, и задачи решаются так же, как в цепях постоянного тока. Только в конце надо выделить действительную часть.

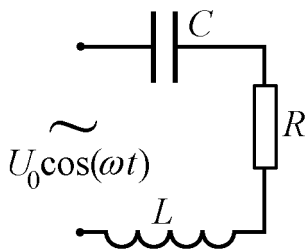


Рис. 1.5.

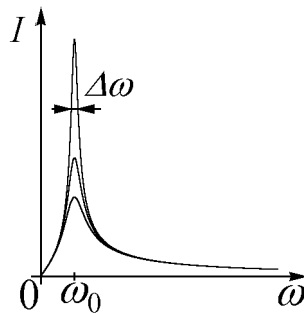


Рис. 1.6.

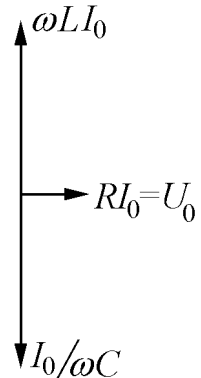


Рис. 1.7.

Для резонансного  $RLC$ -контура (рис. 1.5)  $RI + (1/i\omega C)I + i\omega LI = U$ ;

$$I = \frac{U}{R + i(\omega L - 1/\omega C)} \Rightarrow \frac{U_0}{\sqrt{R^2 + (\omega L - 1/\omega C)^2}} \cdot \cos(\omega t - \varphi),$$

где  $\operatorname{tg}\varphi = (\omega L - 1/\omega C)/R$ . Зависимость амплитуды тока от частоты – резонансная кривая – похожа на «механическую» (на рис. 1.6 показаны три кривые, у которых отличаются сопротивления). Здесь тоже внешняя сила вызывает колебания, причем аналог массы –  $L$ , трения –  $R$ , а упругости –  $1/C$ . Максимум тока достигается, когда

$\omega L = 1/\omega C$ ;  $\omega_0 = 1/\sqrt{LC}$  (собственная частота контура). Вблизи  $\omega_0$  выражение  $(\omega L - 1/\omega C) \cong 2 \cdot \sqrt{L/C} \cdot \Delta\omega/\omega_0$ . При характерных величинах  $L = 10$  мкГ,  $C = 1$  нФ,  $R = 0,1$  Ом  $\omega_0 = 10^7$  Гц, что соответствует длине волны  $2\pi c/\omega_0 = 200$  м. **Волновое сопротивление**  $\sqrt{L/C} = 100$  Ом. Ширина резонансной кривой  $2\Delta\omega$  определяется по падению амплитуды тока в  $\sqrt{2}$  раз, то есть  $(\omega_0 + \Delta\omega)L - 1/(\omega_0 + \Delta\omega)C = R$ ; отсюда  $2 \cdot \sqrt{L/C} \cdot \Delta\omega/\omega_0 = R$ ,  $2\Delta\omega/\omega_0 \sim 10^{-3}$ . Такое резкое падение амплитуды с удалением от резонанса типично в радиотехнике и позволяет настроиться на нужную станцию (частоту), не слыша других.

Если пользоваться импедансами, то задачи о цепях переменного тока до определенного этапа формально решаются так же, как это делается при постоянном токе. Однако при переходе к реальным значениям появляются существенные отличия. Например, в резонансе напряжение на сопротивлении  $RI$  равно входному  $U$ ; можно было бы подумать, что на  $L$  и  $C$  совсем не будет напряжения. Это не так:  $U_L = i\omega LI$ , амплитуда напряжения на индуктивности  $\omega_0 LI = \sqrt{L/C} \cdot I$ , что для «хорошего» контура гораздо больше, чем входное напряжение  $U = RI$  (в описанном выше примере  $\sim 1000 \cdot U$ ). Столько же в резонансе будет на емкости. Но в сумме  $U_L + U_C = 0$ , так как эти напряжения точно в противофазе. Векторная диаграмма показана на рисунке 1.7. На постоянном токе таких интересных эффектов не бывает: все напряжения одного знака, и на любом элементе последовательной цепи напряжение меньше входного. Показания вольтметров, подключенных ко всем элементам цепи и ко входу, могут сильно удивить неподготовленного наблюдателя. Отметим, что на индуктивности напряжение опережает ток по фазе на  $\pi/2$ , а на емкости отстает: действительно, для  $L$  сначала прикладывают напряжение, а затем возникнет ток; для  $C$  вначале нужно пустить ток и после зарядки появится напряжение.

Есть особенности и у мощности в цепях переменного тока. На сопротивлении  $N = RI_0^2 \cos^2(\omega t)$ . Эта величина меняется 100 раз в секунду (для промышленной частоты 50 Гц) и нас, как правило, интересуют не ее колебания, а среднее значение  $\langle N \rangle = RI_0^2/2$ . Приборы показывают **эффективное** значение тока, равное постоянному току, который нагревает сопротивление с той же эффективностью, то есть  $I_{eff} = I_0/\sqrt{2}$  для синусоидального тока. Аналогично  $U_{eff} = U_0/\sqrt{2}$ . В розетке  $U_{eff} = 220$  В, а  $U_0 \cong 311$  В.

Для цепи со сдвигом фаз  $U = U_0 \cos(\omega t)$ ,  $I = I_0 \cos(\omega t - \varphi)$ . Средняя мощность  $N = U_0 I_0 \langle \cos(\omega t) \cos(\omega t - \varphi) \rangle = (1/2) \cdot U_0 I_0 \cos \varphi = U_{eff} I_{eff} \cos \varphi$ . Например, на индуктивности и емкости средняя мощность нулевая (нечему нагреваться). Для мощных потребителей  $N$  и  $U_{eff}$  заданы. Поэтому малый  $\cos \varphi$  приводит к потреблению большого тока, что нежелательно (растут потери). Стараются приблизить  $\cos \varphi$  к 1.

Заметим, что при расчете мощности комплексные функции неудобны. Действительная часть произведения  $IU$  не равна произведению действительных частей (физических величин). Мнимая часть легко отделяется только при линейных операциях.

Из курса анализа известно, что любая периодическая функция раскладывается в ряд **Фурье** – набор простых синусоидальных **гармоник**. Например, последовательность прямоугольных импульсов амплитуды  $\pm 1$  («испорченный косинус»)

$$\text{sign}(\cos(x)) = (4/\pi) \cdot (\cos(x) - \cos(3x)/3 + \cos(5x)/5 - \dots) ,$$

что легко проверяется на компьютере. Когда такое напряжение подано на вход цепи, находят по отдельности токи, вызываемые каждой гармоникой (пользуясь изложенными выше способами), а потом суммируют все токи. Применение тригонометрических функций – **преобразование Фурье** – позволяет дифференциальные уравнения превратить в алгебраические.

## Математическое дополнение: комплексные числа

Обсудим важнейшие характеристики, свойства и способы представления комплексных чисел, стараясь избегать формальных выводов.

### 1. Алгебраическое представление.

Это самая примитивная форма записи комплексного числа (и не самая полезная):

$$z = x + i \cdot y .$$

Здесь  $x$  – действительная, или реальная часть (обозначается также  $\text{Re}(z)$ ),  $y$  – мнимая часть (обозначается также  $\text{Im}(z)$ ),  $i = \sqrt{-1}$  – мнимая единица. Алгебраическое представление удобно, если мы хотим найти сумму и разность комплексных чисел – надо по отдельности складывать действительные и мнимые части:

$$(5 + 3 \cdot i) + (4 + 2 \cdot i) = 9 + 5 \cdot i .$$

Умножение комплексных чисел производится путем раскрытия скобок, с учетом того, что квадрат  $i$  равен  $-1$ . Получается довольно невразумительно:

$$(5 + 3 \cdot i) \cdot (4 + 2 \cdot i) = 20 + 12 \cdot i + 10 \cdot i + 6 \cdot i^2 = 14 + 22 \cdot i .$$

Еще неудобнее будет выглядеть правило деления.

**Сопряженным** числу  $z = x + i \cdot y$  называют  $\bar{z} = x - i \cdot y$ . Произведение любого числа на сопряженное всегда действительно (и даже положительно):

$$(5 + 3 \cdot i) \cdot (5 - 3 \cdot i) = 25 + 15 \cdot i - 15 \cdot i - 9 \cdot i^2 = 34; \quad (x + y \cdot i) \cdot (x - y \cdot i) = x^2 + y^2 .$$

### 2. Геометрическое представление.

Комплексное число можно изобразить и как точку на плоскости с координатами  $x$  и  $y$ . А точка на плоскости – это все равно что вектор с такими проекциями (рис. 1.8):

В алгебраической записи  $i$  можно понимать как единичный вектор по оси  $y$ . Единичный вектор по оси  $x$  – это просто действительная единица и специально не выделяется.



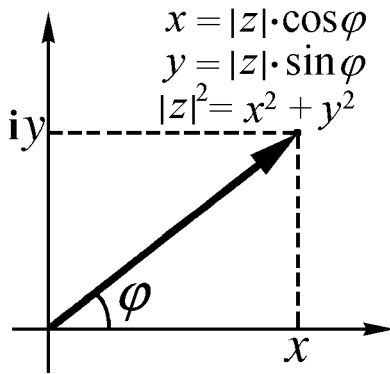


Рис. 1.8.

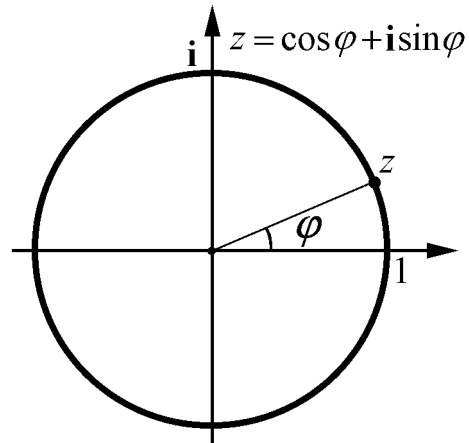


Рис. 1.9.

Такая плоскость называется **комплексной плоскостью**. Геометрическое, или векторное, представление придает простой смысл операциям сложения и вычитания: проекция суммы равна сумме проекций, и выполняются обычные правила соединения стрелок.

Геометрическое представление естественно приводит к новому способу задания чисел. **Модулем** комплексного числа  $|z|$  называется длина вектора, а **аргументом**  $\arg z$  – угол между вектором и осью  $x$ :

$$|z| = \sqrt{x^2 + y^2}, \quad \arg z = \arctg(y/x) = \arccos(x/|z|) = \arcsin(y/|z|).$$

Квадрат модуля равен произведению комплексного числа на сопряженное:  $|z|^2 = z \cdot \bar{z}$ .

### 3. Тригонометрическое представление.

Из предыдущего ясно, что комплексное число можно записать и так:

$$z = |z| \cdot (\cos \varphi + i \cdot \sin \varphi),$$

если  $\varphi$  – это аргумент  $z$ . Здесь вся «комплексность» сидит в скобке. Все числа вида  $(\cos \varphi + i \cdot \sin \varphi)$  имеют единичный модуль и, значит, в комплексной плоскости лежат на единичной окружности (рис. 1.9).

Тригонометрическая форма позволяет наглядно понять смысл произведения двух комплексных чисел. Перемножим два числа с единичным модулем:

$$(\cos \varphi_1 + i \sin \varphi_1)(\cos \varphi_2 + i \sin \varphi_2) = (\cos \varphi_1 \cos \varphi_2 - \sin \varphi_1 \sin \varphi_2) + i(\sin \varphi_1 \cos \varphi_2 + \cos \varphi_1 \sin \varphi_2).$$

Но это попросту  $\cos(\varphi_1 + \varphi_2) + i \cdot \sin(\varphi_1 + \varphi_2)$ . Получается, что при перемножении двух комплексных чисел их аргументы складываются. (Если есть еще модули, они просто перемножаются, как действительные числа, каковыми модули и являются).

### 4. Экспоненциальное представление.

Про действительные числа мы знаем, что при перемножении складываются логарифмы:

$$\ln(ab) = \ln(a) + \ln(b).$$

Это приводит к догадке, что аргумент комплексного числа – это нечто близко связанное с его логарифмом. И действительно,

$$\cos \varphi + i \cdot \sin \varphi = e^{i\varphi}.$$

Эта неожиданная связь называется **формулой Эйлера**. Попробуем в ней убедиться.

Снова рассмотрим «чистые аргументы» – числа с единичным модулем. Хотя мы еще слабо представляем себе, что могут означать экспоненты с мнимым показателем, но если в них хоть какой-то смысл есть, то они тоже лежат на единичной окружности. Действительно, квадрат модуля равен произведению числа на сопряженное,  $e^{i\varphi} \cdot e^{-i\varphi}$ , что при любом разумном понимании обращается в 1. Первая проверка пройдена.

Перемножить два числа с единичным модулем – все равно, что сложить аргументы. Такое же свойство будет иметь и произведение  $e^{i\varphi_1} e^{i\varphi_2}$ . Это, конечно, еще ничего не доказывает: например, для записи  $e^{7i\varphi}$  тоже аргументы суммировались бы. Но для  $e^{i\sqrt{\varphi}}$  суммирования не получится (как и для любой функции угла в показателе экспоненты, кроме прямой пропорциональности).

Рассмотрим случай **малого** угла  $\varphi$ . Тогда экспоненту можно записать в виде

$$e^{i\varphi} \approx 1 + i\varphi.$$

Это число, аргумент которого – малый угол, тангенс которого есть  $\varphi$ . Тангенс малого угла равен углу. Поэтому аргумент числа  $e^{i\varphi}$  стремится к  $\varphi$  для малых углов. Модуль, как уже установлено, равен 1<sup>2</sup>. Значит, для малых углов формула Эйлера верна.

Произвольный угол, например 49 градусов, можно представить как сумму малых, для которых формула выполняется. Перемножая экспоненты с мнимым показателем, мы складываем углы. Выходит, что формула Эйлера правильна для любых углов<sup>3</sup>.

Логарифм произвольного комплексного числа  $z = |z| \cdot e^{i\varphi}$  можно записать в виде

$$\ln z = \ln |z| + i \cdot \varphi.$$

## 1.2 Простейшие импульсные цепи

Возьмем  $RC$  – цепь с разрядом конденсатора. Уравнение:

$$R \cdot dI/dt + I/C = 0,$$

<sup>2</sup>На первый взгляд, модуль больше 1, так как уже действительная часть равна 1. Но это вышло из-за пренебрежения следующими порядками. Например, при учете второго порядка получится  $1 - \varphi^2/2 + i\varphi$ . Можно усмотреть в действительной части разложение косинуса. Этот уточненный модуль равен 1 с точностью до четвертого порядка. Выше мы установили, что  $|e^{i\varphi}| = 1$  точно, и сейчас достаточно разобраться с аргументом этого числа.

<sup>3</sup>Это следует и из указанной выше линейной связи показателя экспоненты с аргументом.

решение:  $I = (U_0/R) \cdot \exp(-t/RC)$ . Характерное время разряда  $\tau = RC$ . Например, при  $C = 10^{-2}$  Ф,  $R = 0,01$  Ом,  $U = 10$  кВ (довольно большая батарея) время разряда 100 мкс, начальная мощность  $10^{10}$  Вт = 10 ГВт – как у крупной электростанции. Полезно заметить, что тепло на сопротивлении  $Q = \int RI^2 dt = CU_0^2/2$  при любом  $R$ : если уменьшать сопротивление, растет ток и уменьшается время разряда, причем так, что вся энергия конденсатора выделяется на сопротивлении.

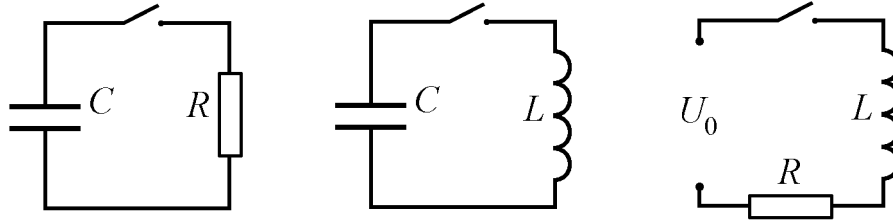


Рис. 1.10.

$LC$  – цепь: уравнение  $Ld^2I/dt^2 + I/C = 0$ ; решение:  $I = I_0 \sin(\omega t)$ , где  $\omega = 1/\sqrt{LC}$ . Амплитуда  $I_0$  определяется из начального условия: при  $t = 0$   $LdI/dt = \omega LI_0 = U_0$ . Окончательно

$$I = (U_0/\omega L) \sin(\omega t).$$

Отношение амплитуды напряжения к амплитуде тока  $\omega L = \sqrt{L/C}$  – уже знакомое волновое сопротивление.

Наконец,  $RL$  – цепь, на которую подано постоянное напряжение  $U_0$ :  $LdI/dt + RI = 0$ ;  $I = (U_0/R) \cdot (1 - \exp(-Rt/L))$ . Время нарастания тока  $\tau = L/R$ .

Сложнее случай полной  $RLC$ –цепи. Качественно при малом сопротивлении естественно ожидать колебаний, как в  $LC$ , но медленно затухающих. При большом же сопротивлении должно быть похоже на  $RC$ –цепь, а индуктивность должна давать завал начального фронта импульса тока на малое время  $L/R$ . Но эти варианты поведения совсем не похожи. Попробуем разобраться. Уравнение цепи:

$$L \frac{d^2 I}{dt^2} + R \frac{dI}{dt} + \frac{I}{C} = 0.$$

Попробуем снова подставить  $I = I_0 \exp(i\omega t)$ . Получается уравнение

$$-\omega^2 + i\omega R/L + 1/LC = 0,$$

которое удовлетворяется, если  $\omega_{1,2} = iR/2L \pm \Omega$ , где  $\Omega = \sqrt{1/LC - R^2/4L^2}$ .

Что означает комплексная частота? При  $R = 0$  мнимая часть отсутствует,  $\omega = \Omega = \sqrt{1/LC}$ : имеем собственную частоту  $LC$  – контура. При  $R > 0$ , подставив  $\omega$  в экспоненту  $\exp(i\omega t)$ , получим колебательный множитель  $\exp(i\Omega t)$ , умноженный на  $\exp(-(R/2L)t)$ . Получается, что мнимая часть частоты означает затухание.

Заметим, что если две функции – решения уравнения, то их сумма – тоже решение (линейность). Попробуем записать решение в виде:

$$A \cdot \exp(i\omega_1 t) + B \cdot \exp(i\omega_2 t) = \exp(-Rt/2L) \cdot (A \exp(i\Omega t) + B \exp(-i\Omega t)).$$

Из формулы Эйлера  $(\exp(ix) - \exp(-ix))/2i = \sin(x)$ . Следует выбрать амплитуды  $A$  и  $B$  так, чтобы получился именно синус (косинус означал бы скачок тока). Тогда  $I = I_0 \sin(\Omega t) \cdot \exp(-Rt/2L)$ . Определяем  $I_0$  из начального условия  $U(0) = L\dot{I}(0)$ :

$$I = \frac{U_0}{\Omega L} \sin(\Omega t) \cdot \exp\left(-\frac{Rt}{2L}\right), \quad \Omega = \sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}}.$$

Это и есть ожидаемые затухающие колебания с частотой, близкой к  $1/\sqrt{LC}$  (на самом деле меньшей). При малом  $R$  затухание медленное – за время  $\sim L/R$  (рис. 1.11, где пунктиром показана ограничивающая экспонента  $\pm \exp(-Rt/2L)$ ).

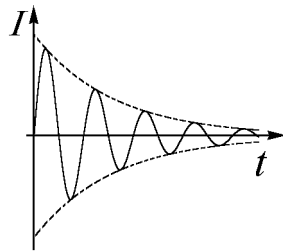


Рис. 1.11.

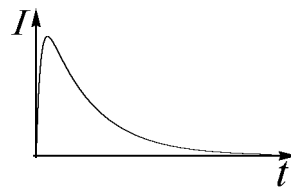


Рис. 1.12.

Что будет при  $R > 2\sqrt{L/C}$ ? Корень  $\sqrt{1/LC - R^2/4L^2}$  тоже получается мнимый – оба решения затухающие. Решение можно записать в виде :

$$I = \frac{U_0}{\Omega L} \left( \frac{\exp(\Omega t) - \exp(-\Omega t)}{2} \right) \cdot \exp\left(-\frac{Rt}{2L}\right), \quad \Omega = \sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}}.$$

Нарастание до максимума занимает примерно время  $L/R$ , после чего будет спад за время  $RC$  (рис. 1.12). Это станет ясно, если воспользоваться биномом Ньютона и упростить  $\Omega$ . Выражение можно сократить, используя гиперболический синус:  $\text{sh}(x) = (\exp(x) - \exp(-x))/2$ , который формально получается как синус мнимого аргумента.

Конденсаторная батарея, упомянутая выше, может иметь индуктивность  $10 \text{ см} = 10^{-8} \text{ Гн}$ . Время нарастания тока при сопротивлении  $0,01 \text{ Ом}$  будет  $1 \text{ мкс}$ . Волновое сопротивление  $10^{-3} \text{ Ом}$ , так что разряд неперiodический.

Теперь можно разобраться, что будет при **внезапном** подключении некоторой цепи к источнику синусоидального напряжения. Например, подключаем  $LRC$  – контур. Известное нам решение, с импедансами и фазовым сдвигом, как правило, не удовлетворяет начальным условиям (ток не нулевой в начальный момент). На самом деле это – установившийся ток. Вначале же к нему добавляются собственные колебания, вызванные подключением – токи, которые могут циркулировать в **замкнутой** цепи, т.е. не

создающие напряжения на входе. Их и следует подобрать так, чтобы решение подходило под начальные условия. Со временем свободные колебания затухают, и остается только стационарная составляющая. Аналогичные **переходные процессы** известны и в механических системах.

Покажем, что импульсные цепи могут выполнять вычисления. Возьмем  $RC$  – цепь с большим  $R$ , так что почти все входное напряжение падает на сопротивлении. Тогда  $I = U/R$ , а напряжение на конденсаторе  $U_C = (1/C) \int I dt = (1/RC) \int U dt$ . На конденсаторе сигнал пропорционален интегралу входного напряжения (интегрирующая цепь). Это будет при  $RC \gg \tau$  – времени изменения сигнала.

Если в той же цепи мала емкость, так что весь входной сигнал «падает» на  $C$ , то  $U = U_C = (1/C) \int I dt$ , ток  $I = C \cdot dU/dt$ . На сопротивлении будет маленький сигнал  $U_R = RC \cdot dU/dt$  – дифференцирующая цепь при  $RC \ll \tau$ .

В обоих случаях, чем точнее цепь выполняет математическую операцию, тем меньше выходной сигнал. В этом можно усмотреть некоторый закон сохранения неприятностей. До появления цифровых компьютеров из таких (или  $RL$  –) цепей строились аналоговые ЭВМ, или интеграторы, иногда довольно сложные. Например,  $LRC$  – цепь «решает» дифференциальное уравнение колебаний и может моделировать, скажем, пролет моста при грамотном выборе номиналов. Если взять связанные цепи, можно получить модель моста.

# ВОЛНОВАЯ ФИЗИКА

Курс лекций для ФМШ

Часть 1. ОПТИКА И ВОЛНЫ

2. ВОЛНОВАЯ ОПТИКА

А. П. Ершов

12 февраля 2007 г.

## Глава 2

# ВОЛНОВАЯ ОПТИКА

### 2.1 Волны. Волновой вектор, частота, длина и скорость волны

Плоская волна задается выражением  $f(x - ct)$ . Функция  $f$  – почти произвольная (на практике есть ограничения), и может быть давлением (звук), высотой поверхности (морские волны), электрическим или магнитным полем (свет). При  $x = ct + \text{const}$  величина  $f$  постоянна: это состояние переносится со скоростью волны  $c$  вправо. Волна,двигающаяся влево, описывается зависимостью  $f(x + ct)$ .

Важен случай синусоидальных, или **монохроматических** волн, например:

$$f = A \cdot \cos(k(x - ct) - \varphi).$$

Здесь  $A$  – **амплитуда**<sup>1</sup> волны. Аргумент косинуса называется **фазой** волны, а постоянная  $\varphi$  – начальной фазой. Подбирая  $\varphi$ , можно получить чистый косинус или чистый синус более простого аргумента  $k(x - ct)$ , а в общем случае выходит их смесь:

$$f = A \cdot (\cos(k(x - ct)) \cos \varphi + \sin(k(x - ct)) \sin \varphi) = B_1 \cdot \cos(k(x - ct)) + B_2 \cdot \sin(k(x - ct)).$$

**Волновой вектор**  $k$  имеет размерность  $1/\text{см}$ . В пространстве  $\mathbf{k}$  задает направление распространения волны, но нам достаточно одного измерения, так что  $k$  – это единственная  $x$  – компонента вектора. **Длина волны**  $\lambda$  есть интервал, на котором значение функции повторяется:  $k\lambda = 2\pi$ ;  $k = 2\pi/\lambda$ . Волна заметно меняется на размере  $1/k = \lambda/2\pi$ . Сама длина волны великовата, чтобы служить характерным масштабом: на ней существенное изменение функции происходит не менее 4 раз (а точнее,  $\lambda$  отличается от  $1/k$  коэффициентом 6,28...).

Другая запись:  $A \cdot \cos(kx - \omega t)$ , где  $\omega = ck$ . Величина  $\omega$  называется **частотой** волны. Если период волны  $T$ , то  $\omega T = 2\pi$ ;  $\omega = 2\pi/T$ . За время  $1/\omega = T/6,28$  с

---

<sup>1</sup>В узком смысле максимального значения.

волной в данной точке пространства будут происходить заметные изменения. В XIX веке пользовались частотой  $\nu = \omega/2\pi = 1/T$  (число колебаний в секунду), а  $\omega$ , под именем круговой частоты, встречалась реже. Теперь больше употребляется  $\omega$ , как более удобная и лучше определяющая временные характеристики. Для сравнения можно сказать, что  $k$  – это как бы пространственная частота, а  $\lambda$  – пространственный период. Частота волны (или соответствующая длина волны) определяют цвет, почему волну с определенной частотой и называют монохроматической (греч.: одноцветной).

Скорость  $c = \omega/k$  называется **фазовой**. С этой скоростью передвигаются гребни и нули синусоиды, то есть точки, в которых фаза косинуса постоянна.

Сложением двух встречных волн получаем **стоячую** волну:

$$A \cdot \cos(kx - \omega t) + A \cdot \cos(kx + \omega t) = 2A \cdot \cos(kx) \cdot \cos(\omega t) .$$

Она никуда не распространяется, хотя в каждой точке происходят колебания. Нули (узлы) и максимумы (пучности) не передвигаются, а стоят на месте. Обычно вторую волну получают отражением. Хорошие стоячие двумерные волны можно наблюдать в прямоугольном баке с водой, если его потрясти<sup>2</sup>.

Для отдельной монохроматической волны нельзя указать ее определенного положения: она присутствует во всем пространстве. Теперь сложим две близкие волны, у которых есть малые отличия волновых векторов ( $2\delta k$ ) и частот ( $2\delta\omega$ ):

$$\begin{aligned} A \cdot \cos((k - \delta k)x - (\omega - \delta\omega)t) + A \cdot \cos((k + \delta k)x - (\omega + \delta\omega)t) = \\ = 2A \cdot \cos(\delta k \cdot x - \delta\omega \cdot t) \cdot \cos(kx - \omega t) . \end{aligned}$$

Получаем (рис. 2.1) «основную» волну (второй  $\cos$ ) с медленно меняющейся (на размере  $\sim 1/\delta k$ ) амплитудой (первый  $\cos$ ). На расстоянии  $\pi/(2\delta k)$  от максимума накапливается разность фаз двух исходных волн, равная  $\pi$ , и составляющие гасят друг друга. На еще больших расстояниях фазовый сдвиг приближается к  $2\pi$  и амплитуда опять вырастает, и т.д. Получается множество областей, в которых волны складываются, разделенных участками «нулей», где происходит взаимное гашение. Заметим, что области большой амплитуды перемещаются со скоростью  $\delta\omega/\delta k$ . Такая волна локализована в пространстве уже несколько лучше монохроматической. Это достигнуто ценой небольшого нарушения монохроматичности: исходная пара волн имела слегка различные частоты.

Добавив еще пару волн с частотами  $\omega_0 \pm \delta\omega/2$ , можно подавить максимумы на расстоянии  $\pi/\delta k$ . Продолжая этот прием, из большого количества почти одинаковых волн, у которых  $\omega$  и  $k$  немного (в пределах  $\delta\omega$  и  $\delta k$ ) отличаются, можно собрать **волновой пакет**, или **цуг волн** (рис. 2.2), у которого амплитуда огибающей практически нулевая (точнее, быстро убывает) вне области размера  $\sim 1/\delta k$ . При малом  $\delta k$  в таком пакете

<sup>2</sup>Раньше такой бак, с готовым механизмом для тряски, имели некоторые типы стиральных машин. Сейчас можно воспользоваться емкостью для овощей из холодильника.



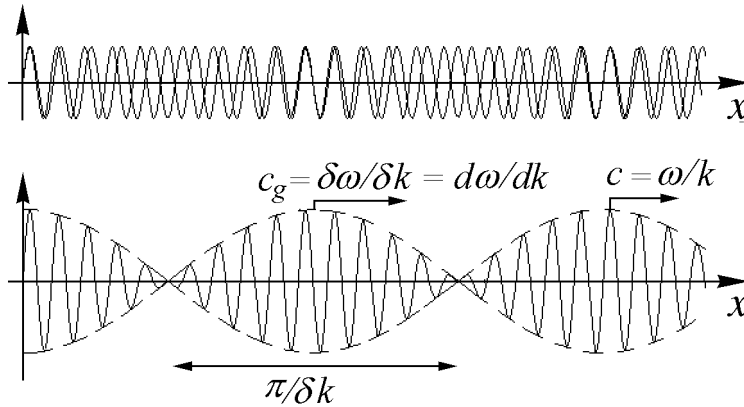


Рис. 2.1.

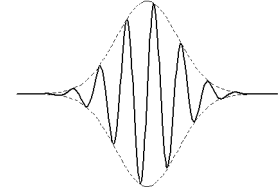


Рис. 2.2.

много ( $\sim k/\delta k$ ) гребней, и он распространяется подобно частице. Можно с точностью до размера  $1/\delta k$  говорить о его положении в пространстве. Скорость перемещения пакета – это скорость его огибающей:  $c_g = \delta\omega/\delta k = d\omega/dk$ .

Эта **групповая** скорость может отличаться от фазовой. Например, для волн на глубокой воде  $\omega \cong \sqrt{gk}$ ;  $c = \sqrt{g/k}$ , а  $c_g = d\omega/dk = (1/2) \cdot \sqrt{g/k}$ . Мы будем видеть эту группу (то есть ее медленную огибающую) распространяющейся со скоростью  $c_g$ . Отдельные же гребни будут «вылезать» из хвоста пакета, достигая максимума в его середине, а затем уменьшаться и пропадать в передней части, двигаясь со скоростью  $c$ , вдвое большей. Бывает, что  $c_g > c$ , и даже они могут быть разных знаков.

Смысл различия в том, что волне нельзя приписать единственной «правильной» скорости, как какой-нибудь материальной точке. Известна аналогия с облаком (Л.И. Мандельштам): капли воды обычно падают вниз, а граница облака может перемещаться и вверх при испарении этих капель. Тоже нельзя однозначно определить, что такое скорость облака. Сидя в лодке, раскачиваемой набегающими волнами, мы интересуемся фазовой скоростью, а ожидая прихода на берег волн от катера, надо прикинуть групповую скорость. С групповой скоростью переносится энергия волны.

К счастью, в акустике и оптике обычно выполняется линейный закон дисперсии:  $\omega = ck$ ,  $c = \text{const}$ . Тогда  $c_g = c$ : групповая и фазовая скорости совпадают. Отклонения возможны, когда частота  $\omega$  близка к частотам внутренних процессов в среде.

## 2.2 Интерференция волн. Интенсивность. Принцип Гюйгенса

Сложим две волны, фазы которых отличаются на константу  $\varphi$  (как говорят, сдвинутые по фазе):

$$A \cdot \cos(kx - \omega t) + A \cdot \cos(kx - \omega t + \varphi) = 2A \cos(\varphi/2) \cdot \cos(kx - \omega t + \varphi/2) .$$

Выходит волна, амплитуда которой в зависимости от  $\varphi$  может быть от 0 до  $2A$ .

Мы не воспринимаем отдельные световые колебания. Свет имеет длину волны  $\lambda \simeq (3 \div 7) \cdot 10^{-5}$  см. Частота  $\omega = 2\pi c/\lambda \approx 4 \cdot 10^{15}$  1/с и характерное время  $1/\omega \approx 3 \cdot 10^{-16}$  с. Раз это время так малó, должна восприниматься какая-то усредненная характеристика. Среднее электрическое поле равно нулю. Разумно назвать **интенсивностью** волны средний квадрат поля:  $I = A^2 \cdot \langle \cos^2(kx - \omega t) \rangle = A^2/2$  для монохроматической волны. Подразумевается усреднение по времени в каждой точке пространства. Интенсивность суммы двух сдвинутых по фазе на  $\varphi$  волн будет

$$I = 4A^2 \cdot \langle \cos^2(kx - \omega t + \varphi/2) \rangle \cdot \cos^2(\varphi/2) = 4 \cos^2(\varphi/2) \cdot \frac{A^2}{2}.$$

Она может меняться от нуля до учетверенной интенсивности одной волны.

Интенсивность – это энергетическая характеристика (вспомним, что для электромагнитных волн – света – энергия как раз  $E^2/4\pi$ ). Явление взаимного гашения или, наоборот, непропорционального – более чем двукратного – увеличения интенсивности при сложении волн называется **интерференцией**. В быту интерференцию двух источников получить не просто. Допустим, что фазовый сдвиг  $\varphi$  сам изменяется, скажем, с частотой  $10^2$  Гц. Именно с такой частотой «мигает» лампа накаливания, что нам незаметно. Тогда надо усреднить суммарную интенсивность еще и по  $\varphi$ ;  $\langle \cos^2(\varphi/2) \rangle$  тоже даст множитель  $1/2$ , и суммарная интенсивность будет равна сумме интенсивностей отдельных волн  $2(A^2/2)$ . Заметим, что в этом примере стабильность  $\varphi$  сравнительно с основной частотой огромная: порядка  $10^{13}$ .

Чтобы получить интерференцию, надо очень жестко зафиксировать фазовый сдвиг. Поэтому, включив две лампы, мы всегда получаем удвоенную освещенность, а не учетверенную либо нулевую в разных местах: фазы волн от ламп случайны и быстро колеблются. Говорят, что для интерференции волны должны быть **когерентны**, то есть их фазы увязаны между собой. Например, вторая (сдвинутая) волна может получаться разделением пучка одного источника зеркалами и т.п. В акустике нет такого запаса порядков величины, и нетрудно заметить интерференцию от разных струн расстроенного рояля («плавание» звука). Потому-то волновая природа звука очевидна давно, а до света дело дошло только в начале XIX века.

Но часто люди наблюдают интерференцию без всяких приспособлений, не осознавая этого. Например, мыльные пузыри или масляные пятна на воде переливаются разными цветами. Это не потому, что они так раскрашены: при смещении глаза перемещаются и цветные полосы, это самая настоящая интерференция. Если где-то глаз видит красное пятно, значит из этого места идут интерферирующие с усилением красные лучи. Более того, всякий объектив, в том числе хрусталик глаза, есть прибор, сводящий расходящийся из точки пучок лучей в другую точку, что и означает интерференцию. Видеть предметы – это значит регистрировать интерференцию исходящих от них лучей (подробнее мы рассмотрим это в следующей главе, п. 3.5).

**Принцип Гюйгенса** в простейшей форме гласит, что каждая точка фронта волны является источником сферических расходящихся волн. Огибающая этих волн через малое время даст новое положение фронта, и т.д. Например, волна на воде – это поднятие уровня. Ясно, что возвышение подействует на соседнюю воду и приведет ее в движение. В акустике изменившееся давление толкнет соседний газ – и для оптики – **поля  $E$  и  $H$**  на фронте создают такие же поля впереди, а вовсе не **фонарь**, зажженный где-то ранее. Вообще волновые процессы явно предполагают близкодействие.

В более детальной форме принцип Гюйгенса – это приближение, которое применяется в задачах дифракции. Рассмотрим плоскую волну, падающую на непрозрачный экран с отверстием. Отверстие можно рассматривать как источник, от которого через малое время в центральной части волна продвинется, оставаясь плоской, но загнется по краям. По мере продвижения огибающих вдали от отверстия получим волну с почти сферическим фронтом. Это не значит, что отверстие эквивалентно точечному источнику: амплитуда волны на фронте, как мы увидим, отнюдь не одинаковая. Память об исходном параллельном пучке сохраняется, но часть световой энергии расходится в стороны. Эту способность волн загибаться за края препятствий и называют **дифракцией**.

Строго говоря, надо решать уравнения Максвелла с граничными условиями на экране и других границах (например, слева от экрана источник, а справа другой экран, на котором наблюдается результат прохождения света). При этом надо знать свойства материала обоих экранов и источника. В частности, правый экран отражает свет налево, этот свет проходит в отверстие и влияет на источник, и пр. Все это сложно.

Но в случае, когда размеры отверстия гораздо больше длины световой волны, волновые свойства света проявляются слабо, и достаточно учесть их в первом приближении. Именно, считается, что поле в отверстии такое, как создавал бы источник в отсутствие экранов, а на остальной поверхности непрозрачного экрана поле точно нулевое. Каждый элемент  $dS$  площади отверстия есть источник проходящей волны, амплитуда которой пропорциональна произведению  $dS$  на амплитуду падающей волны и спадает обратно пропорционально расстоянию, как и положено в сферической волне. Это позволяет вычислить распределения интенсивности в прошедшей волне.

## 2.3 Дифракция от двух щелей. Дифракционная решетка

Пусть плоская волна, то есть параллельный пучок света, падает на непрозрачный экран, имеющий две щели. Расстояние между щелями  $d$ , ширина щелей гораздо меньше. По принципу Гюйгенса, эти щели рассматриваем как два источника, излучающие когерентно: точно в фазе. Что будет видно на втором, удаленном на большое расстояние  $L$ , экране?

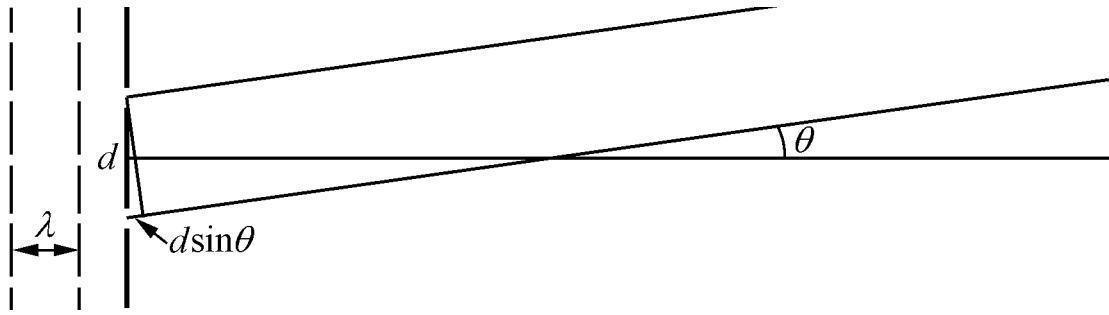


Рис. 2.3.

Выберем угол  $\theta$  отклонения лучей. При удаленном экране ( $L \gg d$ ) лучи, сходящиеся на экране, практически параллельны и имеют **разность хода**  $\Delta = d \cdot \sin \theta$ . (На рис. 2.3 экран не удалось достаточно отодвинуть, и кажется, что лучи не сходятся в одной точке; здесь отдано предпочтение параллельности лучей, а не их сведению, чтобы яснее показать  $\Delta$ ). Поскольку расстояние  $L$  велико, на ослабление волн от обеих щелей малая разница  $\Delta$  не влияет, т.е. на экране складываются волны одинаковой амплитуды. Однако различие **фаз** этих волн  $k\Delta = 2\pi\Delta/\lambda$  учитывать нужно (так как длина волны  $\lambda$  мала). Если разность хода равна целому числу  $n$  длин волн или нулю, то волны от двух источников имеют разность фаз, кратную  $2\pi$ , и будут складываться. Наоборот, когда разность хода равна  $\lambda/2 + n\lambda$ , волны будут друг друга гасить. Поэтому на удаленном экране будет не две светлых полосы на расстоянии  $d$  друг от друга, как, скажем, от двух окон, а чередование большого числа светлых и темных полос. Угол  $\theta$  для первой темной полосы:  $\sin \theta = \theta = \lambda/2d$ , положение ее на экране  $y = L\theta = L\lambda/2d$ . При  $y = L\lambda/d$  будет светлая полоса, дальше опять темная, и т.д.

Найдем распределение интенсивности. Вместо тригонометрической удобнее комплексная запись волны. Пусть волна  $A \cdot \cos(kx - \omega t)$  приходила бы по направлению  $\theta$  на экран из источника, находящегося точно между щелями. Тогда от верхней щели придет сигнал  $A \cdot \cos(kx - kd \sin \theta/2 - \omega t)$  с более ранней фазой, а от нижней придет  $A \cdot \cos(kx + kd \sin \theta/2 - \omega t)$  с запаздыванием по фазе. Складываем, пользуясь преобразованием суммы косинусов:

$$A \cdot (\cos(kx - \omega t - kd \sin \theta/2) + \cos(kx - \omega t + kd \sin \theta/2)) = 2A \cdot \cos(kx - \omega t) \cdot \cos(kd \sin \theta/2).$$

Это, скажем, электрическое поле в точке экрана, определенной углом  $\theta$ . Чтобы найти интенсивность, возводим в квадрат и усредняем по времени:

$$I = 2A^2 \cdot \cos^2(kd \sin \theta/2).$$

Заметим, что при  $\theta = 0$  наблюдается центральный максимум, как раз где должна быть геометрическая тень. Первый нуль будет при  $kd \sin \theta/2 = \pi/2$  или  $\sin \theta = \lambda/2d$ , смысл чего разъяснен выше.

Как и в задачах переменного тока, удобно сложение волн изображать графически. Волну от одной щели в точке наблюдения рисуем в виде горизонтального единичного

вектора, а от второй – вектора такой же длины, повернутого на угол, соответствующий сдвигу фаз (рис. 2.4). При разности хода  $\lambda/4$  сдвиг равен  $\pi/2$ , при  $\lambda/2$  уже будет  $\pi$ , и т.д. Далее вся конструкция вращается против часовой стрелки с угловой скоростью  $\omega$ . Амплитуда в точке наблюдения будет суммой векторов. Видно, что она может меняться от 0 до 2, а интенсивность – от 0 до 4.

Сколько же будет полос от двух щелей? Одна щель тоже дает дифракцию. Пусть ширина щели  $a$ , представляем ее как набор линейных источников, дающих в точке наблюдения амплитуду, пропорциональную их ширине  $dz$ . В прямом направлении ( $\theta = 0$ ) все источники складываются в одной фазе, там будет максимум. Рассмотрим такое направление, при котором разность хода от верхнего и нижнего

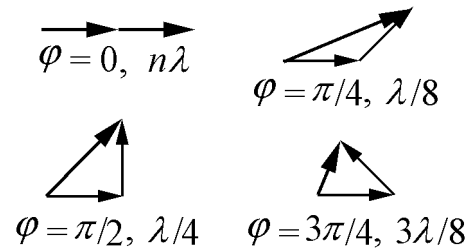


Рис. 2.4.

краев щели  $a \sin \theta = \lambda$ : волны от краев приходят в одной фазе. Зато элементарный источник, расположенный у верхнего края, уничтожается с таким же источником на середине щели. Всю щель можно разбить на такие гасящие друг друга пары. Следовательно, в направлении  $\sin \theta = \lambda/a$  получим нулевую интенсивность. То же можно сказать и о направлениях  $\sin \theta = n\lambda/a$ , где  $n$  – целое число. Между этими нулями располагаются максимумы интенсивности.

В геометрическом представлении (рис. 2.5) амплитуда в направлении точно вперед – это сумма многих малых амплитуд, вытянутых в линию. Считаем, что длина этой линии 1. Под углом  $\theta$  максимальная разность хода  $a \sin \theta$ , фазовый сдвиг  $\varphi = 2\pi a \sin \theta / \lambda$ . Если первый из маленьких векторов горизонтален, то последний будет под углом  $\varphi$ , а образуют они дугу окружности длиной опять же 1. Амплитуда будет хордой этой дуги, то есть  $2R \sin(\varphi/2)$ , а дуга  $1 = R\varphi$ . Выражая  $R$ , получим ту же формулу:  $I \sim (\sin(\varphi/2)/(\varphi/2))^2$ . Первый нуль будет при  $\varphi = 2\pi$ , когда дуга загнется в полный круг, и амплитуда станет нулевой. Отсюда легко оценить первый боковой максимум: надо из единичной веревки свернуть полторы окружности. Диаметр будет  $2/3\pi = 0,2122$ , интенсивность  $(2/3\pi)^2 = 0,045$ .

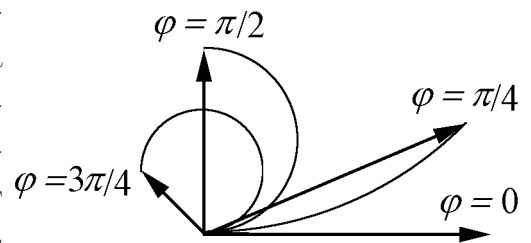


Рис. 2.5.

Здесь, как и в прошлом примере, свет не идет по прямой, а согласно своей волновой природе разбегается в стороны. Угол расхождения, или дифракционный угол, всегда порядка (длины волны)/(поперечник источника). Интенсивность в максимуме  $\sim a^2$ , его угловой размер  $\sim \lambda/a$ , и полная энергия, попадающая на экран, пропорциональна ширине щели  $a$ , как и следовало ожидать. Вклад боковых максимумов – несколько процентов.

Разумеется, дифракция от двух щелей будет наблюдаться, когда дифракционные

пучки отдельных щелей пересекаются. В частности, поэтому мы не видим таких эффектов в комнате с двумя окнами: при макроскопическом отверстии угол расхождения мал<sup>3</sup>; на стене будут видны два ярких пятна даже при когерентном освещении. Если же от достаточно узких щелей главные максимумы пучков перекроются, то в области взаимодействия и будут, собственно, полосы интерференции; их количество  $\sim d/a$ . Отметим, что лазер для наблюдения интерференции необязателен, иначе бы ее не открыл Юнг в 1807 году. Мы часто наблюдаем дифракцию, не подозревая об этом. Например, если смотреть на яркий точечный источник, зажмурившись, можно увидеть цветные полосы; это дифракция на зазорах между ресницами, а экраном служит сетчатка глаза.

Получим распределение интенсивности на экране.

Суммируем сигналы в комплексном виде:

$$\begin{aligned} & \exp(i(kx - \omega t)) \cdot \int_{-a/2}^{a/2} \exp(ikz \cdot \sin \theta) dz = \\ & \frac{e^{i(kx - \omega t)}}{ik \sin \theta} \cdot \left[ \exp\left(\frac{ika \sin \theta}{2}\right) - \exp\left(-\frac{ika \sin \theta}{2}\right) \right] = \\ & e^{i(kx - \omega t)} \cdot a \cdot \frac{\sin(ka \sin \theta/2)}{ka \sin \theta/2}. \end{aligned}$$

Интенсивность получим, усреднив квадрат амплитуды:

$$I \sim \frac{a^2}{2} \cdot \frac{\sin^2(ka \sin \theta/2)}{(ka \sin \theta/2)^2}.$$

При  $\theta = 0$  – главный максимум, при  $0,5 \cdot ka \sin \theta = n\pi$  – нули. Угол на первый ноль  $\theta = 2\pi/ka = \lambda/a$ . Главный максимум вдвое шире боковых (например, между  $\pi$  и  $2\pi$ ). Величина первого бокового максимума около 5% от главного, остальные еще меньше (рис. 2.6).

Пусть в экране много ( $N$ ) щелей. Тогда получается **дифракционная решетка**. Она дает **главные максимумы** в тех же направлениях, что и две щели с тем же разносом  $d$ : при  $d \sin \theta = n\lambda$  все щели работают в фазе. Однако если чуть увеличить  $\theta$ , прибавив всего одну волну на нижнем краю решетки, амплитуда уже обратится в ноль. Как и в случае одной щели, при этом первая щель (источник) сократится с  $N/2$ , вторая с  $N/2 + 1$ . В геометрическом представлении получится замкнутый  $N$  – угольник, при большом  $N$  почти окружность. Видно, что главные максимумы решетки очень узкие ( $\sim \lambda/Nd$ ), зато яркие: в них интенсивность  $\sim N^2$ , а суммарная энергия, разумеется, пропорциональна числу излучателей  $N$ . Между ними находятся  $\approx N$  мелких боковых максимумов, относительная интенсивность первых из них такая же, как в картине дифракции на одиночной щели. Нетрудно получить и распределение интенсивности от решетки подобно уже рассмотренным примерам. Складывая сумму вида  $\sum_n \exp(ikd \sin \theta \cdot n)$ , пользуйтесь формулой геометрической прогрессии.

Дифракционные решетки полезны при исследовании излучения, содержащего волны с разными длинами, позволяя разобраться в его составе. **Спектром** излучения называется набор всех составляющих его волн. Например, спектр волны вида

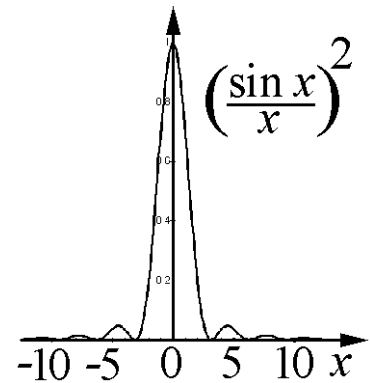


Рис. 2.6.

<sup>3</sup>Не надо путать дифракционное расхождение с размытием границ при освещении протяженным источником (полутени от Солнца, лампы и пр.)



$\cos(\delta k \cdot x - \delta \omega \cdot t) \cdot \cos(kx - \omega t)$ , рассмотренной в п. 2.1 на стр. 14, содержит две компоненты с частотами  $\omega \pm \delta \omega$ . После взаимодействия с дифракционной решеткой каждая компонента образует свои дифракционные линии. На экране (обычно это фотопластинка) получится набор полос, напоминающий наложение двух гребенок с разным шагом. Нулевые главные максимумы ( $n = 0$ ) совпадают, но уже первый максимум ( $n = 1$ ) на фотопластинке распадется на две близко расположенные линии. Это и покажет, что спектр содержит ровно две компоненты. Чем больше штрихов у решетки, тем лучше ее разрешающая способность – линии спектра получают яркими и узкими.

Рентгеновские лучи – это свет с малой длиной волны (например, у меди есть линия 1,54 Å). Кристаллы служат для рентгена дифракционной решеткой с большими углами поворота излучения. Эти углы характерны для каждого кристалла (рентгеноструктурный анализ).

Для отражательной решетки в максимумах отражения должно выполняться **условие Вульфа – Брэгга**:  $d(\sin \theta_i - \sin \theta_r) = n\lambda$ ; углы меряются от вертикали. Обратите внимание, что решетка всегда отражает свет под углом, равным углу падения, но **может** отражать и под совсем другим. Нить лампы накаливания при отражении в граммофонной пластинке видна в нескольких экземплярах (и в радужных цветах). Очень красивые цветные отражения дают компакт-диски. Убедитесь на опыте, что радужные полосы наблюдаются всегда в стороне от геометрического отражения источника света.

## 2.4 Преломление и отражение света. Показатель преломления

Рассмотрим гладкую (в масштабе  $\lambda$ ) поверхность. Опыт показывает, что угол отражения от нее равен углу падения. Это сразу следует из волновой картины: если длина волны  $\lambda$ , то периодичность точек касания поверхности с фронтами волн будет  $\lambda / \sin i$ ,  $i$  – угол падения. Для отраженного пучка та же периодичность  $\lambda / \sin r$ ,  $r$  – угол отражения. Эти величины должны совпадать (рис. 2.7), откуда  $i = r$ .

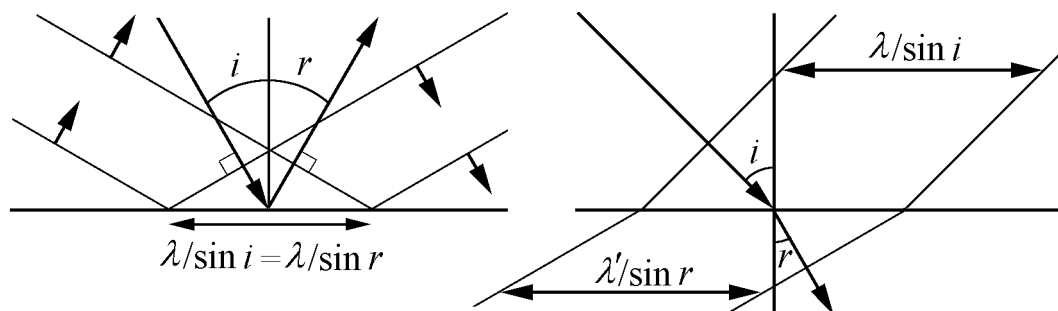


Рис. 2.7.

В прозрачной среде свет имеет меньшую скорость:  $V = c/n$ , **показатель прелом-**

**ления**  $n > 1$ . Частота света в среде не изменится (падающий свет – это вынуждающая сила колебаний). Изменяется длина волны:  $\lambda' = 2\pi c/n\omega = \lambda/n$ . Условие совпадения волн на плоскости:  $\lambda/\sin i = \lambda'/\sin r$ , или  $\sin i/\sin r = n$  (закон Снеллиуса). Преломленный луч идет ближе к нормали. Поток частиц тоже может преломляться таким образом, но только если скорость в веществе, наоборот, увеличивается.

Показатель преломления получается из уравнений Максвелла для поля в среде. В отсутствие токов и зарядов среда полностью характеризуется проницаемостями  $\varepsilon$  и  $\mu$ . Уравнения

$$\oint (\mathbf{E}d\mathbf{L}) = -\frac{1}{c} \frac{d}{dt} \int (\mathbf{B}d\mathbf{S}), \quad \oint (\mathbf{H}d\mathbf{L}) = \frac{1}{c} \frac{d}{dt} \int (\mathbf{D}d\mathbf{S})$$

приводят, аналогично п. 5.4 «Электродинамики», к соотношениям:

$$E = V\mu H/c, \quad H = V\varepsilon E/c, \quad \text{откуда } V = c/\sqrt{\varepsilon\mu}, \quad \sqrt{\mu}H = \sqrt{\varepsilon}E,$$

т.е. энергии полей в волне равны:  $\mu H^2/8\pi = \varepsilon E^2/8\pi$ . Обычно для прозрачных сред  $\mu = 1$  с большой точностью, и показатель преломления  $n = \sqrt{\varepsilon}$ . Соотношение между полями при этом:  $H = E\sqrt{\varepsilon} = nE$ . Для воздуха ( $\varepsilon = 1,00057$ )  $n \approx 1,0003$ . Таким слабым отличием от 1 обычно можно пренебречь. Для стекол типичное значение  $n \approx 1,5$ . Отметим, что  $n$  воды (1,333) не согласуется с известной нам диэлектрической проницаемостью ( $\varepsilon \simeq 80$ ). Оказывается,  $\varepsilon$  зависит от частоты. Значение 80 – статическое, для постоянного или медленно меняющегося поля. Видимый же свет имеет частоту порядка  $3 \cdot 10^{15}$  Гц, при которой  $\varepsilon \approx 1,78$ . Видно, что зависимость  $\varepsilon(\omega)$  может быть существенной.

Рассмотрим простую модель среды, в которой электроны держатся на пружинках жесткости  $k$  (например, как в атоме Томсона). Уравнение движения электрона

$$m \frac{d^2x}{dt^2} = -kx + eE_0 \exp(i\omega t).$$

Длина волны  $\gg$  размеров атома (у света в  $\sim 5000$  раз), так что  $E$  для атома однородно. Как обычно, временную зависимость пишем в экспоненциальном виде. Решение:  $x = eE_0 \exp(i\omega t)/m(\omega_0^2 - \omega^2)$ . Собственная частота  $\omega_0 = \sqrt{k/m}$ . Дипольный момент атома будет  $ex$ , а вектор поляризации  $P = nex = (ne^2/m)/(\omega_0^2 - \omega^2) \cdot E = \alpha E$ . Наконец, диэлектрическая проницаемость

$$\varepsilon = 1 + \frac{4\pi ne^2}{m(\omega_0^2 - \omega^2)}.$$

Условие  $\varepsilon > 1$  выполняется, когда частота света меньше собственной частоты среды. Если же приближаться к собственным частотам (реально их много), то начинается резонанс, колебания уже не малы; надо учитывать затухание. Свет начинает сильно поглощаться. Большинство тел как раз непрозрачны в видимом свете, попадающем в полосу поглощения, а такие, как стекло, случайно имеют «окна» прозрачности. В ультрафиолете стекло уже не пропускает.

Плазма содержит свободные электроны, для которых  $k$  и  $\omega_0$  равны нулю. Для таких сред  $\varepsilon = 1 - 4\pi ne^2/m\omega^2 < 1$ . Фазовая скорость волны больше скорости света в вакууме:



$\omega/k = c/\sqrt{\varepsilon} = c/\sqrt{1 - 4\pi ne^2/m\omega^2}$ , откуда  $\omega = \sqrt{4\pi ne^2/m + c^2 k^2}$ . Величина  $\Omega = \sqrt{4\pi ne^2/m}$  называется плазменной частотой. Групповая скорость

$$c_g = \frac{d\omega}{dk} = c \cdot \frac{ck}{\sqrt{\Omega^2 + c^2 k^2}} < c.$$

При  $\omega < \Omega$  диэлектрическая проницаемость даже меньше нуля. Корень из  $\varepsilon$  – показатель преломления – и вместе с ним скорость волны  $c/n = \omega/k$  получаются мнимыми. При действительной  $\omega$  будет мнимый  $k$ , что означает затухание волны с расстоянием:  $\exp(ikx) = \exp(-|k| x)$ . Это не поглощение, так как наша теория не содержит какого-либо трения: будет отражение низких частот и прохождение высоких. Дальняя радиосвязь на коротких волнах облегчается их отражением от ионизированных слоев (даже от метеоритных следов), а на УКВ необходима прямая видимость, почему и строят телевизионные башни.

Для металлов ( $n \sim 10^{23} \text{ 1/см}^3$ )  $\omega_p \approx 10^{16}$  Гц, примерно как у видимого света. Однако свет даже более высокой частоты не проходит через металл из-за его большой электропроводности. Для меди  $\sigma = 5 \cdot 10^{17} \text{ 1/с}$ , и ток проводимости  $\sigma E$  много больше тока смещения  $\omega E/4\pi$ . В разреженной слабоионизированной плазме током проводимости бывает возможно пренебречь.

Если свет идет из плотной (точнее, оптически плотной) среды в менее плотную (значит, с меньшим  $n$ ), и падение луча далеко от нормали,  $\sin r$  (в воздухе)  $= n \sin i$ , что бывает больше 1. Такой луч, неспособный найти себе угол выхода, отражается (полное внутреннее отражение). Этот эффект используется в световодах.

Мы не учитывали пока поперечность световой волны. Вектор  $\mathbf{E}$  всегда перпендикулярен направлению распространения волны (т.е. вектору  $\mathbf{k}$ ). Можно выбрать два таких взаимно перпендикулярных направления, и по ним разложить произвольный вектор  $\mathbf{E}$ . Значит, произвольный свет можно представить как сумму двух **плоскополяризованных** волн, в каждой из которых электрический вектор колеблется в одной из взаимно перпендикулярных плоскостей, пересекающихся по вектору  $\mathbf{k}$ .

Рассмотрим падение на прозрачную среду волны, поляризованной в плоскости падения (рис. 2.8). В среде электроны колеблются перпендикулярно прошедшему лучу. Довольно ясно, что в направлении колебаний они излучать не могут. Поэтому при совпадении направления колебаний и отраженного луча никакого отражения не будет:  $i + r = \pi/2$ ,  $\sin i / \sin r = n$  или  $\operatorname{tgi} = n$  (угол Брюстера). Под этим углом будет отражаться только свет, поляризованный перпендикулярно плоскости падения: так можно получать полностью поляризованный свет. Для получения линейно поляризованного света используют также специальные поляризаторы, избирательно пропускающие свет с заданным направлением поляризации.

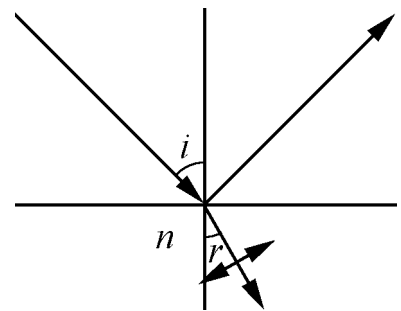


Рис. 2.8.

Коэффициенты отражения в зависимости от  $i$ ,  $n$  получаются из уравнений Максвелла (формулы Френеля). Они довольно громоздки и зависят от поляризации.

*Задача:* найти коэффициент отражения при нормальном падении.

## 2.5 Принцип Ферма. Приближение геометрической оптики

Мы перешли от интерференции и дифракции к прямолинейному распространению света по школьным законам. Есть ли между этими явлениями какое-то соответствие?

Принцип Ферма как раз позволяет в этом разобраться. (Пьер Ферма – автор так называемой Великой теоремы). Согласно этому принципу, свет из одной точки в другую идет по пути, требующему наименьшего времени. Для простого движения из лампы в глаз тогда получится прямая  $AB'$  (рис. 2.9). Если свет попадает из  $A$  в  $B$ , отражаясь от зеркала, все равно из таких путей надо выбрать тот, для которого время наименьшее. Это явно ломаная, положение точки излома определяется минимальностью пути. Отрасим симметрично  $B' \rightarrow B$ ; путь из  $A$  в  $B$  такой же и, очевидно, минимальный. Получается известный закон: угол падения равен углу отражения.

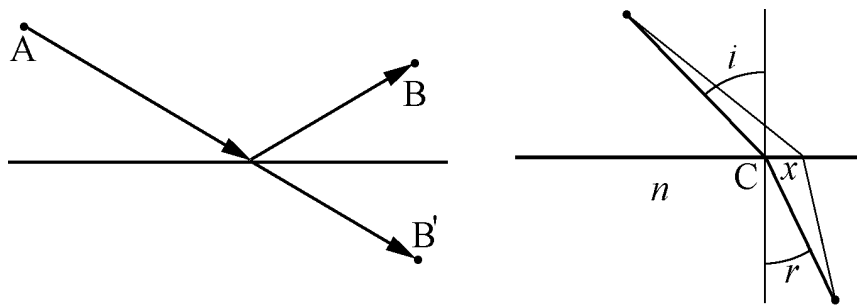


Рис. 2.9.

Так же получается закон преломления. Пусть луч преломляется в точке  $C$ . Если сместиться вдоль поверхности на малое расстояние  $x$ , то верхний путь луча (в воздухе) удлинится на  $x \sin i$ , а нижний (в стекле) укоротится на  $x \sin r$ . Если  $x = 0$  соответствует минимальному времени, то оба изменения должны компенсироваться: в минимуме производная времени по  $x$  равна нулю. Получаем  $x \sin i/c - x \sin r/(c/n) = 0$ , или  $\sin i/\sin r = n$ .

Почему же время должно быть минимальным? Пусть свет шел бы какими-то извилистыми путями, чем они хуже?

Правильный ответ: он именно идет **всеми** возможными путями, а прямолинейность – это суммарный эффект всех путей. Возьмем прямую, которую мы воспринимаем как путь света. Для простоты рассмотрим пути типа показанных на рис. 2.10. На таком

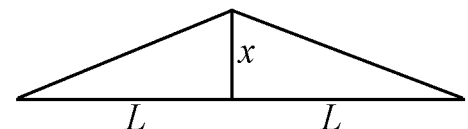


Рис. 2.10.

пути свет набирает разность хода  $s = 2\sqrt{L^2 + x^2} - 2L$  по сравнению с прямым, если  $x$  – высота треугольника. Очевидно, что минимум  $s = 0$  при  $x = 0$ .

Теперь позволим свету распространяться по таким путям. Если пути равноправны и надо учесть все, остается сложить волны, прошедшие все пути, и считать оконча-

тельную волну результатом интерференции. Пусть по прямой придет свет с нулевой фазой. Соседние волны, для которых высота  $x$  мала, придут с малой разностью хода и будут складываться амплитуды. С возрастанием  $x$  появятся большие сдвиги фаз. Найдем, когда сдвиг фаз будет  $\pi$  (то есть полволны). При малом  $x$  из бинома Ньютона  $s = x^2/L = \lambda/2$ ; соответствующее  $x_1 = \sqrt{\lambda L/2}$ . Круг радиуса  $x_1$  называется первой зоной Френеля. Для  $L = 2$  м,  $\lambda = 5 \cdot 10^{-5}$  см  $x_1 = 0,7$  мм.

Представим результат интерференции графически, изображая амплитуды волн, прошедших разными путями, в виде векторов почти одинаковой длины. В духе принципа Гюйгенса эти векторы можно понимать как результат действия источников – тонких колец, расположенных в средней плоскости и имеющих одинаковые площади (малые по сравнению с площадью зоны Френеля).

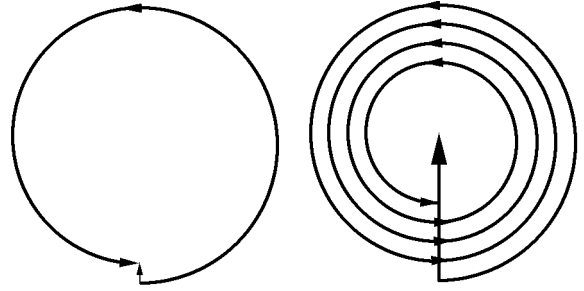


Рис. 2.11.

Сдвиг фаз изображается поворотом векторов. Свет, прошедший через первую зону Френеля, изобразится последовательностью большого числа векторов, образующих почти правильный полукруг, так как фаза конечного вектора сдвинута на  $\pi$  (рис. 2.11). В сумме амплитуда от первой зоны будет равна диаметру круга. Далее начнется вторая зона от  $x_1$  до  $x_2 = \sqrt{2\lambda L/2} = 1,4x_1 \approx 1$  мм. Ее площадь равна площади первой зоны. Сдвиг фаз, или направление векторов, изменяется в ней от  $\pi$  до  $2\pi$ , а последний вектор приходит почти в исходную точку и дает практически нулевую интенсивность. Значит, в точке наблюдения свет, прошедший через вторую зону Френеля, гасит свет, прошедший через первую зону.

Конечно, общее количество света, прошедшего через отверстие радиуса  $x_2$ , вдвое больше, но на оси пучка при таких данных будет темное пятно. На этот парадокс указал Пуассон, когда Френель докладывал свои результаты Французской академии наук. Эксперимент показал, что именно так и есть.

Дальше третья зона практически повторяет первую, четвертая – вторую и т.д. Однако с удлинением пути следует постепенно уменьшать длины векторов, и для большого числа зон образуется спираль, стремящаяся к центру круга, образованного первыми двумя зонами. Амплитуда будет равна половине амплитуды от первой зоны, а интенсивность – четверти.

Видим, что свободно идущий сквозь всю плоскость свет дает в точке приема такой же результат, как если бы он проходил только сквозь часть первой зоны Френеля с поперечником  $\sim 0,1$  мм. Поэтому у нас и возникает впечатление прямолинейности распространения света. При всей очевидности и практической полезности это впечатление совершенно ложное.

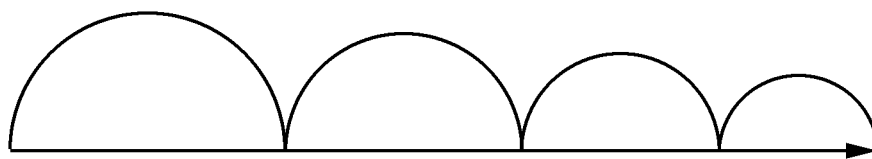


Рис. 2.12.

Поместим посередине между источником и изображением **зонную пластинку**, у которой прозрачны только нечетные зоны Френеля, а четные зачернены. Тогда четные зоны не аннулируют свет, прошедший через нечетные, и интенсивность в точке наблюдения возрастет многократно (рис. 2.12). При этом мы не добавляем новый свет, а, наоборот, убираем часть имеющегося. Такая пластинка называется еще линзой Френеля и действительно дает изображение, как обычная стеклянная линза<sup>4</sup>.

Этот опыт доказывает, что свет **не распространяется прямолинейно**, а использует все возможные пути. Суммирование же вкладов с разными фазами обычно маскирует этот плюрализм и симулирует прямолинейность распространения.

Аналогично при отражении свет не ограничивается точкой геометрического отражения, а прощупывает все пути, в том числе неожиданные. Если выкрасить зеркало в области, далекой от геометрического отражения, полосками черной краски, из него может получиться дифракционная решетка, отражающая свет далеко «вбок».

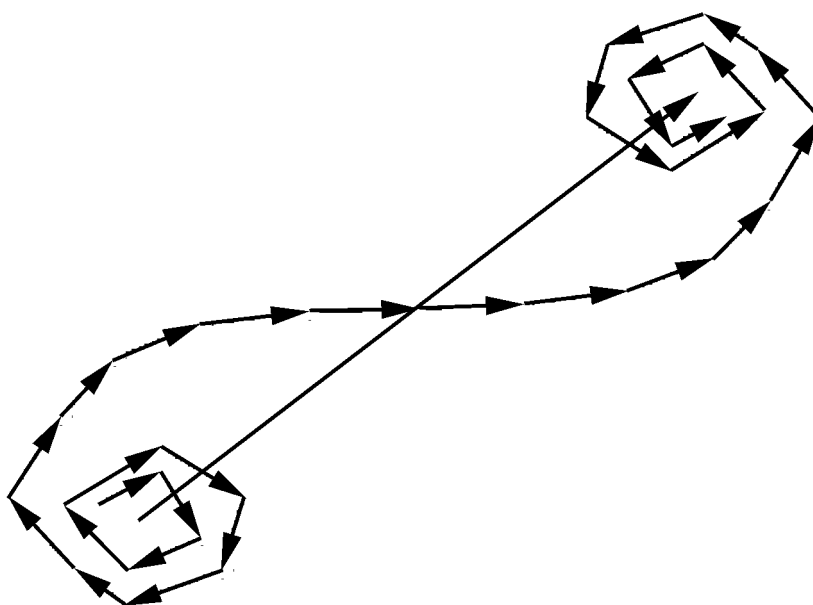


Рис. 2.13.

Тем не менее возможен и другой вывод: законы геометрической оптики в пределах своей компетенции вполне пригодны. Мы можем считать, что свет распространяется прямолинейно и отражается зеркально, когда на его пути нет разных хитрых линз Френеля, малых отверстий или дифракционных решеток. Поскольку результат не зависит от наших представлений, удобно говорить о лучах света, с тем же основанием, как атомы можно представлять себе блестящими шариками.

<sup>4</sup>В самом деле, она собирает свет от точки – источника в точку на экране с другой стороны пластинки. Из нескольких светящихся точек получится столько же точек на экране. Любой светящийся или освещенный предмет – это совокупность точек, испускающих свет. Совокупность их изображений на экране и даст изображение предмета. Например, хорошо получается изображение нити лампы накаливания, причем меняющее цвета при небольшом сдвиге экрана. Цвет определяется длиной волны, для которой выполняется условие Ферма при данном расположении лампы, пластинки и экрана.

Лучами же будут те линии, вблизи которых пути света удачно интерферируют: складываются амплитуды. Это будет, когда разности хода (а в среде разности времени хода) для близких путей малы, что достигается как раз для пути наименьшего времени (см. рис. 2.13). На рисунке схематически показано суммирование вкладов от путей различной длины при отражении от зеркала. Приблизительно прямой участок векторной диаграммы соответствует пути наименьшего времени, когда при малом сдвиге разность хода пропорциональна квадрату смещения. Для произвольного же пути при малом его шевелении разность хода пропорциональна смещению; эти участки соответствуют медленно сходящимся в точку спиралям. Результирующая амплитуда – длинная прямая стрелка, протянутая из одного фокуса в другой. Она примерно совпадает со вкладом «почти прямого» участка.

Еще укажем, что реально существующие тонкие лучи света тоже способны ввести в заблуждение. Они тонкие сравнительно со своей длиной, но очень толстые сравнительно с длиной волны. Луч лазера, похожий на спицу, на самом деле – плоская волна, в масштабе  $\lambda$  практически бесконечная поперек направления распространения.

# ВОЛНОВАЯ ФИЗИКА

Курс лекций для ФМШ

Часть 1. ОПТИКА И ВОЛНЫ

3. ГЕОМЕТРИЧЕСКАЯ ОПТИКА

А. П. Ершов

12 февраля 2007 г.

## Глава 3

# ГЕОМЕТРИЧЕСКАЯ ОПТИКА

Хотя свет имеет волновую природу, в быту и при рассмотрении многих оптических приборов обычно можно применять простейшие представления о лучах света, их прямолинейном распространении в однородной среде, отражении и преломлении на границах. Причины работоспособности такого **приближения геометрической оптики** разъяснялись в п. 2.5. В данной главе мы в основном и будем пользоваться этим приближением. Выход за его рамки понадобится только при рассмотрении предельных возможностей оптики.

### 3.1 Линзы, сферические зеркала. Построение изображений

Хорошо бы иметь устройство, собирающее лучи из целого пучка, выходящего из одной точки, опять в точку. Одна такая система – глаз – позволяет нам видеть. Точки предметов отображаются на заднюю стенку глаза – сетчатку. Из опыта известно, что годится еще выпуклая линза.

Пусть источник света расположен на расстоянии  $a$  от линзы, а изображение – на расстоянии  $b$  с другой стороны (рис. 3.1). Длина пути по прямой  $a + b$ , а по ломаной высоты  $x$  она больше:  $\sqrt{a^2 + x^2} + \sqrt{b^2 + x^2}$ . Для малого  $x$ , то есть тонкого пучка, это преобразуется в  $a + b + x^2/2a + x^2/2b$ , что явно длиннее.

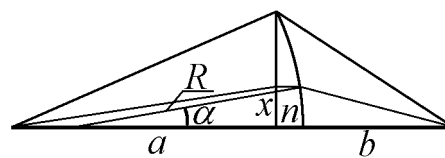


Рис. 3.1.

По принципу Ферма, для схождения лучей в точку надо обеспечить одинаковое время для всех путей. Поэтому короткие пути надо сделать более медленными. Линза, в стекле которой свет замедляется, толще в середине, и можно ожидать компенсации.

Берем плоско-выпуклую линзу из стекла с показателем преломления  $n$  и радиусом кривизны одной поверхности  $R$ . Если максимальная толщина линзы  $h_0$ , то на высоте  $x$  будет меньшее значение  $h$ ; разница  $h_0 - h = R(1 - \cos \alpha) = R\alpha^2/2 = x^2/2R$ . Угол

$\alpha = x/R$  проведен из центра кривизны и тоже мал. Получаем  $h = h_0 - x^2/2R$ .

Для упрощения записи вместо времени прохождения светом по любому пути (т.е. суммы геометрических расстояний, деленных на  $c$  или  $(c/n)$ ), принято употреблять **оптическую** длину пути, у которой «медленные» участки внутри стекла берутся с коэффициентом  $n$ . При малых углах путь внутри линзы можно считать горизонтальным. Оптическая длина пути с высотой  $x$  будет

$$\left[ a + \frac{x^2}{2a} \right] + \left[ b + \frac{x^2}{2b} - \left( h_0 - \frac{x^2}{2R} \right) \right] + n \left( h_0 - \frac{x^2}{2R} \right) = a + b + (n-1)h_0 + \frac{x^2}{2} \left( \frac{1}{a} + \frac{1}{b} - \frac{n-1}{R} \right).$$

В левой части вначале записаны пути «по воздуху», а затем – внутри линзы. Полученное выражение не зависит от  $x$  при условии

$$\frac{1}{a} + \frac{1}{b} = \frac{1}{f},$$

где  $f = R/(n-1)$  называется фокусным расстоянием. Это и есть **формула линзы**. Параллельный пучок, для которого  $a = \infty$ , как раз собирается на фокусном расстоянии от линзы. Применяется и другая форма записи формулы линзы:

$$xx' = f^2,$$

где  $x$  и  $x'$  – расстояния от предмета до фокуса и от изображения до другого фокуса. Она легко получается из основной формулы.

Если линза имеет два радиуса кривизны, то

$$\frac{1}{f} = (n-1) \cdot \left( \frac{1}{R_1} + \frac{1}{R_2} \right).$$

У собирающей линзы один из радиусов может быть отрицательным (выпукло-вогнутая линза), однако вогнутая поверхность должна быть менее кривой.

Изображение наблюдается на некотором экране, положение которого  $b$  определяется из формулы линзы по данным  $a$  и  $f$ . Если сдвинуть экран, каждая точка источника перейдет не в точку, а в кружок. Изображение тогда будет размытым («не в фокусе»), примерно как на сетчатке у людей с дефектами зрения (подробнее см. п. 3.4).

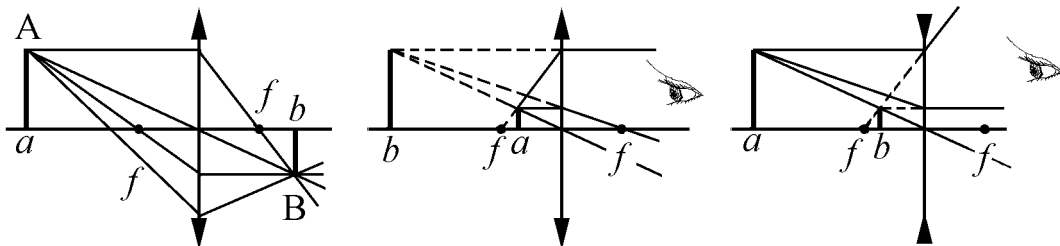


Рис. 3.2.

Для геометрического построения изображений полезны параллельные оси лучи (они после линзы проходят через фокус), проходящие через фокус (становятся параллельными), а также проходящие через центр (так и идут). Пересечение двух лучей, исходящих



из одной точки предмета (А), даст изображение этой точки В (рис. 3.2, слева). Отметим, что там же пересекаются **все** лучи, вышедшие из А и прошедшие линзу. Изображение предмета находится в плоскости, положение которой (b) можно определять из формулы линзы. Геометрическое построение дает независимый способ найти плоскость изображения; кроме того, оно еще указывает и его размер (увеличение). **Действительное** изображение, которое можно видеть на экране, перевернутое, в том числе и у нас на сетчатке. Мы, однако же, воспринимаем окружающую среду в правильном положении (обыкновенно люди видятся ногами вниз, головой вверх) благодаря обработке изображения в мозгу. Обработка эта не сводится к простой инверсии, см. далее п. 3.5.

*Задача:* Доказать, что прямой луч, проходящий из А через центр линзы, попадет в точку В, определенную как пересечение лучей, прошедших через левый и правый фокусы.

При  $a < f$  выходит отрицательное  $b$ . Построение показывает, что лучи нигде не пересекаются, экран ставить некуда. Однако пересекаются продолжения лучей с той же стороны линзы, где находится предмет (рис. 3.2, центр). Глаз наблюдателя, смотрящего через линзу, воспринимает расходящиеся лучи как вышедшие из этой точки пересечения. Предмет видится увеличенным и висящим на некотором удалении за линзой. Такое изображение называется **мнимым**<sup>1</sup>; оно не перевернутое.

Если оба радиуса отрицательны (двояковогнутая линза) или хотя бы один, но он меньше по модулю, то фокусное расстояние отрицательно. Это рассеивающая линза, она не собирает параллельный пучок, а делает из него расходящийся. Изображения в такой линзе мнимые и уменьшенные; они также наблюдаются только глазом. Этот случай показан на рис. 3.2 справа.

Действительное изображение тоже можно наблюдать глазом без всякого экрана. Если собирающую линзу отодвинуть заметно дальше фокусного расстояния, то, например, лектор увидит студентов сильно уменьшенными и висящими вниз головой где-то между ним и линзой (что ясно из левой картинки на рис. 3.2).

Все изложенное применимо к **тонким** линзам. Поэтому и не возникал вопрос, откуда отмерять  $a$  и  $b$ .

Для вогнутого **зеркала** формула такая же, но расстояния  $a$  и  $b$  – с одной стороны:

$$\frac{1}{a} + \frac{1}{b} = \frac{1}{f},$$

где фокусное расстояние  $f = R/2$ , что легко получить отражением луча, идущего под малым углом. Разумеется, зеркало должно быть тонким: сегмент в виде полусферы весь свет в точку не собирает (сферическая аберрация). Зато весь свет собирает параболическое зеркало. Пусть уравнение параболы  $y = x^2/4f$ ; путь луча, падающего с

<sup>1</sup>Потому что без специальных мер оно не видно на каком-либо экране. На сетчатке глаза будет, конечно, действительное изображение, благодаря дополнительному преломлению на входе в глаз. Функция глаза в том и состоит, чтобы сводить расходящиеся пучки от любых предметов на сетчатку. Разумеется, не обязателен именно глаз (с его недостатками как прибора субъективного): мнимое изображение можно вывести на экран с помощью любой собирающей линзы и, в частности, сфотографировать.

высоты  $H$  на расстоянии  $x$  от оси  $y$ , и отражающегося в фокус ( $x = 0$ ,  $y = f$ ) будет  $H - x^2/4f + \sqrt{x^2 + (x^2/4f - f)^2} = H + f$  независимо от  $x$ . По принципу Ферма, это и означает фокусировку.

*Задача:* убедитесь в том же, исходя из закона отражения.

Аналогично рассеивающей линзе, выпуклое зеркало дает мнимое прямое изображение, что можно наблюдать в елочных шарах, блестящих чайниках и т.п.

Зеркало свободно от хроматической аберрации (см. п. 3.3), так как все цвета отражаются одинаково. Поэтому наиболее крупные телескопы – зеркальные.

Для эллиптического зеркала все пути из одного фокуса в другой точно одинаковы. Отсюда видно, что принцип Ферма не минимальный, а экстремальный. Если провести касательную прямую к эллипсу, то в точке касания для отражения от этой прямой путь минимальный (рис. 3.3). Если же провести касательную окружности большей кривизны, чем у эллипса, то в точке отражения путь максимальный.

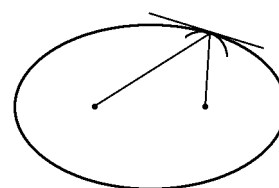


Рис. 3.3.

## 3.2 Фотометрия\*

Эта область оптики, изучающая энергетические характеристики света, изобилует новыми названиями. **Световой поток**  $\Phi$  – это попросту мощность, или энергия/сек, Вт в системе СИ, испускаемая источником. То же самое, но определяемое по световому ощущению (в виде света), измеряется в специальных единицах – **люменах**. Коэффициент перевода: для стандартной зеленой волны  $5550 \text{ \AA}$  (555 нм):  $1 \text{ Вт} = 680 \text{ лм}$ , или  $1 \text{ лм} = 0,0016 \text{ Вт}$ . При этой длине волны глаз наиболее чувствителен. Таблица видности ( $\lambda$  в нанометрах,  $K$  – относительная чувствительность в процентах):

$\lambda$	420	440	460	480	500	520	540	560	580	600	620	640	680
$K$	0,4	2,3	6	13,9	32,3	71	95,4	99,5	87	63,1	38,1	17,5	6,1

Кривая видности устанавливается усреднением индивидуальных ощущений, но когда она зафиксирована, можно сделать искусственный глаз на основе фотоэлементов и светофильтров, вполне объективный.

**Сила света**  $I = \Phi/\Omega$  – световой поток на единицу телесного угла, размерность Вт/стер, а по световому ощущению единица кандела = лм/стер (бывшая свеча, ср. канделябр). Эта единица основная, для нее есть эталон:  $1 \text{ кд} = 1/60$  силы света, излучаемой в направлении нормали абсолютно черным телом площади  $1 \text{ см}^2$  при  $2042,5 \text{ К}$  (температура затвердевания платины). Такой эталон представляет собой малое отверстие, через которое излучение выходит из нагретой полости. Практически получается точечный источник, почему и удобно брать излучение в телесный угол.

Вообще  $1 \text{ см}^2$  излучает  $\sigma T^4$  в полусферу, а на телесный угол 1 стер в среднем придется доля  $1/2\pi$ . Но по нормали излучается больше, и получается  $\sigma T^4/\pi = 31,4 \text{ Вт/стер}$ . Поскольку

ку это 60 кд, то выходило бы, что 1 кд соответствует 0,5235 Вт/стер, или  $0,5235 \cdot 680 = 356$  лм/стер, если бы мы видели все это излучение с эффективностью 1. Реально же при температуре эталона большая часть энергии приходится на инфракрасный диапазон, и видна только малая доля излучения ( $\cong 1/356$ ), откуда и получается цифра 1 кд = 1 лм/стер.

Лампа 100 Вт излучает  $100/4\pi \approx 8$  Вт/стер. Будь это зеленый свет, вышло бы  $8 \cdot 680 \approx 5500$  кд, но такое количество свечей явно завышено. На самом деле светоотдача ламп накаливания 10 – 20 лм/Вт, или  $\sim 1$  кд/Вт. Световой КПД порядка  $2 \div 3$  %. Заметно лучше у люминесцентных ламп «дневного света», излучающих в более видимом диапазоне (десятки процентов). Поскольку фотометрия – это стык физики и физиологии, мало кому охота в этом разбираться. Даже в эталоне уцелел  $1 \text{ см}^2$ , а на лампочках пишут не канделы, а ватты.

**Освещенность**  $E = \Phi/S$  – энергия, опять-таки по световому ощущению, падающая на единицу поверхности. Размерность лм/м<sup>2</sup> = люкс. Для точечного источника с силой света  $I$ :  $\Phi = 4\pi I$ ,  $E = I/R^2$  на расстоянии  $R$ . Освещенность в солнечный день  $10^5$  лк, требуемая для тонких работ на рабочем столе 100 лк. Следовательно, надо лампу накаливания около 100 Вт на расстоянии 1 м. Если поверхность не нормальна к свету,  $E = I \cos \theta / R^2$ . Наклонное падение солнечного света зимой уменьшает поток энергии и вместе с ростом отражения позволяет снегу накапливаться.

**Яркость**  $B = I/S$  – сила света с единицы излучающей поверхности. 1 кд с  $1 \text{ м}^2$  – это будет нит, а с  $1 \text{ см}^2$  – стильб (сб). Под углом следует брать «видимую» поверхность – проекцию на направление наблюдения. Яркость черного тела не зависит от наклона поверхности, так как сила света и видимая поверхность пропорциональны  $\cos \theta$ . Яркость Солнца в зените  $16,5 \cdot 10^4$  сб, волоска лампы накаливания  $\approx 1000$  сб, люминесцентные лампы и пламя свечи дают  $\approx 0,5$ , полная Луна и дневное небо  $\approx 0,25$  сб.

### 3.3 Дисперсия, спектры

В п. 2.4 мы видели, что показатель преломления может зависеть от частоты. Изменение  $n$ , то есть скорости волны, называется **дисперсией**. В области прозрачности эта зависимость обычно слабая. Например, для воды  $n_F = 1,337$ ,  $n_D = 1,333$ ,  $n_C = 1,331$ . Обозначения – по Фраунгоферу:  $F$  – 486,1 нм, голубая линия Н,  $D$  – 589,3 нм, желтая линия Na,  $C$  – 656,3 нм, красная линия H. В веществе  $\lambda$  меняется, и правильнее было бы указывать частоту, но исторически принято употреблять длины волн в вакууме (при  $n = 1$ ). Линия  $D$  находится в середине видимого диапазона, к ней обычно и относится справочный показатель преломления. Как количественная мера дисперсии, используется разность  $n_F - n_C$  (средняя дисперсия). Например, для стекол она может быть 0,0213 (тяжелый флинт,  $n_D = 1,65$ ), и 0,0085 (легкий крон,  $n_D = 1,517$ ), то есть в пределах видимого света  $n$  может изменяться примерно на процент.

Дисперсия света хорошо видна на призме. Угол отклонения луча порядка (угол призмы)  $\cdot (n - 1)$  и тоже меняется с  $\lambda$ . В быту красивые радужные блики дают граненое стекло и хрусталь. Блеск драгоценных камней прямо связан с показателем преломления

$n$ . Действительно, коэффициент отражения растет с  $n$ ; кроме того, при большом  $n$  шире диапазон углов полного внутреннего отражения.

По величине  $n$  алмаз, имеющий  $n_D = 2,42$ , превосходит все камни (рубин – 1,76, изумруд – 1,58). Благодаря специальной огранке, выработанной ювелирами в течение веков, почти весь свет, вошедший в камень в лицевой стороны, испытывает полное отражение от невидимых задних граней и выходит опять в переднюю полусферу (рис. 3.4). У алмаза дисперсия почти максимальная:  $\Delta n = 0,044$ . От дисперсии зависит «игра» камня, то есть его способность переливаться разными цветами, разлагая свет. Алмаз виден испускающим цветные лучи большому числу приглашенных на бал, нежели кусок стекла той же формы с  $\Delta n = 0,015$  (в отношении  $44/15 \approx 2,9$ ). У рубина  $\Delta n = 0,018$ , и такое отношение будет  $44/18 \approx 2,44$ . Рубин, однако, ценится еще и за красивый натуральный цвет. У искусственного камня фианита<sup>2</sup> (кубическая окись циркония)  $\Delta n = 0,06$  – даже больше, чем у алмаза, а  $n = 2,15 \div 2,18$ . На вид фианит практически неотличим от алмаза, пока не поцарапать его заведомым алмазом, к тому же легко получать цветные камни, если варить его с добавками<sup>3</sup>.

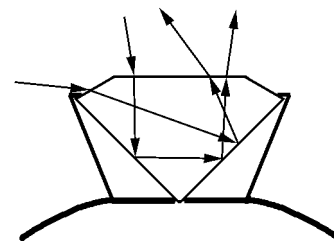


Рис. 3.4.

Пропуская луч через призму, получаем спектр в виде разноцветных полос. Если сильнее преломляются короткие волны (большие частоты), это называется нормальной дисперсией; в редких обратных случаях говорят об аномальной дисперсии. Последовательность цветов видимого света:

$\lambda$	380÷450	450÷480	480÷510	510÷562	563÷585	585÷620	620÷760
цвет	Ф	С	Г	З	Ж	О	К

В сторону коротких волн далее идет УФ, рентген, гамма – диапазоны, границы между которыми несколько расплывчаты. В сторону длинных волн – ИК, миллиметровые, дециметровые, УКВ, КВ, СВ, ДВ и далее сверхдлинные. С обеих сторон нет естественных пределов<sup>4</sup>. Линза тоже разлагает свет – фокусы для разных волн будут в разных местах (хроматическая аберрация). В составных объективах фотоаппаратов она компенсирована применением разных сортов стекла.

<sup>2</sup>Название дано в честь Физического Института АН.

<sup>3</sup>Первым удачным заменителем алмаза был иттриево-алюминиевый гранат (ИАГ), имеющий  $n_D = 1,83$ ,  $\Delta n = 0,026$  и приближающийся по твердости к алмазу. Успеху продукта способствовал удачный опыт Элизабет Тейлор, которой Ричард Бартон в 1969 г. подарил алмазный кулон в 69,42 карата, стоимостью более 1 млн долларов. Страховка при ношении натурального камня обходилась Э. Тейлор в 1000 долларов за вечер, а стоимость ИАГ-копии была только 3500 долларов. Камни можно было отличить, если они находились рядом, но только эксперт мог определить синтетический камень, если рассматривать его отдельно. См.: Д. Элуэлл. Искусственные драгоценные камни. М.: Мир, 1986.

<sup>4</sup>Если не говорить о таких, не очень естественных, как размер Вселенной или ее видимой части, или равенство энергии кванта  $\hbar\omega = 2\pi\hbar c/\lambda$  характерной энергии Вселенной.

Спектры, которые мы видим в домашних условиях, непрерывные: в них присутствуют все волны. Такой спектр дают **оптически толстые** тела типа Солнца, нити накаливания лампочки и т.п. Излучение находится в равновесии с веществом; предельный случай полного равновесия – абсолютно черное тело. Наоборот, **оптически тонкие** тела, в которых длина пробега света больше размера, могут излучать в отдельных ярких линиях, соответствующих переходам между атомными и молекулярными уровнями энергии. Если посмотреть через призму на лампу накаливания, видна радужная расплывшаяся нить лампы. Взглянув же на ртутный уличный фонарь, увидим два фонаря, желтый и зеленый, расположенные рядом (ртуть среди прочих излучает яркую линию 546зел и дублет 577&579желт). Атомные переходы между возбужденными электронными состояниями дают серии, называемые линейчатыми спектрами; в молекулах наблюдается большое количество близко лежащих колебательных и вращательных переходов (полосатые спектры). Плазма дает и непрерывный спектр от свободных электронов.

Каждое вещество имеет свой характерный спектр, чем пользуются при анализе. Само собой, поглощение в таких спектральных линиях гораздо больше, чем вне их. Поэтому из непрерывного спектра Солнца, проходящего через солнечную атмосферу, вырезаются темные линии поглощения (фраунгоферовы). Можно найти, из чего состоят звезды. Например, на Солнце был в 1868 г. открыт новый элемент – гелий (Ж. Жансен, Н. Локьер), на Земле выделенный только в 1895 г. (У. Рамзай).

На некоторых звездах обнаружен искусственный элемент технеций, не имеющий стабильных изотопов (максимальный период полураспада 4,2 млн лет<sup>5</sup>). Можно думать, что это сигнал братьев по разуму. В лабораториях используют для разложения в спектр монохроматоры с призмами либо дифракционными решетками. Много спектральных приборов применяют химики (УФ, ИК, ЯМР).

Радуга, разумеется, тоже связана с дисперсией. Пусть горизонтальный солнечный луч попадает в каплю воды в точке, характеризуемой углом  $\varphi$  к нормали. Преломленный луч пойдет под углом

$r$ ,  $\sin r = \sin \varphi / n$ . Дойдя до поверхности, он частично отразится опять внутрь и снова подойдет к поверхности под тем же углом  $r$ , после чего выйдет снова под углом  $\varphi$  к радиусу. Угол поворота луча:  $\psi = \pi + 2\varphi - 4r$ ; угол луча с горизонтом (направление, откуда он придет к вам в глаз) будет  $\theta = 4 \arcsin(\sin \varphi / n) - 2\varphi$ .

При малом  $\varphi$  и показателе преломления воды  $n = 4/3$   $\theta \approx \varphi$ . При максимальном  $\varphi = \pi/2$   $\theta = 4 \arcsin(3/4) - \pi = 0,25$ , причем вблизи этого значения убывает. Максимум  $\theta$  где-то посередине, а именно при  $\sin \varphi_0 = \sqrt{(4 - n^2)/3} = 0,86$ , что соответствует прицельному углу  $\varphi_0 = 1,04 = 59^\circ$ . Угол прихода  $\theta_0 = 4r_0 - 2\varphi_0 = 0,734 = 42^\circ$ .

<sup>5</sup>Для <sup>97</sup>Tc: Физическая энциклопедия, т. 5, стр. 112; в Химической энциклопедии (т. 4, стр. 560) приводится цифра 2,6 млн лет.

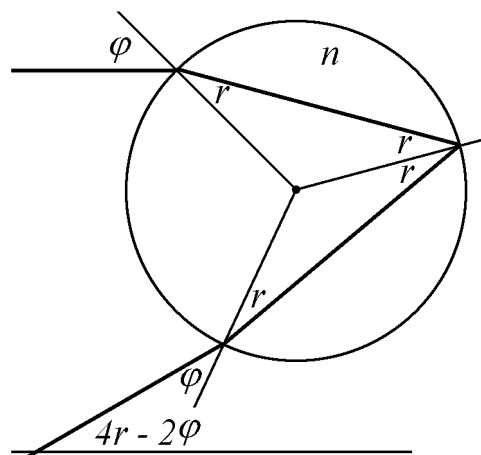


Рис. 3.5.

После дождя, когда в воздухе много капель воды, под этим углом будет наблюдаться повышенный уровень рассеянного излучения. Как всегда вблизи максимума, близкие по прицельному углу  $\varphi \approx \varphi_0$  лучи будут выходить практически под таким же углом  $\theta_0$  и, таким образом, сгущаться. Для прочих прицельных углов такого увеличения яркости не будет. Такое рассеянное излучение, именно по дуге с углом  $42^\circ$  впереди себя, если стоять спиной к Солнцу, мы и видим в подходящую погоду. Из-за дисперсии для разных цветов углы  $\theta_0$  будут немного отличаться и вместо просто яркой полосы «солнечного» цвета на небе мы видим ряд полос – радугу. При хорошем свете бывает видна вторая, менее яркая, радуга, вызванная двукратным отражением в капле.

### 3.4 Оптические приборы

Главный прибор – это глаз. Поскольку он не совсем похож на линзу, рассмотрим преломление в нем подробнее. Разница в том, что экран – сетчатка – тоже находится в преломляющей среде. Для разнообразия применим не принцип Ферма, а закон преломления на сферической поверхности: угол падения  $i = x/R + x/a$ , угол преломления  $r = i/n = x/R - x/b$ , откуда «формула глаза»

$$\frac{1}{a} + \frac{n}{b} = \frac{n-1}{R}.$$

При  $a = \infty$   $b = nR/(n-1)$ . Поскольку это порядка  $2R$ , получается  $n \approx 2$ . На самом деле радиус, входящий в фокусное расстояние, меньше, так как спереди у глаза выступает хрусталик, имеющий бóльшую кривизну. Соответственно  $n$  получается завышенным. Но мы знаем, что  $n$  больше, чем у воды (иначе в воде глаз совсем бы не работал). Преломляющая, или оптическая сила глаза  $1/f = (n-1)/R \approx 50$  обратных метров (диоптрий). К другим расстояниям глаз приспособляется, изменяя кривизну передней поверхности (хрусталика). Для этого нужно не так много: при  $a = 10$  см надо увеличить оптическую силу на 10 дптр, т.е. внести малую поправку к основной форме глаза. Расстоянием наилучшего зрения считается 25 см, при  $R \approx 2$  см практически бесконечность ( $1/0,25$  увеличивает  $1/f$  на  $\approx 10\%$ ). С одной стороны, чем ближе, тем больше размер изображения на сетчатке, больше деталей (пикселей) удастся различить. Но на расстояниях меньше 25 см средний глаз будет испытывать напряжение при фокусировке, а ближе некоторого расстояния вообще откажется фокусировать.

В небольших пределах исправляют зрение очки ( $\pm 50$  не бывает). Считаем, что линза с фокусным расстоянием  $f_1$  находится близко к глазу. Она строит изображение в точке  $c$ :  $1/a + 1/c = 1/f_1$ . Для глаза это изображение в точке  $-c$ : глаз видит лучи, сходящиеся где-то в затылок;  $-1/c + n/b = 1/f$ . Складывая уравнения, имеем  $1/a + n/b = 1/f_1 + 1/f$ , откуда виден смысл оптических сил  $1/f$ : для близких линз они складываются.

Близорукий глаз не может настроиться на удаленные предметы. Параллельный пучок из бесконечности сводится таким глазом в точку впереди сетчатки, а на сетчатке



вместо точки получится кружок. Пусть человек с таким дефектом может увидеть резко предмет, находящийся на максимальном расстоянии  $a$  от глаза. Можно написать  $1/a + n/b = 1/f$ ; либо  $b$  великовато, либо  $f$  мало. Исправим дефект линзой с фокусным расстоянием  $F$ :  $1/\infty + n/b = 1/f + 1/F$ . Лучи из бесконечности сведутся в точку, если  $1/F = -1/a$ , то есть  $F = -a$  (рассеивающая линза). Например, человек, видящий не дальше 20 см = 0,2 м, нуждается в очках  $(-1/0,2) = -5$ . Сложные манипуляции окулистов при подборе очков представляют собой род практической магии, рассчитанной на малограмотных, и дают примерно такую же точность ( $\pm 0,5$  дптр).

Дальнорядный глаз не фокусируется на близкие предметы: его оптическая сила, наоборот, мала. Бывает, что при дальнорядности не фокусируются предметы ни на каких расстояниях: даже параллельный пучок сводится за сетчатку. Поэтому такое же простое измерение, как для близорядного глаза, не всегда осуществимо. Однако можно вооружить глаз достаточно сильной собирающей линзой, которая сведет задачу к предыдущему случаю (близорядкости). Полученные «отрицательные» диоптрии надо будет вычесть из оптической силы этой пробной линзы. Зато фокусное расстояние собирающей линзы гораздо проще находится экспериментально – с помощью удаленного источника (например, Солнца) и линейки.

Следующий по распространенности прибор – лупа, применяемая для «увеличения» мелких объектов. Предмет располагается ближе фокуса линзы. Линза создает прямое мнимое изображение, которое глаз видит «висящим» за линзой, на расстоянии  $|b| = af/(f - a)$  от нее. При размере предмета (или его мелкой детали)  $X$  видимый размер изображения  $Y = X |b| / a$ . Ясно, что чем ближе предмет к фокусу, тем крупнее (но и дальше) изображение. Обычно изображение помещают на бесконечность, тогда  $a \rightarrow f$ . Следует сравнивать угловой размер изображения  $Y/b$  с размером  $X/L_{25}$ , под которым смотрел бы невооруженный глаз на расстоянии наилучшего зрения  $L_{25} = 25$  см. Из подобия треугольников с катетами  $Y$  и  $X$  (рис. 3.6) получаем отношение

$$\frac{Y}{b} = \frac{X}{a} \approx \frac{X}{f}, \quad \text{откуда увеличение} \quad U = \frac{Y}{b} \bigg/ \frac{X}{L_{25}} = \frac{L_{25}}{f}.$$

Лупа должна иметь по возможности малое фокусное расстояние (большую «силу»). Линза с фокусным расстоянием 2,5 см увеличивает в 10 раз (обозначается  $10\times$ ). Для близорядких, у которых расстояние наилучшего зрения меньше, меньше и толку от лупы.

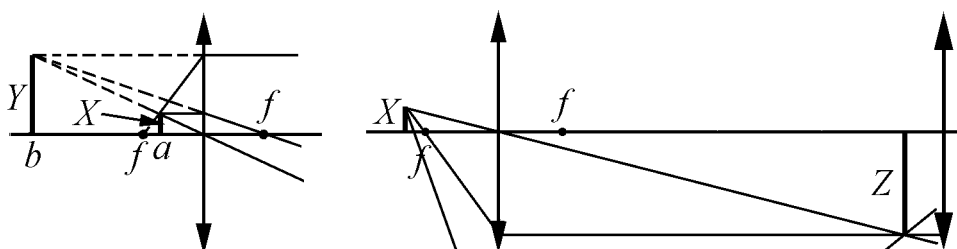


Рис. 3.6.

Левенгук изготовлял однолинзовые микроскопы (попросту лупы) с увеличением в сотни раз, для чего надо  $f \sim 2,5$  мм. Поскольку поперечник линзы должен быть заметно меньше ее радиуса кривизны, ясно, какая требовалась ювелирная работа.

Сейчас имеются микроскопы с увеличением  $\sim 10^3$ . Поскольку одна линза не позволяет этого добиться, не поможет ли еще одна? Очевидный прием – создать изображение одной линзой (объективом), а затем рассматривать его через другую линзу (окуляр).

Пусть предмет размера  $X$  находится чуть дальше фокуса объектива и создает действительное изображение размера  $Z$  на расстоянии  $L$ . Увеличение  $Z/X = L/f_1$ . Далее, окуляр дает изображение с увеличением  $Y/Z = L_{25}/f_2$ , так что в итоге  $Y/X = (LL_{25})/(f_1f_2)$ . При  $L \approx L_{25}$ , что примерно и выполняется для микроскопов, увеличение порядка квадрата величины, характерной для одной отдельно взятой лупы. На рис. 3.6 проведен луч, явно не проходящий через первую линзу. Такие лучи тем не менее можно использовать для построения, воображая, что линза продолжена вверх и вниз на сколько нужно. Этим несколько улучшается наглядность для второй линзы: она «видит» расходящийся из нижней точки изображения пучок, а добавочный луч по крайней мере не идет мимо.

У близко расположенных линз складываются оптические силы. Рассмотрим оптические свойства разнесенных линз. Как ни странно, система из двух линз  $f_1$  и  $f_2$  аналогична одной тонкой линзе, свойства которой зависят от  $f_1$ ,  $f_2$  и расстояния между линзами  $d$ . Единственно, что расстояния  $A$  и  $B$  надо отмерять не от линз, а от некоторых **главных плоскостей**.

Найдем положение главных плоскостей и фокусное расстояние пары линз. Пусть для определенности расстояние  $d < f_{1,2}$ . Пустим слева параллельный пучок. Первая линза собирает его в фокус  $f_1$ . Но стоящая на дороге вторая загнет лучи еще сильнее (рис. 3.7). Для нее «источник» находится на расстоянии  $-(f_1 - d)$ ; знак обозначает правую сторону второй линзы. По формуле  $-1/(f_1 - d) + 1/b = 1/f_2$ .

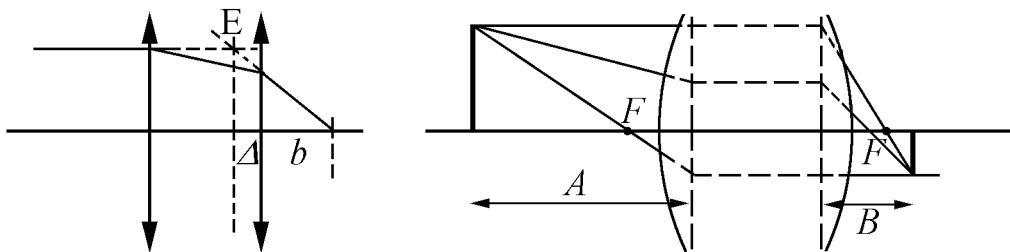


Рис. 3.7.

Получаем положение точки сбора лучей относительно второй линзы  $b = f_2(f_1 - d)/(f_1 + f_2 - d)$ . Теперь продолжим преломленный луч до пересечения с продолжением исходного параллельного оси луча. Расстояние от второй линзы до точки E пересечения  $\Delta_2 = bf_1/(f_1 - d) - b = df_2/(f_1 + f_2 - d)$ . Если пустить параллельный пучок справа, аналогичная точка будет на расстоянии  $\Delta_1 = df_1/(f_1 + f_2 - d)$  от первой линзы.



Вот это и будут главные плоскости: параллельный пучок для внешнего наблюдателя как бы преломляется один раз – в точке  $E$  – и дальше идет напрямиком в фокус. Фокусное расстояние такой системы  $F = b + \Delta = f_1 f_2 / (f_1 + f_2 - d)$  симметрично по отношению к перестановке линз и, значит, не зависит от направления пучка. Сумма  $\Delta_1 + \Delta_2 = d(f_1 + f_2) / (f_1 + f_2 - d) > d$  при принятых ограничениях (когда линзы не слишком далеко), так что между главными плоскостями имеется область «перекрывтия».

Правила построения в системе с различными главными плоскостями (рис. 3.7): параллельный оси луч ведется до главной плоскости, затем прыгает к другой плоскости и дальше ведется через фокус; луч, проходящий через фокус, доходит до главной плоскости, перепрыгивает на другую, сохраняя расстояние до оси, и идет параллельно. Все как в тонкой линзе, только выбрасывается расстояние между главными плоскостями.

Надо еще убедиться, что для составной системы выполняется формула линзы, то есть  $1/A + 1/B = 1/F$ , где  $A = a + \Delta_1$  и  $B = b + \Delta_2$  отмеряются от главных плоскостей. Для этого используем уже известный факт фокусировки параллельного пучка. Если луч на любой высоте  $h$  от оси приходит в фокус, то можно записать равенство оптических путей осевого и пробного луча:

$$\Delta(h) + F + \frac{h^2}{2F} = \Delta(0) + F,$$

где  $\Delta(h)$  – разность хода, которая набирается в системе для пробного луча (по построению мы можем считать, что это происходит между главными плоскостями),  $\Delta(0)$  – для осевого,  $h^2/2F$  – геометрическое удлинение пути пробного луча. Теперь пускаем луч из точки  $A$  и допустим, что он приходит в точку  $B$ :

$$A + \frac{h^2}{2A} + \Delta(h) + \frac{h^2}{2B} + B = A + \Delta(0) + B.$$

Сравнивая с предыдущим уравнением, немедленно получаем

$$\frac{1}{A} + \frac{1}{B} = \frac{1}{F},$$

что и требовалось. Для двух тонких разнесенных линз можно получить то же самое и прямым отображением в линзах по очереди, но вычисление довольно громоздкое.

В существовании главных плоскостей легко убедиться, выкрутив объектив фотоаппарата и фокусируя удаленный объект (типа Солнца). Отмерив от фокусов по 50 мм, увидим, что порядка сантиметра внутри объектива еще остается. Несмотря на внушительные порой размеры, объектив работает как тонкая линза (с учетом раздвижки главных плоскостей)! У одной тонкой линзы обе плоскости совпадают с самой линзой.

Если вернуться к микроскопу, то удобнее букву  $d$  заменить на длину тубуса  $L$ . Аналогичное построение для случая  $L > f_1 + f_2$  дает

$$F = \frac{f_1 f_2}{L - f_1 - f_2}, \quad \Delta_1 = \frac{f_1(L - 2f_2)}{L - f_1 - f_2}, \quad \Delta_2 = \frac{f_2(L - 2f_1)}{L - f_1 - f_2},$$

причем главные плоскости расположены снаружи микроскопа вблизи фокусов окуляра и объектива (при  $f_1 = f_2$  точно в фокусах). Так как обе линзы короткофокусные,  $F \approx f_1 f_2 / L$ , и микроскоп аналогичен лупе с таким фокусным расстоянием.

Интересная комбинация линз – **телескоп**. Простейший **кеплеровский** телескоп состоит из двух линз, фокусы которых совмещены ( $d = f_1 + f_2$ ). Поскольку он применяется для рассмотрения удаленных объектов, играет роль угловое увеличение. Если параллельный пучок света идет под углом  $\varphi$  к оси, он сфокусируется в фокальной плоскости объектива (рис. 3.8).

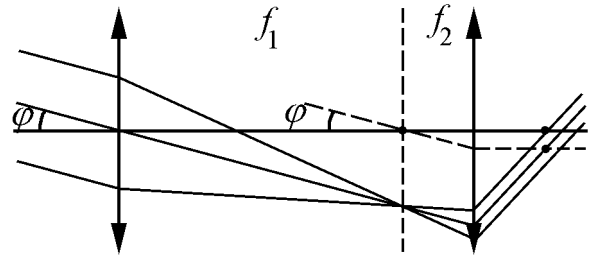


Рис. 3.8.

Дальше лучи расходятся как бы из этой точки, и из второй линзы выходит опять параллельный пучок, но уже под углом  $\varphi f_1 / f_2$ . Действительно, проведем пунктирный (на самом деле несуществующий) луч через общий фокус параллельно центральному лучу пучка. Этот вспомогательный луч после окуляра идет параллельно оси. В задней фокальной плоскости окуляра он должен пересекаться с преломленным центральным лучом, что и помогает направить выходящий пучок.

Делая окуляр короткофокусным, а объектив длиннофокусным, получаем большое угловое увеличение  $f_1 / f_2$ . Например, угол  $\varphi$  может быть между двумя звездами, или деталями на поверхности Луны.

Такой телескоп переворачивает изображение, что неудобно почти всегда (например, в военном или морском деле). Исправить этот эффект можно, взяв третью линзу с фокусным расстоянием  $f_3$  и увеличив длину телескопа на  $4f_3$ . Третья линза будет «связывать» фокусы окуляра и объектива, находясь на  $2f_3$  от каждого. Поскольку увеличение для расстояния  $2f_3$  равно 1, она будет только переворачивать изображение еще раз, устраняя неудобство простого телескопа. Получается очень длинная подзорная труба. В биноклях изображение переворачивается двумя призмами, отчего на них заметно характерное «колени». Объективы наиболее крупных телескопов (до 5 – 6 м в диаметре) делаются зеркальными, так как большие стеклянные линзы деформируются под собственной тяжестью.

**Галилеевский** телескоп состоит из длиннофокусного объектива и рассеивающей линзы – окуляра, помещенной на расстоянии  $f_1 - |f_2|$  (рис. 3.9). Объектив пытается собрать параллельный пучок в фокальной плоскости, общей для обеих линз, а окуляр делает опять параллельный пучок, но идущий уже под углом к оси, увеличенным в  $f_1 / |f_2|$  раз. Здесь вспомогательный луч проведен через центр окуляра и должен пересекаться с продолжением центрального луча в передней фокальной плос-

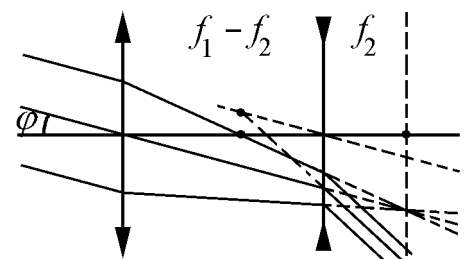


Рис. 3.9.

кости окуляра. Изображение не перевернутое; третья линза или призмы не нужны. Поскольку к тому же фокусные расстояния не складываются, а вычитаются, то система получается короче и используется, например, в театральные биноклях и в других случаях, когда компактность важнее качества изображения.

Увеличение бинокля или зрительной трубы примерно равно отношению диаметров объектива и окуляра, поскольку весь входящий пучок должен попасть в глаз.

**Фотография.** По существу, фотоаппарат отличается от глаза лишь техническим исполнением. Но только в последнее время удалось создать цифровые фотоаппараты, не уступающие глазу по качеству изображения и записывающие изображение примерно в такой же форме (набор электрических потенциалов в точках дискретной матрицы). Одновременно объективы стали, подобно глазу, изменять фокусное расстояние (вместо того чтобы, как в старину, только передвигаться как целое относительно пленки). Еще недавно даже фотограф-любитель владел массой технических знаний. Сейчас же каждый может навести фотоаппарат, нажать кнопку и получить приемлемый результат; искусство фотографии все больше смещается в сферу выбора кадра.

**Голография** получила развитие с появлением мощных источников когерентного света – лазеров. Пусть фотопластинка освещается одновременно пучком света (так называемым опорным), исходящим из лазера, и более слабым светом того же лазера, рассеянным на каком-то предмете (например, Вашем лице). На пластинке оба пучка дадут интерференционную картину, а после проявления получится соответствующая картина почернения. Где волны складывались, пластинка почернеет сильнее, где вычитались – слабее. Фазовые различия пучков запомнятся в виде причудливого узора.

Если теперь направить на проявленную пластинку один только опорный пучок, в самом грубом приближении получится ослабленный проходящий пучок, соответствующий среднему почернению. Это тривиально. Но оказывается, будет эффект и от структуры пластинки. Свет будет проходить лучше в светлых местах и слабее – в темных. Отклонения от среднего фона можно описать как распределение источников, находящихся в одной фазе, но имеющих разную амплитуду, по поверхности пластинки. По существу, получается дифракционная решетка.

Оказывается, в результате взаимодействия пучка с пластинкой появятся еще два пучка: один дает мнимое изображение предмета в том месте, где он был при съемке, а другой – действительное изображение предмета в симметричном относительно пластинки месте. Рассматривая точечный предмет, нетрудно убедиться, что для этих двух симметричных точек выполняется условие одинаковой разности хода лучей. Протяженные предметы можно считать наборами точек. Изображение получается объемным, и его можно, в определенных пределах, разглядывать с боков. Этим и объясняется название техники (ср. whole – весь, целый, полный). Интересно, что даже небольшая часть пластинки даст то же изображение, правда с некоторой потерей качества. Как мы уже видели, дифракционная решетка отклоняет лучи по тем же направлениям независимо от размера, но расхождение этих пучков и, значит, резкость дифракционной картины обратно пропорциональны размеру решетки.

**Томография** стала актуальной в последние годы. Тела, непрозрачные для видимого света, можно просвечивать рентгеновскими лучами. Старые способы рентгенографирования давали

плоский отпечаток на пленке. В каждой точке снимка запоминается интегральное поглощение вдоль луча, проведенного от источника к этой точке. Если же провести просвечивание с нескольких направлений и/или точек, становится возможно восстановить пространственную структуру тела. При этом необходимо решить большое количество интегральных уравнений, что стало доступным благодаря развитию компьютеров и математического обеспечения. Обычно результатом исследования является представление объекта в виде набора слоев, откуда и идет название метода (греч. *τομοσ* – сечение, слой, ломоть; микротом – прибор для шинкования биологических образцов). В медицине рентгеновская и ЯМР-томография позволяют обнаружить неоднородности типа раковых опухолей, инородные тела и пр., не разрезая пациента. Вероятно, почти каждому, кто читает этот текст, предстоит хотя бы раз пройти томографическое обследование.

### 3.5 Пределы увеличения. Условие Аббе. Возможности оптики

Выясним ограничения на увеличение микроскопа. В формулу увеличения  $Y/X = (LL_{25})/(f_1 f_2)$  входит длина тубуса  $L$ ; при тубусе длиной 10 м и фокусных расстояниях объектива и окуляра 1 см увеличение будет 25000. Еще имеется некоторая свобода для роста увеличения.

Поскольку для большого увеличения предмет следует располагать близко к фокусу, а фокусное расстояние делать по возможности малым, надо рассмотреть особенности изображения широкими пучками лучей. Для источника, находящегося на оси, желательно, чтобы все лучи, а не только близкие к оси, сходились в точке изображения. Ясно, что зеркало или линза со сферическими поверхностями таким свойством не обладает – имеется **сферическая аберрация**. Особенно очевидна она для зеркала в виде полусферы. Для зеркальной фокусировки параллельного пучка подходящая поверхность – параболоид, а для двух точек – эллипсоид вращения, у которого источник и изображение являются фокусами. У линз тоже можно компенсировать сферическую аберрацию, немного регулируя кривизну поверхности.

Но сферическая аберрация – это не единственная неприятность. Для неискаженности изображения все точки предмета, или хотя бы близкие к оси, должны изображаться резко любыми лучами, в том числе сильно наклонными. Другими словами, все части оптической системы должны давать одинаковое увеличение. Возьмем «предмет» в виде малого отверстия в экране, с центром на оси линзы, освещаемого слева параллельным пучком. За линзой получится перевернутое изображение

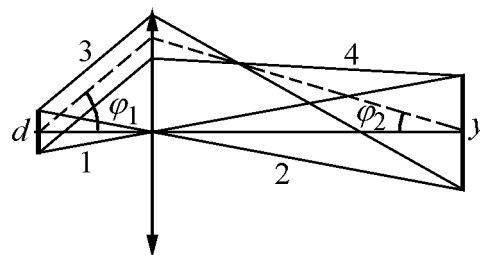


Рис. 3.10.

(рис. 3.10). Лучи 1 и 2, прошедшие от краев предмета через центр линзы, проходят до точек изображения одинаковые оптические пути (не потому, что они интерферируют между собой – они идут в разные точки изображения – а просто из симметрии).

По принципу Ферма, такими же должны быть оптические пути всех лучей, попадающих в концы изображения. Возьмем пару лучей от краев источника, идущих под углом  $\varphi_1$  к оси и проходящих, возможно, вблизи края линзы. Перед линзой у них будет разность хода  $d \sin \varphi_1$  (нижний луч 4 длиннее). После линзы лучи пересекаются в фокальной плоскости и приходят на изображение с разностью хода  $y \sin \varphi_2$ , но уже длиннее луч 3. (Считаем пучок тонким, так что углы лучей 3 и 4 к оси оба практически равны  $\varphi_2$ ). Линза сама не дает разности хода, так что эти отклонения должны взаимно компенсироваться. Путь 3 должен интерферировать с путем 2, а путь 4 – с путем 1:

$$d \sin \varphi_1 = y \sin \varphi_2$$

(условие синусов Аббё). Оно должно выполняться для всех углов, в том числе максимальных.

Для простейшей геометрической оптики с близкими к оси лучами, то есть когда линзы не только тонкие, но и маленькие, синусы можно заменить углами или тангенсами. Тогда условие переходит в  $dh/a = yh/b$ , где  $h$  – высота прохода луча через линзу. Получаем известную формулу увеличения:  $d/a = y/b$ , которая выполняется автоматически, для всех положений предмета на оси. Условие Аббе для широких пучков уже нетривиально и выполняется в лучшем случае для пары точек на оси.

Пример (из книги: М. Борн, Э. Вольф. Основы оптики): тонкая линза с фокусом  $f$  из стекла с  $n = 1,5$ , если мы хотим изобразить предмет, находящийся в бесконечности, должна иметь радиусы кривизны  $5f$  и  $5f/9$ , причем более выпуклой стороной обращена к предмету. Сферическая aberrация устраняется теперь небольшими поправками к форме поверхности (кривизна несколько уменьшается с удалением от оси). Уравнение поверхности сферы вблизи полюса  $x = r^2/2R + r^4/8R^3$ , а надо взять в данном случае  $x = r^2/2R + 0,26r^4/8R^3$ .

Бывает, что вообще не удастся убрать одновременно сферическую aberrацию и выполнить условие Аббе. Например, зеркало без сферической aberrации – это эллипсоид. Но эллипсоид с очевидностью не дает одинаковых увеличений для всех углов: один его полюс делает изображение меньше предмета, а другой – больше.

Всего при рассмотрении первых поправок к простейшей оптике возникает шесть видов aberrаций. Кроме сферической, имеется кома (исчезающая при выполнении условия синусов), астигматизм, кривизна поля изображения, дисторсия и хроматическая aberrация. Поэтому объективы делаются многолинзовыми, причем исправление одной aberrации может увеличить другие. Остается удивляться хорошему качеству изображения фотоаппаратов (и телевизоров, с их электронной оптикой). Глаз почти ни на что не исправлен, и неискаженное изображение получается пересчетом в мозгу (что зависит также, видимо, от жизненного опыта).

Перейдем к увеличению микроскопа. На его объективе происходит дифракция с минимально разрешаемым углом  $\lambda/d$  и соответствующим размером изображения

$L\lambda/d$ . По условию Аббе, для минимально различаемого размера предмета  $x \sin \varphi_{max} = (L\lambda/d) \cdot (d/2L)$ , где  $d/2L$  – синус малого угла, под которым изображению видна линза объектива. Получаем  $x \sim \lambda/(2 \sin \varphi_{max})$ , то есть нельзя различать предметы меньше примерно длины волны. Заметим, что в школьном приближении вместо синуса стоит тангенс, и ограничение исчезает. Мы хорошо видим глазом  $\sim 1$  мм, и для наблюдения предмета порядка  $\lambda = 5 \cdot 10^{-5}$  см нужно увеличение порядка 1000. Поэтому оптические микроскопы и не делают длиной в 10 м – в них мы увидели бы большие, но совершенно не похожие на предмет изображения.

Для наблюдения очень мелких предметов нужно уменьшить длину волны (ультрафиолетовая микроскопия). Пока плохо идет рентгеновская оптика, но применяются электронные и ионные микроскопы до размеров, уже приближающихся к ангстрем (атомных). Поскольку электрон в атоме как раз локализован в таком размере, возникает новая проблема – разрушение объектов «взглядом» (точнее, освещением).

Хотя форма малых предметов перестает улавливаться, заменяясь на дифракционное пятно, свет сам по себе виден, и на темном поле присутствие мелких частиц наблюдается (ультрамикроскопия). Применяется и освещение предмета через тонкую трубочку размером, скажем, 100 Å, и запись отраженного сигнала при сканировании. Здесь дифракционные ограничения заменяются очень сложными технологическими.

Рекордно малые объекты сейчас наблюдают в сканирующих туннельных и атомно-силовых микроскопах. Первые регистрируют ток с острой иглы на поверхность подложки при напряжении в несколько вольт. Ток этот, называемый туннельным, идет прямо через вакуумный зазор и возможен потому, что электрон – тоже волна и может просачиваться через потенциальные барьеры. Игла «ездит» в двух направлениях вдоль поверхности под действием пьезоэлементов. Всякий выступ, например, отдельный атом, вызывает увеличение тока, и на экране компьютера появляется цветное пятно. Атомно-силовой прибор визуализирует притяжение иглы к поверхности, и тоже различает отдельные атомы. Те же иглы позволяют перемещать атомы по подложке, составляя из них рисунки и надписи.

Оценим максимальное увеличение телескопа. Сомнительно, чтобы не было никакого предела, кроме технического, определяемого отношением фокусных расстояний (или длины телескопа к  $f_2$  окуляра). На входном зрачке (линзе или зеркале) диаметра  $d$  свет испытывает дифракцию. Угол центрального максимума, в котором содержится большая часть света,  $\theta_0 \sim \lambda/d$ . Поэтому телескоп не сможет разделить две звезды, угловое расстояние между которыми меньше  $\theta_0$ . Их изображения в фокальной плоскости будут накладываться. Увеличивая углы в  $\Gamma = f_1/f_2$  раз, получим хоть и большое, но искаженное изображение.

Звезда типа Солнца (диаметр 1,5 млн км) с расстояния 4 св. года  $= 4 \cdot 3 \cdot 10^7 \text{ с} \cdot 3 \cdot 10^5 \text{ км/с} = 3,6 \cdot 10^{13}$  км видна под углом  $4 \cdot 10^{-8}$ . Чтобы разрешить ее диск, нужен телескоп с таким же отношением  $\lambda/d$ , то есть  $d \approx 12$  м. Сейчас это технически достижимо, поскольку успехи информатики позволяют объединять изображения от нескольких больших зеркал. Однако на Земле наблюдению мешает атмосфера. Преломление света на флуктуациях плотности приводит к размытию изображения, заметно большему дифракционного предела. Сейчас на орбите



работает 2400 мм «хаббловский» телескоп<sup>6</sup>, который на 1993 г. дал разрешение лучше земных ( $0,007'' \approx 3,4 \cdot 10^{-8}$ ). Это выше дифракционного предела для типичного видимого света  $\lambda/d \approx 5 \cdot 10^{-5}/240 = 2 \cdot 10^{-7}$  благодаря возможности (недоступной на Земле) работать в ультрафиолетовом диапазоне. Получается, что диск звезды виден «на пределе». Видеть же планеты, как имеющие размер объекты, даже у ближайших звезд тем более пока невозможно.

Тем не менее число обнаруженных вне Солнечной системы планет уже исчисляется сотнями (по <http://www.obspm.fr/encycl/encycl.html> 153 планеты в 134 планетных системах на 14 февраля 2005; по <http://planetquest.jpl.nasa.gov/> 199 планет на 11 января 2007). Большая часть из них выявлена по возмущениям движения звезд, а меньшая – в некотором смысле видна, но способами, далекими от прямого визуального наблюдения. Скажем, замечают слабые изменения светимости при «затмении» звезды планетой. На пределе удастся увидеть и отраженный планетой свет, если заблокировать свет звезды (который сильнее в миллионы или миллиарды раз); этот способ похож на ультрамикроскопию. Сейчас крупнейший земной телескоп расположен на горе Мауна Кеа (о. Гавайи). Собственно, это два десятиметровых зеркала, каждое из 36 гексагональных сегментов, разнесенных на расстояние 85 м. Это и позволяет использовать машину как интерферометр, в частности, вычитанием сигналов «отключать» свет звезды. База интерферометра позволяет получить угловое разрешение в тысячные доли угловой секунды ( $\approx 10^{-8}$ ). Планируется постройка европейского телескопа E-ELT с эквивалентным диаметром 42 м; зеркало будет состоять из 906 шестиугольных отдельно управляемых сегментов.

Спутник, летающий на высоте 200 км и снабженный оптикой со входным зрачком 20 см, различает углы  $\lambda/d \sim 2,5 \cdot 10^{-6}$ , что соответствует расстоянию между деталями на Земле  $2,5 \cdot 10^{-6} \cdot 200 \cdot 10^3 = 0,5$  м.

**Оптический вечный двигатель второго рода.** Линза, дающая изображение Солнца, собирает лучи с большой поверхности в малое пятно. Можно ли получить более высокую температуру, чем на Солнце? Солнце излучает  $4\pi R^2 \cdot \sigma T^4$  энергии, и на орбите Земли через линзу радиуса  $r$  пройдет  $4\pi R^2 \cdot \sigma T^4 \cdot \pi r^2 / (4\pi L^2)$  эрг/с. Здесь  $R$  – радиус Солнца,  $L$  – расстояние до него. Весь этот свет сконцентрируется в пятно площадью  $\pi(Rf/L)^2$ , и на единицу поверхности будет падать  $\sigma T^4 r^2 / f^2$  эрг/с. В равновесии изображение должно излучать столько же:  $\sigma T_1^4$ , и если  $r > f$ , то температура экрана  $T_1$  в этом месте будет больше. Ставим тепловую машину, в которой нагревателем служит изображение Солнца, а холодильником само Солнце, и извлекаем работу.

Из второго закона термодинамики сомнительно, чтобы такой проект получился. Но физический этикет требует доказать невозможность любого вечного двигателя в рамках постановки, в данном случае – в оптике. Угол  $\varphi_2 = \arctg(r/f)$  со стороны изображения получается большой,  $\varphi_2 > \pi/4$ , и приближение малых углов не работает. Как это повлияет на результат?

<sup>6</sup>Телескоп планировалось запустить в 1986 г. Однако из-за гибели шаттла «Challenger» в январе этого года запуск был отложен до апреля 1990, вывод на орбиту осуществлял шаттл «Discovery». Только после запуска обнаружилась небрежность сборки – плоскость, где фиксировалось изображение, располагалась немного не в фокусе. В декабре 1993 года неисправность была устранена прямо на орбите, для чего запускался шаттл «Endeavour»; космическими ремонтниками были Story Musgrave и Jeffrey Hoffman.

Трудно сомневаться, что на  $1 \text{ см}^2$  изображения приходится  $\sigma T^4 \cdot (R^2/L^2) \cdot (\pi r^2/\pi y^2)$ , где  $y$  – радиус изображения. Как раз его мы определили неправильно:  $y \neq fR/L$ ! Если наша линза хорошая, то есть изображает Солнце широким пучком резко, то должно выполняться условие Аббе:  $Rr/L = y \sin \varphi$ . Получаем  $y = Rr/(L \sin \varphi)$ . Подставим в выражение для потока энергии:

$$\sigma T^4 \cdot \frac{R^2}{L^2} \cdot \frac{\pi r^2 L^2 \sin^2 \varphi}{\pi R^2 r^2 y^2} = \sigma T^4 \cdot \sin^2 \varphi = \sigma T_1^4.$$

Видно, что температура изображения всегда меньше.

Повторим подобные же рассуждения более последовательно. Малый участок  $\Delta S_1$  светящегося предмета с яркостью  $B_1$  излучает в телесный угол  $\Delta \Omega$  мощность  $B_1 \Delta S \cos \varphi \cdot \Delta \Omega$ . Напомним, что для черного тела, вроде Солнца,  $B = \sigma T^4/\pi$ ; здесь применяем не визуальную, а энергетическую яркость. Для света, приходящего на линзу под углом  $\varphi$ ,  $\Delta \Omega = 2\pi \sin \varphi \cdot \Delta \varphi$ , так что к линзе подается мощность  $2\pi \int B \Delta S \cos \varphi \sin \varphi d\varphi = \pi B_1 \Delta S_1 \sin^2 \varphi_1$ . Можно ввести понятие яркости изображения  $B_2$ , по энергии приходящего света (она, пройдя место фокусировки, будет распространяться как от источника с такой яркостью). Аналогично мощность, сходящаяся к изображению, будет  $\pi B_2 \Delta S_2 \sin^2 \varphi_2$ , и может быть только меньше из-за потерь в линзе. Площади пропорциональны квадратам размеров:  $\Delta S_1 = x^2$ ,  $\Delta S_2 = y^2$ . По условию синусов, если мы желаем получить резкое изображение,  $x \sin \varphi_1 = y \sin \varphi_2$ , так что  $B_2 = B_1$ . Реально в линзе часть энергии теряется, и  $B_2 < B_1$ .

Как и в случае с разрешением микроскопа, учет условия синусов вместо бездумного применения школьных формул исправляет положение. Яркость или температура в фокусе линзы всегда меньше, чем у излучающего предмета. На этом обыкновенно останавливаются, предпочитая не замечать возможности фокусировки без выполнения условия Аббе. Заранее не очевидно, что плохое изображение всегда будет менее ярким. Но это уже примем без доказательства.

В отсутствии усиления яркости легко убедиться, посмотрев через линзу на протяженный светящийся объект (небо, экран телевизора). Яркость участка, наблюдаемого мимо линзы, на глаз такая же, как участка, видимого через линзу. Хотя линза уменьшает изображение (концентрирует свет), она увеличивает телесный угол (растягивает этот концентрированный свет по большему интервалу углов). Освещенность же в фокусе, разумеется, возрастает по сравнению с освещенностью без линзы, почему линза и позволяет зажигать бумагу и т.п.

Другое дело – свет от точечных источников, типа звезд. Телескоп вместо точки рисует дифракционный кружок, в который при росте диаметра зеркала попадает больше световой энергии. Пока звезды остаются точками, растяжения их телесного угла не происходит. Поэтому в большой телескоп видно больше звезд. В принципе, можно видеть через телескоп звезды и днем, так как фон неба не увеличивается, а яркость изображений звезд растет пропорционально площади объектива.

Улучшение видимости через зрительную трубу или бинокль в сумерках («ночезрительная труба» по М.В.Ломоносову) чисто физиологическое. При слабом свете падает разрешающая способность глаза – от  $1'$  до  $1^\circ$ . Труба, уменьшая несколько яркость, дает более крупное изображение на сетчатке. Тоже физиологией объясняется опасность наблюдения Солнца в бинокль – организм не справляется с действием яркого пятна большой площади на сетчатку.



# ВОЛНОВАЯ ФИЗИКА

Курс лекций для ФМШ

Часть 1. ОПТИКА И ВОЛНЫ

4. ТЕОРИЯ ОТНОСИТЕЛЬНОСТИ

А. П. Ершов

12 февраля 2007 г.

## Глава 4

# ТЕОРИЯ ОТНОСИТЕЛЬНОСТИ

### 4.1 Принцип относительности Эйнштейна. Предельная скорость

Мы не раз пользовались переходом в разные инерциальные системы отсчета. Многие явления выглядят проще не в «нашей родной» (лабораторной), а именно в какой-то движущейся относительно нас системе. Допустим, удар по мячу – в системе ноги футболиста ( $V$  относительно лабораторной), где мяч просто меняет знак скорости ( $-V \Rightarrow V$ ). Вернувшись в лабораторную систему, находим скорость отскока  $2V$ , что не так очевидно сразу. Многие отвечают, что мяч отлетает со скоростью  $V$ .

Принцип относительности впервые выдвинул Галилей (1610). Согласно этому принципу, все явления протекают одинаково во всех инерциальных системах отсчета. Сидя в корабле, нельзя узнать, движется ли он относительно берега (если, конечно, не выглядывать наружу). Качание маятников, соударения, работа механизмов и т.д. в «движущемся» корабле происходят так же, как в «неподвижном». Само понятие «движется» бессмысленно, пока не указано, относительно чего. В пространстве нет никаких колев или гвоздей, которые помогли бы нам уловить абсолютное движение. Такие метки бывают только на телах, с которыми и связывают системы отсчета.

В классической механике с принципом относительности Галилея тесно связаны преобразования Галилея (ПГ):

$$\begin{cases} x' = x - Vt, \\ t' = t; \end{cases} \quad \begin{cases} x = x' + Vt', \\ t = t'. \end{cases}$$

В лабораторной системе имеем событие  $(x, t)$ , то есть в точке  $x$  в момент  $t$  что-то произошло, хотя бы извержение вулкана. Тогда в штрихованной системе (на корабле), имеющей скорость  $V$  относительно лабораторной, координата того же события будет  $x'$ , а время  $t'$  (по Галилею – то же самое). Подразумевается для простоты, что в обеих системах синхронизованы часы и в нулевой момент начала координат совпадают. Пре-

образования помогают, описав явление – последовательность событий – в одной системе отсчета, перенести результаты в другую. Согласно принципу относительности, основное уравнение механики – второй закон Ньютона – должен быть инвариантным по отношению к ПГ. Поскольку ускорение – производная от скорости, постоянная скорость «сноса»  $V$  не сказывается, и для инвариантности достаточно, чтобы силы не зависели от системы отсчета. Мы уже пользовались таким условием в механике.

Из ПГ следует известная формула сложения скоростей. Если на корабле бежит матрос,  $x' = u't'$ , то в лабораторной системе  $x = x' + Vt' = (u' + V) \cdot t' = (u' + V) \cdot t$ , скорость  $u = u' + V$ .

Если ограничиваться механикой, преобразований Галилея «хватило бы» еще долго. Но мы уже знаем пример, противоречащий этим преобразованиям. Электромагнитная волна имеет скорость, равную электродинамической постоянной  $c$ , введенной в уравнения Максвелла вначале просто для уравнивания размерностей и единиц полей  $E$  и  $H$ . Из вывода (Электродинамика, п. 5.4) следует, что скорость одна и та же в любой системе отсчета, потому что волна может распространяться в пустоте, а токи и заряды нужны только вначале – для возбуждения волны. Контуры, пересекаемые фронтом волны, рисуются во всех системах одинаково, и скорость фронта всегда будет  $c$ .

Либо ПГ, либо уравнения Максвелла неправильны. В поддержку уравнений Максвелла приведем аргумент **близкодействия**. Конечная скорость распространения взаимодействий есть результат локального действия через поле в данной точке. Если сдвинуть заряд в другой галактике, то сначала вокруг него изменится поле, потом это изменение распространится дальше, а мы почувствуем это, может быть, через миллионы лет. В 1987 г. наблюдался взрыв сверхновой 1987А в Большом Магеллановом облаке; это значит, что звезда взорвалась примерно 200000 лет назад – в 198013 году до н.э.

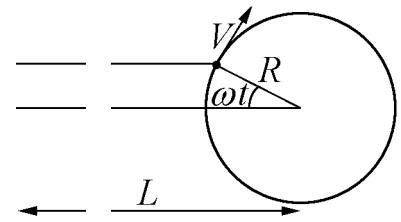
ПГ, с их сложением скоростей, никак не ограничивают скорость распространения сигнала. Это согласуется с противоположной идеей **дальнодействия**: непосредственно через пространство и с какой угодно, возможно бесконечной, скоростью. Принципиально были бы возможны сообщения быстрее света, причем каждая частица Вселенной могла бы чувствовать все другие мгновенно.

Может быть, электродинамика это исключение, а все остальные силы (тяжесть, ядерные и т.д.) дальнодействующие? Во-первых, наш мир в основном электрический – механика, молекулярная физика, химия и биология основаны, хотя не всегда это наглядно, на взаимодействиях зарядов. Но, кроме того, опыт показывает, что и все основные силы – близкодействующие.

Яркий пример нарушения галилеевского сложения скоростей – парадокс двойных звезд. Пусть далеко от Земли имеется двойная звездная система, в которой звезды вращаются вокруг общего центра масс с периодом  $T \approx 1$  год. Это отнюдь не экзотика: большинство звезд входит в состав кратных систем. Рассмотрим последствия очень естественной, хорошо согласующейся с бытовым здравым смыслом **баллистической гипотезы** – пусть скорость света складыва-

ется со скоростью источника, по Галилею. Выберем систему, видную с ребра, в которой звезды движутся попеременно к нам или в обратном направлении. С расстояния  $L$  свет удаляющейся звезды дойдет по баллистической гипотезе за время  $t_1 = L/(c - V)$ , а когда она же будет приближаться, пройдя половину орбиты, мы увидим сигнал через  $t_2 = T/2 + L/(c + V)$ . Разность этих времен  $t_2 - t_1 = T/2 - 2LV/(c^2 - V^2)$  при достаточном расстоянии  $L$  может быть отрицательной – более поздний сигнал придет раньше из-за большей скорости распространения. Мы увидели бы, что поведение звезды никак не согласуется с законами Кеплера, особенно если  $t_2 - t_1$  близко к нулю.

Почувительно рассмотреть не только два момента, а всю динамику видимого движения звезды. Как видно из рис. 4.1, для момента  $t$  «баллистическая скорость»  $c_x = c - V \sin(\omega t)$ . Время прихода света с расстояния  $L$



$$t_1 = t + \frac{L}{c - V \sin(\omega t)} \approx t + \frac{L}{c} + \frac{LV}{c^2} \sin(\omega t).$$

Рис. 4.1.

Большое постоянное слагаемое  $L/c$  можно отбросить (это просто выбор летоисчисления – начала отсчета времени наблюдения). Тогда время  $t_1$  как функция времени  $t$  будет суммой линейной функции и синусоиды. При небольших расстояниях  $L$  роль синусоиды невелика,  $t_1 \approx t$ , и движение звезды выглядит «нормальным»: ее координата  $y$  колеблется, как и положено при вращении:  $y = R \sin(\omega t_1)$ . Конечно, **видим** мы зависимость именно от времени **наблюдения**  $t_1$ .

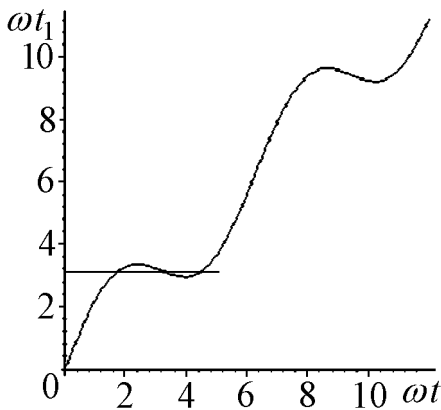


Рис. 4.2.

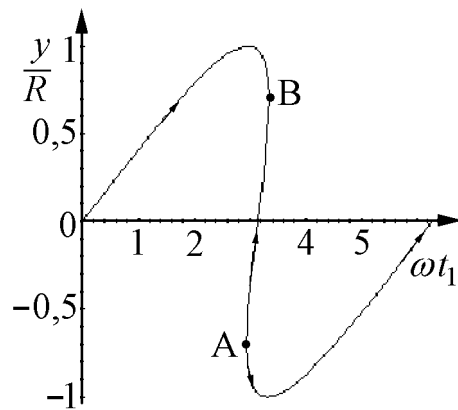


Рис. 4.3.

С ростом  $L$  увеличивается амплитуда синусоиды, и положение меняется. На рис. 4.2 показана зависимость  $t_1(t)$  для  $L = 1,4c^2/\omega V$ . Каждому времени наблюдения отвечает горизонтальная линия, а ее пересечения с кривой определяют фазу, в которой видна звезда. Обратим внимание на промежуток  $2,9 < \omega t_1 < 3,3$ . Здесь есть три пересечения горизонтали с кривой, следовательно, звезда видна в трех положениях одновременно.

Попробуем описать, что видит наблюдатель. Самое левое пересечение – это «нормальное» решение, продолжающее более ранний ход событий. Но в момент  $\omega t_1 \approx 2,9$  «возникает» еще одна звезда и тут же раздваивается. На рис. 4.3, где показана зависимость вертикальной координаты от времени наблюдения, новое изображение возникает а точке А. Один из фантомов

летит навстречу первому изображению, а второй движется в прямом направлении, причем в момент возникновения скорость обоих бесконечна. При  $\omega t_1 \approx 3,3$  пятающийся фантом аннигилирует с тем, который мы считали нормальным (точка В), и остается только одно видимое изображение звезды. Надо полагать, что теперь оно и будет «нормальным».

Заметим, что выбранное расстояние можно записать и как  $L = cT \cdot (1,4c/2\pi V)$ . Здесь  $cT$  – это примерно световой год. В скобке стоит число довольно значительное – порядка  $1,4 \cdot 3 \cdot 10^5 / (2\pi \cdot 30) \approx 2200$ , столько будет световых лет. Но астрономически это не такое большое расстояние, и с него звезды отлично можно наблюдать. Точнее, на таких расстояниях звезды видны хорошо, но телескоп не разделяет детали кратных систем: видна одна яркая точка вместо двух. Тем не менее можно различать отдельные компоненты по их спектрам (так называемые спектрально-двойные звезды). Свет звезды, движущейся по направлению к Солнечной системе, становится более коротковолновым, а удаляющейся – длинноволновым (эффект Доплера). Легко показать (в качестве упражнения сделайте это сами), что относительное изменение длины волны пропорционально продольной скорости:  $\Delta\lambda/\lambda = V \sin(\omega t)/c$ . Поэтому наблюдаемый сдвиг спектра зависит от времени наблюдения так же, как поперечная координата (изображенная на рис. 4.3 зависимость), в частности, будут появляться и исчезать «лишние» источники. При большем удалении можно было бы увидеть пять, семь и т.д. изображений одновременно, если бы была правильна баллистическая гипотеза. Таких чудес на небе не нашли. Пришлось признать, что **скорость света не зависит от движения источника**.

Баллистическая гипотеза неявно заменяет свет потоком песчинок или горошин. Но свет – это электромагнитная волна, вещь совсем другая. Лучи света могут интерферировать (например, гасить друг друга, если направления полей от двух источников противоположны) и дифрагировать (огibtать препятствия). Такие явления несвойственны летящим горошинам. Не удивительно, что баллистическая гипотеза не работает. Чем же ее заменить? По аналогии с упругими волнами ожидали, что световые волны распространяются в специальной среде (эфире), как звук – в воздухе. Предполагалось, что эфир нужен только для электромагнитных волн, что он не имеет заметной плотности и не оказывает сопротивления движению тел в том смысле, что свободно через них протекает. (Сейчас понятно, что это довольно странный набор свойств). Тогда движение источника и не должно влиять на скорость света, поскольку, оторвавшись от источника, свет шел бы по эфиру. Система отсчета, связанная с эфиром, явно выделена, как в акустике – система Земли, в которой неподвижна атмосфера.

Но тогда должны складываться скорости света и приемника. Майкельсон и Морли (1887) поставили решающий опыт, чтобы измерить скорость Земли относительно эфира, но никакого движения найти не удалось. Для измерений был изобретен новый прибор – **интерферометр** (рис. 4.4). На пути луча света стоит полупрозрачное зеркало  $M$ . Свет частично проходит через  $M$  прямо, отражается от хорошего зеркала  $M_1$  и снова от  $M$  отражается вниз на рисунке к экрану  $S$ . Другая часть света отражается вверх к зеркалу  $M_2$  и затем тоже проходит к экрану. На экране лучи встречаются и интерферируют. Некоторым точкам соответствует разность хода в целое число волн (там поля складываются), а некоторым – полуцелое (поля вычитаются). На экране получается «забор» из светлых

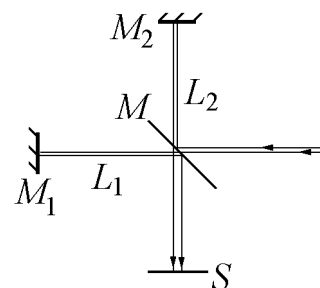


Рис. 4.4.

и темных полос.

Пусть прибор вместе с Землей летит относительно эфира горизонтально со скоростью  $V$ . Тогда время прохода горизонтальной петли  $M - M_1$  равно  $t_1 = L_1/(c + V) + L_1/(c - V) = (2L_1/c)/(1 - V^2/c^2)$ . При  $V \ll c$  получим  $(2L_1/c) \cdot (1 + V^2/c^2)$ . Время для вертикальной петли  $M - M_2$  будет  $2L_2/\sqrt{c^2 - V^2} = (2L_2/c)/\sqrt{1 - V^2/c^2} \approx (2L_2/c) \cdot (1 + V^2/2c^2)$ . Разность времен, определяющая фазу интерференции, содержит большое постоянное слагаемое, которое несущественно, а от скорости зависит как

$$\Delta t_x = \frac{V^2}{c^2} \cdot \left( \frac{2L_1 - L_2}{c} \right), \quad \Delta_x = c\Delta t_x = \frac{V^2}{c^2} \cdot (2L_1 - L_2).$$

Эти величины очень малы, но при интерференции время  $\Delta t$  надо сравнивать с периодом волны, а оптическую разность хода  $\Delta$  – с длиной волны.

Заметим, что если прибор движется поперек исходного луча, надо будет перенести коэффициент 2 в последних формулах:

$$\Delta t_y = \frac{V^2}{c^2} \cdot \left( \frac{L_1 - 2L_2}{c} \right), \quad \Delta_y = \frac{V^2}{c^2} \cdot (L_1 - 2L_2).$$

Разность хода станет другой. Изменение это равно

$$\Delta_x - \Delta_y = \frac{V^2}{c^2} \cdot (L_1 + L_2).$$

Если взять  $V = 30$  км/с (скорость движения Земли по орбите), то  $V^2/c^2 = 10^{-8}$  и получится разность хода  $\lambda/2 \simeq 2.5 \cdot 10^{-5}$  см при  $L_1 + L_2 \approx 25$  м. Где было темно, станет светло (интерференционные полосы сдвинутся). Прибор располагался на каменной плите, которая плавала в баке со ртутью, а измерения проводили при непрерывном медленном вращении плиты. Необходимая длина плеч интерферометра набиралась с помощью ряда дополнительных зеркал. Заметного эффекта от поворота не получилось – полосы не двигались. Не было результата и через три месяца, когда Земля изменила направление движения по орбите на  $90^\circ$ . Таким образом, идея эфира тоже провалилась. Стали рассуждать об увлечении эфира Землей, но и из этого тоже ничего не вышло. По нынешним представлениям, поля «живут» без всякой среды, просто в вакууме, и им никакой эфир не нужен, другими словами, его попросту нет.

Но как быть со сложением скоростей? Разве не надо складывать встречные скорости, отнимать «догоняющие» и компенсировать «снос» при движении поперек? И если нет среды для электромагнитных волн, то относительно чего они имеют скорость  $c$ ? Недостаточно отказаться от эфира, а надо иметь что-то взамен. Мы увидим, что теория относительности вполне согласуется с экспериментом и решает все вопросы такого рода. В частности, толкование опыта Майкельсона приведено в п. 4.3.

Если все передается через какие-то поля, должна существовать **предельная скорость** распространения взаимодействий (или сигналов). Важно, что эта скорость, если она вообще существует, **одинакова** во всех системах отсчета (в силу принципа относительности). Естественный выбор в пользу ограниченной предельной скорости приводит,

как видим, к неожиданному результату – **инвариантности** этой скорости, что противоречит преобразованиям Галилея и известному нам закону сложения скоростей.

Принцип относительности Эйнштейна утверждает, что все природные явления (не только механические) протекают одинаково во всех инерциальных системах отсчета. Закипание чайника, игра в домино, прием пищи в каюте корабля приведут к таким же результатам, как на суше (если система корабля инерциальна – нет, например, качки). В частности, во всех системах одинакова предельная скорость распространения взаимодействий. С точностью современного эксперимента, с предельной скоростью распространяются электромагнитные волны, отчего короче ее называют скоростью света.

Важное и заранее не очевидное следствие инвариантности скорости света – **относительность одновременности** событий, происходящих в разных точках пространства. Скажем, мы на корабле решили синхронизовать часы на носу и корме. Став посередине, посылаем лучи света к часам. Свет идет одинаковое время, и увидев вспышку, ассистенты устанавливают часы на 0 одновременно. Этот способ синхронизации (указанный Эйнштейном) выглядит вполне разумным и естественным.

Если же на этот процесс смотреть из лабораторной системы, то передние часы убегают от луча света, а задние едут навстречу ему. Свет же имеет, как мы приняли, следуя экспериментальным данным и принципу Эйнштейна, всегда одну и ту же скорость. Явно передние часы будут отставать после такой синхронизации. События пуска часов в лабораторной системе не одновременны. Более того, при наблюдении из лабораторной системы все часы на корабле показывают разное время, именно благодаря усилиям команды по их синхронизации. Наоборот, команда видит свои часы установленными правильно, а береговые рассогласованными. Более подробно относительность одновременности рассматривается ниже, в п. 4.3.

Если принцип Галилея «произошел» из механики, то основа принципа Эйнштейна – электродинамика, более фундаментальная область знаний. Имеющаяся в электродинамике предельная скорость распространена на все явления природы. Значит, принцип относительности Эйнштейна не опровергает принцип Галилея, а наоборот, расширяет его. Теории относительности противоречит не принцип, а преобразования Галилея.

## 4.2 Преобразования Лоренца

Получим правильные преобразования. Довольно общий вид перехода от лабораторной системы  $K$  (событие  $x, t$ ) к системе  $K'$  ( $x', t'$ ):

$$\begin{cases} x' = \gamma(x - Vt), \\ t' = \delta t - \varepsilon x. \end{cases}$$

Здесь использовано требование однородности пространства и времени – это проявилось в линейности преобразований – и соответствие начала координат  $x' = 0$  точке  $x = Vt$ .

Коэффициенты  $\gamma$ ,  $\delta$  и  $\varepsilon$  не зависят от  $x, t$  (но могут зависеть от  $V$ ) и пока неизвестны. ПГ – частный случай, в них  $\gamma = \delta = 1$ ,  $\varepsilon = 0$ . Непривычный вариант  $\varepsilon \neq 0$  необходим, иначе одновременность событий будет абсолютной; остальные нововведения ( $\gamma, \delta \neq 1$ ), как мы увидим, также понадобятся.

Рассмотрим начало координат системы К ( $x = 0$ ). В системе К' эта точка движется: для нее  $x' = -\gamma Vt$ ,  $t' = \delta t$ . Скорость движения  $x'/t' = -V\gamma/\delta$ . Если мы хотим равноправия систем отсчета, эта скорость должна быть  $-V$ , откуда  $\delta = \gamma$ . Остается два неизвестных коэффициента:

$$\begin{cases} x' = \gamma(x - Vt), \\ t' = \gamma t - \varepsilon x. \end{cases}$$

Из электродинамики и эксперимента известно, что скорость  $c$  не зависит от системы отсчета. Пусть фронт световой волны движется по закону  $x = ct$  в системе К. Получаем

$$x' = \gamma(c - V)t, \quad t' = (\gamma - \varepsilon c)t, \quad x'/t' = \gamma(c - V)/(\gamma - \varepsilon c) \equiv c.$$

Из последнего равенства выражаем  $\varepsilon = \gamma V/c^2$ . Снова упрощаем:

$$\begin{cases} x' = \gamma(x - Vt), \\ t' = \gamma(t - Vx/c^2). \end{cases}$$

Найдем обратные преобразования:  $x, t(x', t')$ . Решая два уравнения, получим

$$\begin{cases} x = \frac{x' + Vt'}{\gamma(1 - V^2/c^2)}, \\ t = \frac{t' + Vx'/c^2}{\gamma(1 - V^2/c^2)}. \end{cases}$$

В силу равноправия систем прямые и обратные преобразования должны иметь одинаковый вид. Для этого должно выполняться равенство  $\gamma(1 - V^2/c^2) = \gamma^{-1}$ , или  $1 - V^2/c^2 = 1/\gamma^2$ . Окончательно имеем **преобразования Лоренца**:

$$\begin{cases} x' = \gamma(x - Vt), \\ t' = \gamma(t - Vx/c^2); \end{cases} \quad \begin{cases} x = \gamma(x' + Vt'), \\ t = \gamma(t' + Vx'/c^2), \end{cases} \quad \gamma = \frac{1}{\sqrt{1 - V^2/c^2}}. \quad (4.1)$$

Прямые и обратные преобразования отличаются только знаком  $V$ . Множитель, или фактор,  $\gamma$  всегда больше 1. При бытовых условиях, например для скорости Земли по орбите 30 км/с, под корнем стоит  $1 - 10^{-8}$ ,  $\gamma = 1,000000005$ . Это число отличается от 1 на границе точности бытовых калькуляторов. Поэтому мы не замечаем неправильности ПГ. При больших же скоростях ( $V \approx c$ )  $\gamma \gg 1$ . Для полноты добавим еще два преобразования:

$$y = y', \quad z = z',$$

то есть координаты, поперечные к скорости системы  $V$ , не меняются.



### 4.3 Относительность одновременности. Сокращение масштаба и замедление времени

Самое поразительное следствие преобразований Лоренца – это уже упомянутая в п. 4.1 модификация понятия одновременности, которую теперь легко рассмотреть количественно. Пусть в системе  $K$  в разных местах, но одновременно происходит несколько событий. Момент этот примем за начальный,  $t = 0$ . Преобразуем событие  $(x, 0) \Rightarrow (x', t')$ :  $(x' = \gamma x, t' = -\gamma Vx/c^2)$ . Уже и по Галилею  $x'$  отличалось от  $x$ . Новое в том, что  $t'$  зависит от положения события. С корабля эти события не одновременны. Те, которые впереди, произойдут в движущейся системе раньше. Аналогично можно получить, что очередность удаленных событий может смениться при переходе в другую систему отсчета.

Это не иллюзия, связанная с распространением и запаздыванием сигналов: мы можем наблюдать события с большого поперечного расстояния  $y \gg x$ . Тогда пути света к наблюдателю будут практически параллельны и одинаковы. Или можно расставить много наблюдателей на длинном корабле, а потом они обменяются впечатлениями. Когда говорят «движущийся наблюдатель видит то-то», всегда следует понимать такую процедуру. В буквальном же смысле это будет неверно: близкого к оси  $x$  наблюдателя действительно может сбить с толку запаздывание света, и раньше всего он увидит ближайшее событие при  $x = 0$ .

Пусть на разных планетах одновременно решено провести дефолт<sup>1</sup>. Для движущегося наблюдателя эти события не одновременны. Может ли он, узнав о более раннем событии, использовать эту экономическую информацию? Пусть космонавт – разведчик наблюдает событие ( $t = 0, x > 0$ ), например в системе Альдебарана; для него крах случится в момент  $t' = -\gamma Vx/c^2$ . На Земле, при  $x = 0$ , реформа произойдет позже, при  $t = t' = 0$ . Казалось бы, есть еще некоторое время, чтобы связаться со своим партнером на Земле и посоветовать ему скупать (либо продавать) все подряд. Но чтобы попасть на Землю хотя бы в момент «Ч», пройдя расстояние  $x' = \gamma x$ , телеграмма должна идти со скоростью  $x'/|t'| = c^2/V > c$ , что невозможно. Информация придет слишком поздно – уже при  $t' > 0$ . Тот же сигнал в системе Земли послан в момент  $t = 0$ , и опоздание его становится очевидным.

Возьмем стержень, покоящийся в системе  $K'$  и имеющий в ней длину  $L_0$  (так называемую собственную длину, которую мы обычно и меряем). Координата левого конца  $x'_1 = 0$ , правого  $x'_2 = L_0$ . Как выглядит этот стержень из лабораторной системы  $K$ ? Из преобразований Лоренца  $x' = \gamma(x - Vt)$ . Надо брать координаты  $x$  в один момент лабораторного же времени  $t$ , проще всего в нулевой. Получаем  $L = x_2 - x_1 = (x'_2 - x'_1)/\gamma = L_0/\gamma$ . Размер стержня сокращается в  $\gamma$  раз. Этот эффект называют **лоренцевым сокращением**, или **сокращением Лоренца – Фицджеральда**.

<sup>1</sup>В прошлом веке примером служило повышение цен.

Может показаться, что более прямой путь – взять формулу обратного преобразования  $x = \gamma(x' + Vt')$ . Если в ней считать  $t' = 0$ , получим, наоборот, удлинение. Но здесь не учтена относительность одновременности. Неправильно брать события, одновременные на корабле. Тогда в лабораторной системе левый конец отметится при  $t = 0$ , а правый – при  $t_{\Pi} = \gamma VL_0/c^2$ . Естественно, за это время правый конец проедет  $Vt_{\Pi}$  лишнего, и правильная длина стержня будет  $\gamma L_0 - Vt_{\Pi} = \gamma L_0(1 - V^2/c^2) = L_0/\gamma$ .

Аналогично стержень, покоящийся в лабораторной системе отсчета, сокращается для движущегося наблюдателя. Поэтому принцип относительности не нарушается – все изменения симметричны.

Рассмотрим часы, установленные на корабле при  $x' = 0$  и в начальный момент  $t = 0$  синхронизованные с неподвижными часами. Когда-то они проедут мимо других неподвижных часов, показывающих время  $t$ . Мы уже чувствуем, что движущиеся часы покажут другое. Пишем  $t = \gamma(t' + Vx'/c^2) = \gamma t'$ , так как  $x' = 0$ . Время  $t'$ , прошедшее по движущимся часам, меньше в  $\gamma$  раз (**эйнштейновское замедление** времени).

Возникает кажущееся неравноправие систем отсчета. Если летящие часы покажут меньше, вроде бы можно определить, какая система движется. Однако посмотрим на окрестность  $x = 0$ . Там стоят неподвижные часы, показывающие  $t$ , и пролетают мимо подвижные (их, наверно, много на корабле). Показания подвижных часов

$$t' = \gamma(t - V \cdot 0) = \gamma t,$$

наоборот, больше. Получается вполне симметричная картина. Сидя в любом начале координат, увидим, что пролетающие мимо часы показывают в  $\gamma$  раз больше, чем наши собственные. Опять проявляется относительность одновременности. Систему отсчета надо представлять как размеченные оси координат, на оси  $x$  густо наставлены синхронизированные между собой часы. Синхронизировать их можно хотя бы по телефону, учитывая запаздывание сигнала, либо удалившись далеко по оси  $y$ . В другой системе проделываются такие же манипуляции. Мы можем еще синхронизовать пару часов в разных системах. Обычный выбор в преобразованиях Лоренца – устанавливаются на 0 часы в началах координат. Но после этого другие пары часов в разных системах почти всегда не будут синхронизированы.

Проследим из лабораторной системы за теми часами корабля, которые в момент  $t_0$  как раз подойдут вровень к лабораторным часам в начале координат. Сколько показывали те же корабельные часы при  $t = 0$ ? Находились они при  $x = -Vt_0$ . Тогда

$$t'_{\text{нач}} = \gamma(0 - V(-Vt_0)/c^2) = \gamma V^2 t_0 / c^2.$$

Для лабораторной системы эти часы сильно убежали вперед. На сколько изменятся их показания через лабораторное время  $t_0$ ? Часы как раз подойдут к лабораторному началу координат, и покажут

$$t'_{\text{кон}} = \gamma(t_0 - V \cdot 0/c^2) = \gamma t_0.$$

Мы уже получали, что при встрече корабельные часы покажут в  $\gamma$  раз больше, чем показания  $t_0$  лабораторных в начале координат. Но изменение их показаний  $t'_{\text{кон}} - t'_{\text{нач}} = \gamma t_0(1 - V^2/c^2) = t_0/\gamma$ : такое же, как у корабельных часов, находившихся при  $t = 0$  в начале координат. Приведем пример при  $t_0 = 1, V/c = 0,8 \Leftrightarrow \gamma = 5/3$  (рис. 4.5): Разница показаний двух верхних (корабельных) часов как была  $16/15$ , такой и

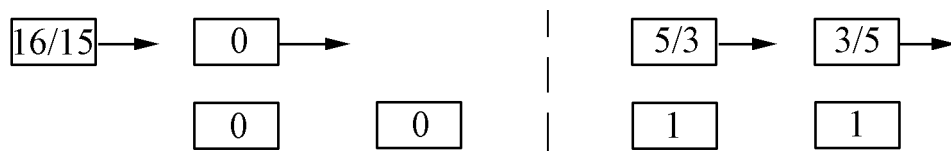


Рис. 4.5.

осталась. Увеличились же показания часов на одинаковое время  $3/5$ . Это и означает замедление времени в движущейся системе отсчета в  $\gamma$  раз.

Мы видим, что ряд величин и соотношений, которые ранее считались абсолютными, переходят в разряд относительных. Соответственно теорию явлений, происходящих при больших скоростях, называют теорией относительности (англ. *relativity theory*, нем. *Relativitätstheorie*); скорости, сравнимые со скоростью света  $c$ , и характерные для таких скоростей эффекты называют релятивистскими.

Иногда лоренцево сокращение и эйнштейновское замедление называют кажущимися. Мол, на самом деле у стержня длина  $L_0$ . Но сокращение происходит именно на самом деле. Например, если надо пропустить летящий стержень через отверстие в стене, надо открыть заслонку на время  $L_0/(\gamma V)$ . Тот же стержень можно поймать в футляр меньшей, чем  $L_0$ , длины. Короткоживущие частицы, получающиеся в физике высоких энергий, пролетают гораздо большие расстояния, чем им положено по времени жизни. Ежедневно теория относительности проверяется на ускорителях в миллионах событий.

Релятивистские парадоксы многим не нравятся. Однако в галилеевской картине мира тоже есть свои неудобства. Вспомним последствия баллистической гипотезы из предыдущего раздела. Теория относительности объясняет все известные эффекты, причем довольно просто, если к ней привыкнуть.

Основная трудность в теории относительности – осознание важности понятия события. От рождения у нас, вероятно, аристотелевское мышление. Как когда-то оппоненты Галилея, мы в детстве думаем, что движущаяся Земля убежала бы из-под ног, что камень, сброшенный с мачты корабля, отстанет от его хода, а также испытываем дискомфорт, видя из окна поезда отъезжающий вокзал. Нужны обучение и практика, чтобы усвоить галилеевскую относительность. Но у Галилея событие не является центральным понятием, и именно из-за инвариантности времени. Преобразования Галилея преобразуют только координаты; время же еще абсолютно. Теперь известно, что вера в абсолютное время – это тоже пережиток прошлого, но более живучий из-за слабости нарушения одновременности в нашем мире малых скоростей и плохой вообще нагляд-

ности времени. Момент времени и место пространства образуют событие. Вот событие – это и есть элемент реальности. Что произошло в одной системе, произошло и в другой<sup>2</sup>, и преобразования Лоренца говорят нам, где и когда. Дополнительный материал о релятивистских парадоксах изложен в приложении к данной главе.

Теперь мы можем понять причину отрицательного результата опыта Майкельсона. В разделе 4.1 мы находили воздействие «эфирного ветра» на плечи интерферометра. Для параллельного скорости плеча время прохода равно

$$\frac{L}{c-V} + \frac{L}{c+V} = \frac{2L}{c} \frac{1}{1-V^2/c^2} = \frac{2L}{c} \cdot \gamma^2,$$

а для перпендикулярного

$$\frac{2L}{\sqrt{c^2-V^2}} = \frac{2L}{c} \cdot \gamma.$$

Вся проблема в том, что движение по-разному удлинняет время по сравнению с простым  $2L/c$ . Фицджеральд в 1889 г. указал, что если бы параллельное плечо сокращалось при движении в  $\gamma$  раз, несимметрия исчезла бы. Как мы видели, такое сокращение действительно происходит. Лоренц и позднее Пуанкаре вывели из электродинамики преобразования, названные именем Лоренца, из которых это сокращение следует. Эйнштейн независимо получил те же преобразования и верно их истолковал – как следствие относительности. В частности, он интерпретировал преобразование времени – не как формально удобное «местное» время, а как самое настоящее время, которое будут показывать часы в движущейся системе. Отметим, что время прохода и параллельного, и перпендикулярного плеча в данном примере в  $\gamma$  раз больше времени  $2L/c$ , которое получилось бы при неподвижном приборе (и которое измерит движущийся вместе с прибором наблюдатель). Это – следствие не изменения скорости света, а эйнштейновского замедления времени. По принципу относительности, такое же время получилось бы при любом повороте «световых часов» (и по любым часам вообще).

Наконец, рассеем мнение, что у Эйнштейна все относительно. Например, скорость света абсолютна – не зависит от системы отсчета. Другой пример – **интервал** между событиями. По определению, квадрат интервала называется величина

$$s^2 = c^2 \Delta t^2 - \Delta x^2,$$

где  $\Delta x$  и  $\Delta t$  разделяют события в пространстве и во времени. Пусть у первого события  $x = t = 0$ , второе задано как  $(x, t)$ . В системе, движущейся со скоростью  $V$ ,

$$s'^2 = c^2 t'^2 - x'^2 = \gamma^2 (c^2 (t - Vx/c^2)^2 - (x - Vt)^2) = c^2 t^2 - x^2 = s^2,$$

что проверяется прямым вычислением. Такое свойство называют **инвариантностью**.

Например, если в одной системе отсчета квадрат интервала положителен, он положителен во всех системах. Тогда существует система, в которой  $\Delta x = 0$  и события происходят в одной

<sup>2</sup>В общей теории относительности бывает еще интереснее. Например, при путешествии в черную дыру космонавт упадет туда за довольно недолгое время по своим часам, а для отдаленного наблюдателя время падения будет бесконечным. Это крайний предел относительности времени – событие, происшедшее в одной (уже неинерциальной) системе отсчета, не происходит в другой.

точке. Только при  $s^2 > 0$  (**времениподобный** интервал) события могут быть причинно связанными. Если же  $s^2 < 0$  (**пространственноподобный** интервал), то есть система отсчета, в которой события происходят в один момент в разных точках. При нулевом интервале события связаны лучом света.

Интервал смешивает время и пространство. Минковский предложил считать  $ct$  четвертой (либо нулевой) координатой наравне с  $x, y$  и  $z$ . Тогда интервал будет аналогом квадрата такого четырехмерного вектора (у нас – двумерного), только квадраты времени и координат имеют разные знаки. Вспомним, что квадрат вектора или скалярное произведение векторов в обычном пространстве не меняются при вращениях, хотя все отдельные компоненты изменяются. Преобразование Лоренца в мире Минковского будет аналогом вращения. Интервал можно понимать как расстояние между событиями в четырехмерном пространстве.

## 4.4 Сохранение энергии-импульса

Сначала выясним, как складываются скорости. На корабле пусть движется матрос со скоростью  $u'$ , то есть  $x' = u't'$ . В лабораторной системе  $x = \gamma(x' + Vt')$ ,  $t = \gamma(t' + Vx'/c^2)$ . Получим лабораторную скорость матроса

$$u = \frac{x}{t} = \frac{u' + V}{1 + u'V/c^2}. \quad (4.2)$$

Это выражение от галилеевского отличается только знаменателем. При малых скоростях знаменатель почти равен 1. Но легко показать, что если складывать  $c$  и любую скорость, получится  $c$ ; при сложении двух скоростей, меньших  $c$ , результат всегда меньше  $c$ .

Скорость преобразуется не по Галилею (что не удивительно), но даже, в отличие от координат, явно не по Лоренцу. Желательно найти связанную со скоростью величину  $s$  с хорошим законом преобразования. Попробуем вычислить  $\gamma(u) = 1/\sqrt{1 - u^2/c^2}$ . Этот множитель пригодится хотя бы для перехода из лабораторной системы в систему идущего по кораблю матроса. На первый взгляд эта величина, выраженная через  $V$  и  $u'$  согласно (4.2), будет очень сложной и громоздкой. Однако, если ее честно упростить, выходит не так уж длинно:

$$\frac{1}{\sqrt{1 - u^2/c^2}} = \frac{1 + u'V/c^2}{\sqrt{1 - V^2/c^2} \sqrt{1 - u'^2/c^2}}. \quad (4.3)$$

Теперь умножим это на  $u$ ; тогда пропадет громоздкий множитель  $1 + u'V/c^2$  и останутся только сумма скоростей и уже привычные гаммы:

$$\frac{u}{\sqrt{1 - u^2/c^2}} = \frac{u' + V}{\sqrt{1 - V^2/c^2} \sqrt{1 - u'^2/c^2}}. \quad (4.4)$$

С такой скоростью может двигаться какая-то частица. Умножим на ее массу  $m$ :

$$\frac{mu}{\sqrt{1 - u^2/c^2}} = \frac{mu' + (V/c) \cdot mc}{\sqrt{1 - V^2/c^2} \sqrt{1 - u'^2/c^2}}.$$

Это уже недалеко от преобразований Лоренца. Чтобы лучше их разглядеть, запишем координатно-временные преобразования, введя вместо времени  $t$  «нулевую» координату  $x_0 = ct$ , имеющую одинаковую размерность с обычной координатой  $x$ , которая считается первой ( $x_1$ ;  $y$  и  $z$  – вторая и третья координаты  $x_{2,3}$ ).

$$x = \gamma(x' + (V/c)ct'), \quad ct = \gamma(ct' + (V/c)x').$$

Формулы стали симметричнее по сравнению с (4.1): входит только  $\gamma$  и безразмерный коэффициент  $V/c$  (часто обозначаемый  $\beta$ ). Во избежание путаницы здесь и далее под  $\gamma$  понимается  $1/\sqrt{1 - V^2/c^2}$ , то есть величина, вычисленная по фиксированной скорости системы отсчета. Обозначаем

$$\frac{mu}{\sqrt{1 - u^2/c^2}} = p_1, \quad \frac{mu'}{\sqrt{1 - u'^2/c^2}} = p'_1,$$

$$\frac{mc}{\sqrt{1 - u^2/c^2}} = p_0, \quad \frac{mc}{\sqrt{1 - u'^2/c^2}} = p'_0,$$

Тогда вместо (4.3,4.4) получим пару преобразований

$$p_1 = \gamma(p'_1 + (V/c)p'_0), \quad p_0 = \gamma(p'_0 + (V/c)p'_1).$$

Это – преобразования Лоренца; остается установить смысл величин  $p_1$  и  $p_0$ . Для этого используем соответствия в нерелятивистском пределе. При малых скоростях  $\sqrt{1 - u^2/c^2}$  в знаменателе  $p_1$  стремится к 1, и  $p_1$  переходит попросту в нерелятивистский импульс  $p_x = mu$ . Разумно принять, что правильное выражение релятивистского импульса

$$p = \frac{mu}{\sqrt{1 - u^2/c^2}}, \quad \text{а в векторном виде} \quad \mathbf{p} = \frac{m\mathbf{u}}{\sqrt{1 - u^2/c^2}}. \quad (4.5)$$

Чтобы разобраться с нулевой компонентой  $p_0$ , умножим ее на  $c$ :

$$\frac{mc^2}{\sqrt{1 - u^2/c^2}} = mc^2 \cdot (1 - u^2/c^2)^{-1/2} \approx mc^2 \cdot \left(1 + \frac{u^2}{2c^2}\right) = mc^2 + \frac{mu^2}{2}.$$

Приближенное равенство годится при малых скоростях (бином Ньютона). Видим большую постоянную величину с добавкой  $mu^2/2$  – нерелятивистской кинетической энергии. Эйнштейну потребовалась определенная смелость, чтобы провозгласить, что  $cp_0$  есть релятивистское выражение энергии:

$$E = \frac{mc^2}{\sqrt{1 - u^2/c^2}}. \quad (4.6)$$

Тогда получается, что даже при нулевой скорости любое тело имеет **энергию покоя**:

$$E = mc^2. \quad (4.7)$$

Эта формула (Эйнштейн, 1905) относится к числу самых популярных и красуется на сникерсах, вывесках фитнес-заведений и пр. Энергия покоя макроскопических тел

огромна, так как (по нашим масштабам) скорость света велика. Например, все атомные электростанции страны вырабатывают в год порядка  $10^{18}$  Дж, что соответствует 10 кг массы. Если научиться массу переводить в энергию полностью, выйдет большая экономия. Реально перегружается порядка 1000 т/год ядерного топлива из-за сильного разбавления, неполного выгорания, скромного теплового КПД и, главное, неполной эффективности ядерной реакции: используется около 0,1% массы распавшегося ядра  $^{235}\text{U}$ . Все же это гораздо меньше, чем потребовалось бы угля ( $\sim 10^8$  т/год).

Что мы выиграли, изменив выражения для энергии и импульса? Они теперь гораздо сложнее зависят от скорости, но, правда, проще стали преобразования к другой системе отсчета. Может быть, лучше пользоваться старыми, нерелятивистскими выражениями?

Основная польза от энергии и импульса – в законах их сохранения. В то же время в теории относительности шагу не ступить без перехода между системами отсчета. Вряд ли будет смысл в законе сохранения, если он выполняется только в одной системе отсчета. По принципу относительности, если уж импульс сохраняется, то сохраняется в любой системе.

Именно такой плохой величиной будет нерелятивистский импульс  $mu$ . Рассмотрим простую задачу – неупругое столкновение одинаковых тел. В «штрихованной» системе их скорости возьмем противоположными, и нерелятивистский импульс сохраняется:

$$\text{(до удара)} \quad m \cdot u' + m \cdot (-u') = 2m \cdot 0 \text{ (после)}.$$

Нулевая скорость после слипания получается из симметрии. В лабораторной же системе, в которой одно из тел до удара покоится, получившееся тело имеет скорость  $u'$ , а ударник налетает с релятивистской суммарной скоростью  $u = (u' + u')/(1 + u'^2/c^2)$ . Равенства не получается:

$$\text{(до удара)} \quad 2m \cdot u'/(1 + u'^2/c^2) \neq 2m \cdot u' \text{ (после)}.$$

Можно только предположить, что почему-то масса выйдет не  $2m$ , а как раз такая, чтобы получилось равенство. Но если рассмотреть процесс из системы с произвольной скоростью  $V$ , то масса уже будет зависеть от скорости  $V$  (явно, а не через изменившуюся скорость  $u'$ ). Проку из этого не выйдет никакого. По принципу относительности, измеряя массу, нельзя найти скорость вашей системы. Математически корень зла в том, что преобразование скорости сложное (нелинейное).

Наоборот, релятивистские выражения импульса и энергии преобразуются по Лоренцу – линейно. Если были какие-то импульсы и энергии до взаимодействия и сумма по всем частицам после взаимодействия такая же, то преобразование к любой системе не нарушит равенства. Для рассмотренного примера в движущейся системе отсчета

$$\frac{mu'}{\sqrt{1 - u'^2/c^2}} + \frac{m(-u')}{\sqrt{1 - u'^2/c^2}} = 0,$$



$$\frac{mc^2}{\sqrt{1-u'^2/c^2}} + \frac{mc^2}{\sqrt{1-u'^2/c^2}} = Mc^2.$$

Из второго уравнения – закона сохранения энергии – видим, что масса получившегося покоящегося тела больше, чем  $2m$ : кинетическая энергия при ударе перешла в массу.

В системе покоя одного из тел до удара у этого тела скорость 0, у налетающего  $u = 2u'/(1+u'^2/c^2)$ , импульсы до удара 0 и  $mu/\sqrt{1-u'^2/c^2}$ , что дает  $2mu'/(1-u'^2/c^2)$  (получается прямым вычислением). После удара скорость  $u'$ , масса  $M = 2m/\sqrt{1-u'^2/c^2}$ , импульс  $Mu'/\sqrt{1-u'^2/c^2}$  такой же, как до удара. Точно так же энергия сохраняется независимо от системы отсчета. Мы получили это максимально «тупым» образом через преобразования скоростей; через преобразования Лоренца для импульса и энергии то же получится автоматически при любой скорости системы. Линейные преобразования не могут нарушить равенства двух сумм.

В действительности при больших скоростях сама скорость – нехарактерная величина, а более удобны как раз энергия и импульс. Аналогично интервалу, легко ввести инвариантную величину: вместо  $c^2t^2 - x^2$  пишем  $E^2/c^2 - p^2$ . Раз это вещь инвариантная, ее можно найти в системе покоя частицы, где она равна  $m^2c^2$ . Или можно то же получить прямым вычислением из выражений  $E$  и  $p$  через скорость. Отсюда имеем явную связь энергии частицы с импульсом:

$$E = \sqrt{c^2p^2 + m^2c^4}.$$

Это выражение заменяет параметрическую зависимость  $p(u)$ ,  $E(u)$ .

Из последней формулы видна возможность существования безмассовых частиц, обладающих, однако, энергией и импульсом. При  $m = 0$  будет  $E = cp$ . Скорость при этом, очевидно, равна  $c$ . По имеющимся сведениям, такими частицами являются фотоны. Вместо скорости характерной кинетической величиной у них будет частота  $\omega$  или длина волны  $\lambda$ :  $E = \hbar\omega$ ,  $p = \hbar k = 2\pi\hbar/\lambda$ . Здесь  $\hbar = 1,054 \cdot 10^{-27}$  эрг·с – постоянная Планка<sup>3</sup>.

Вернемся к рассмотрению неупругого соударения двух частиц в более реальной постановке. Пусть соударяются электрон и позитрон с противоположными импульсами. При достаточной энергии  $E$  частиц в результате может получиться пара протон + антипротон. Для этого нужно как минимум  $2E = 2Mc^2$ , если родившиеся частицы покоятся.  $M$  – это масса протона, превышающая массу электрона  $m$  в  $\approx 1840$  раз. Такая постановка описывает опыт на встречных пучках. Скорость частиц  $V$ , следовательно, такова, что  $\gamma = 1/\sqrt{1-V^2/c^2} \geq M/m$ .

Если же разгонять позитрон, а затем ударить им по покоящемуся в мишени электрону, потребуется энергия позитрона, которую можно найти из преобразования Лоренца:  $E_1 = \gamma(E + Vp) \approx \gamma(E + VE/c) \approx 2\gamma E = (2M/m) \cdot Mc^2$ . Видна выгода метода встречных пучков: для того же эффекта нужна в тысячи раз меньшая энергия пучка.

<sup>3</sup>Разумеется, выражения, содержащие  $\hbar$ , никак не следуют из уже рассмотренного материала; мы получим их обоснование в следующих разделах.



Релятивистские зависимости импульса и энергии от скорости в популярной литературе обычно объясняют как увеличение массы:

$$m = \frac{m_0}{\sqrt{1 - u^2/c^2}}.$$

При этом  $m_0$  называют массой покоя. Переопределение массы приносит некоторые частные удобства. Например, полная энергия через «увеличенную» массу выражается так же, как энергия покоя:  $mc^2$ , а импульс будет  $m\mathbf{u}$ . Сейчас такой подход признается устаревшим (хотя по традиции еще может использоваться в технике). В физике «массу покоя» называют просто массой и обозначают  $m$ . Импульс тогда равен  $m\mathbf{u}/\sqrt{1 - u^2/c^2}$ , энергия  $mc^2/\sqrt{1 - u^2/c^2}$ . До некоторой степени дело вкуса, что считать массой. Но мы увидим, что в динамике переменная масса неудобна. Вообще полезнее найти что-то постоянное («физическая» масса является лоренц-инвариантом), нежели меняющееся.

В теории относительности импульс и энергия – не независимые величины, как в нерелятивистской механике, а взаимосвязанные. Они преобразуются по Лоренцу друг через друга при переходе в другую систему. То же относится ко времени и координатам события. Для компактного представления таких величин вводится понятие **четырёхвектора**, т.е. вектора в четырёхмерном пространстве. В частности, событие  $(t, x, y, z)$  описывается четырёхвектором события  $x_i$ ,  $i = 0..3$ :

$$x_i : (x_0 = ct, x_1 = x, x_2 = y, x_3 = z) \quad \text{или} \quad x_i = (ct, \mathbf{r}). \quad (4.8)$$

Аналогично энергия и три компоненты импульса есть компоненты четырёхмерного вектора энергии-импульса (кратко – четырёхимпульса или 4-импульса):

$$p_i : (p_0 = E/c, p_1 = p_x, p_2 = p_y, p_3 = p_z) \quad \text{или} \quad p_i = (E/c, \mathbf{p}). \quad (4.9)$$

Правильное уравнение, каковым является закон сохранения, связывает значения четырёхимпульса до и после взаимодействия.

## 4.5 Релятивистская динамика\*

Нам надо получить аналог второго закона Ньютона, правильный при любых скоростях. Пусть скорость частицы  $u$  сравнима с  $c$ , тогда выберем такую систему отсчета, в которой частица почти покоится или хотя бы движется нерелятивистски. Скорость такой почти сопутствующей системы  $V$  близка к  $u$ , но в отличие от  $u$  строго постоянна (для инерциальности).

В движущейся системе можно описать движение точно. Если на частицу действует продольная сила  $F'$ , то ее скорость  $u' = F't'/m$ . Переходим в лабораторную систему:

$$u = \frac{V + u'}{1 + u'V/c^2} = (V + u')(1 - u'V/c^2) = V + u'(1 - V^2/c^2).$$

Второе слагаемое – изменение лабораторной скорости  $\Delta u$  из-за действия силы. Так как  $u'$  зависит от  $t'$ , преобразуем еще время:

$$t = \gamma(t' + Vx'/c^2) = \gamma t',$$

поскольку  $x' = F't'^2/2m$  мало. Или можно сказать, что  $t'$  это сокращенное собственное время. Получаем

$$\Delta u = (F't'/m)(1 - V^2/c^2) = (F'/m) \cdot (1 - V^2/c^2)^{3/2} \cdot t.$$

Время  $t$  – это, собственно,  $\Delta t$ , так что

$$m \frac{\Delta u}{\Delta t} = F'(1 - V^2/c^2)^{3/2} \quad \text{или} \quad \frac{m\Delta u/\Delta t}{(1 - V^2/c^2)^{3/2}} = F'.$$

Вычислим производную

$$\frac{d}{dt} \left( \frac{mu}{\sqrt{1 - u^2/c^2}} \right) = m \frac{du}{dt} \frac{1}{\sqrt{1 - u^2/c^2}} + \frac{mu^2}{c^2} \frac{du}{dt} \frac{1}{(1 - u^2/c^2)^{3/2}} = m \frac{du}{dt} \frac{1}{(1 - u^2/c^2)^{3/2}}.$$

Вспоминаем, что система квазисопутствующая. Под корнем после дифференцирования можно считать  $u = V$ , и уравнение движения

$$\frac{dp}{dt} = \frac{d}{dt} \frac{mu}{\sqrt{1 - u^2/c^2}} = F'. \quad (4.10)$$

Сила  $F'$ , например, может быть  $eE'$  – электрическая, вычисленная по полю в сопутствующей системе.

Теперь пусть действует поперечная сила, так что скорость по оси  $z$  равна  $w' = F't'/m$ . Смещение «поперек» не изменяется:

$$z = F't'^2/2m = (F'/2m) \cdot (1 - V^2/c^2) \cdot t^2.$$

Ускорение, перпендикулярное скорости, уменьшается в  $\gamma^2$  раз. То же самое напишем так:

$$\frac{d}{dt} \frac{mw}{\sqrt{1 - V^2/c^2}} = F' \sqrt{1 - V^2/c^2}.$$

Скорость  $\mathbf{u}$  – это векторная сумма  $iV$  и  $\mathbf{k}w$ , взаимно перпендикулярных. Если появляется малая скорость  $w$ , на величине  $u$  это скажется во втором порядке – по теореме Пифагора. Поэтому опять можно поменять  $V$  на  $u$ :

$$\frac{d}{dt} \frac{mw}{\sqrt{1 - u^2/c^2}} = F' \sqrt{1 - u^2/c^2}. \quad (4.11)$$

В уравнениях движения (4.10,4.11) в левой части под знаком производной стоит понятная величина – соответствующая компонента импульса. В правой части пока что «не та» сила. Из лабораторной системы нужно послать делегатов в сопутствующую, чтобы узнать  $F'$ . В (4.11) стоит  $F'$ , еще и деленная на  $\gamma$ . Как преобразуются силы?

Продольную силу получим, рассмотрев конденсатор, пластины которого перпендикулярны скорости  $V$ . По теореме Гаусса поток электрического поля  $ES$  равен  $4\pi Q$  в любой системе отсчета. Поперечные размеры не меняются,  $S$  то же самое. Отсюда получаем преобразование продольного поля:  $E_{\parallel} = E'_{\parallel}$ .

Для преобразования поперечной силы берем неподвижный конденсатор, пластины которого параллельны скорости. В сопутствующей системе поток  $E'S'$ , а в лабораторной  $ES$ . Из-за лоренцева сокращения движущаяся площадь  $S'$  меньше в  $\gamma$  раз, так что  $E'_{\perp} = E_{\perp}/\sqrt{1 - V^2/c^2}$ .

В сопутствующей системе есть еще поле  $H$  от летящего конденсатора. Но оно не влияет на покоящуюся частицу. Теперь видно, что в обоих случаях в правой части уравнения движения стоит попросту компонента  $eE$ , или сила в лабораторной системе отсчета. Окончательно векторное уравнение движения выглядит так:

$$\frac{d\mathbf{p}}{dt} = \mathbf{F}, \quad \mathbf{F} = e \left( \mathbf{E} + \frac{\mathbf{u} \times \mathbf{H}}{c} \right).$$

Это уравнение годится при любых скоростях.

Рассмотрим релятивистское движение в простейших случаях. В магнитном поле сила поперечна скорости. Поэтому величина скорости, как и в нерелятивистском случае, не меняется. Тогда

$$\frac{d}{dt} \frac{m\mathbf{u}}{\sqrt{1-u^2/c^2}} = \frac{m}{\sqrt{1-u^2/c^2}} \frac{d\mathbf{u}}{dt} = e \frac{[\mathbf{u}\mathbf{H}]}{c},$$

и меняется только направление вектора  $\mathbf{u}$ . Верны те же нерелятивистские формулы, но с заменой массы на  $m/\sqrt{1-u^2/c^2}$ . Например, частота вращения  $\omega = (eH/mc)\sqrt{1-u^2/c^2}$ , так что релятивистский циклотрон будет работать плохо. На практике пользуются не скоростью, а энергией частицы  $\mathcal{E}$ . Для ускорителей важен радиус обращения  $R = muc / (eH\sqrt{1-u^2/c^2}) \approx \mathcal{E}/eH$ . При энергии 1 ГэВ, поле 1 кЭ радиус будет  $(10^9/300)/10^3 \approx 3000$  см или 30 м. Сейчас законсервировано строительство ускорителя в Техасе с окружностью 100 км.

В электрическом поле, параллельном движению, можно написать<sup>4</sup>

$$\frac{dp}{dt} = \frac{d}{dt} \left( \frac{mu}{\sqrt{1-u^2/c^2}} \right) = eE = \text{const}.$$

Отсюда  $u/\sqrt{1-u^2/c^2} = (eE/m) \cdot t$ . В скобках стоит ускорение в сопутствующей системе отсчета. Поэтому мы одновременно решаем задачу о полете космонавта к звездам с субъективно постоянным ускорением  $a' = eE/m$ , что, конечно, занимательнее.

Находим скорость:  $u/c = (a't/c)/\sqrt{1+(a't/c)^2}$ . Разумеется, космонавту удобнее постоянное ускорение  $g = 10^3$  см/с, создающее в ракете земную тяжесть. Время будем мерять в годах ( $3 \cdot 10^7$  с). Тогда за 1 год  $a't \approx c$  и безразмерная скорость (в единицах  $c$ )

$$u = t/\sqrt{1+t^2}.$$

Видно что скорость всегда меньше 1, но стремится к 1 (то есть скорости света) при больших  $t$  (должно пройти много лет). При малых  $t$  скорость пропорциональна времени – нерелятивистское поведение. Пройденное расстояние:  $dx/dt = u$ , откуда

$$x = \int \frac{t}{\sqrt{1+t^2}} dt = \sqrt{1+t^2} - 1.$$

При малых  $t$  получается  $t^2/2$  (в размерных единицах школьное выражение  $a't^2/2$ ), а при больших практически  $x = t - 1$ : движение почти со скоростью света. Напомним, что расстояние здесь выражено в световых годах, а в сантиметрах будет  $c(t - c/a')$ .

<sup>4</sup>Заметим, что сторонники школы увеличивающейся со скоростью массы здесь пишут, как правило, неверное уравнение  $(m/\sqrt{1-u^2/c^2}) \cdot du/dt = eE$  по аналогии со вторым законом Ньютона. Корень нельзя выносить из-за знака производной.

Найдем еще затраты собственного времени космонавта. Собственное время  $\tau$  меньше  $t$ :  $d\tau = dt \cdot \sqrt{1 - u^2/c^2} = dt/\sqrt{1 + t^2}$ . Такой интеграл берется:  $\tau = \ln(t + \sqrt{1 + t^2})$ . При малых временах (меньше года)  $\tau \approx t$ , но при больших очень слабо зависит от времени полета:  $\tau \approx \ln(2t)$ . Для наглядности нарисуем таблицу:

$t$ , лет	0	1,0	3,0	10	100	$10^3$	$10^4$	$10^5$	$10^6$	$10^8$
$x$ , св. лет	0	0,41	2,2	9,1	99	999	$10^4 - 1$	$10^5 - 1$	$10^6 - 1$	$10^8 - 1$
$\tau$ , лет	0	0,88	1,8	3,0	5,3	7,6	9,9	12,2	14,5	19,1

До ближайшей звезды около 4 св. лет, так что за 3 земных года полета пролетим половину расстояния. Затем надо тормозить; в результате туда и обратно затратится  $3 \cdot 4 = 12$  лет по земным часам и  $1,8 \cdot 4 = 7,2$  года по часам космонавта. Дальше релятивистские эффекты проявляются во всей красе. До центра галактики ( $10^4$  св. лет) лететь меньше 10 лет (значит, типичный полет по Галактике займет 40 лет). Однако на Земле пройдут все  $4 \cdot 10^4$  лет, за которые неизбежно сменится не одна цивилизация. Отсюда видна невозможность так называемых галактических империй. До границы наблюдаемой Вселенной лететь около  $2 \cdot 23 = 46$  лет; правда, туда и обратно будет уже 92 года. Вернувшись, такой путешественник застанет погасшее Солнце. Если слетать всего на  $10^8$  св. лет, то в лучшем случае на Земле обнаружатся существа столь же близкие космонавтам, как динозавры. Эти примеры показывают, что в полетах к звездам нет никакого смысла, не говоря об огромных затратах энергии. Лучше пользоваться телефоном. С другой стороны, нам нечего опасаться нашествия космических завоевателей.

Отметим, что логарифмическая зависимость собственного времени получена в случае постоянного ускорения. Если разогнаться даже до большой скорости и затем выключить мотор, то после этого собственное время будет нарастать линейно, хотя и с малым коэффициентом.

В заключение рассмотрим последний релятивистский эффект. Вокруг провода с током есть магнитное поле  $H = 2I/cR$  и на движущийся со скоростью  $V$  заряд действует сила Лоренца. Например, если  $V$  направлена вдоль тока, положительный заряд притягивается к проводу с силой  $qVH/c^2 = 2qVI/c^2R$ . Притяжение должно наблюдаться и из сопутствующей системы. Но в ней нет силы Лоренца, так как скорость заряда нулевая. Что же притягивает?

В лабораторной системе провод не заряжен и электрического поля нет. Мы знаем, что при движении происходят искажения длин. Может, это как-то влияет?

Пусть в проводе двигались положительные заряды со скоростью  $u$ . Их плотность  $n_0$  такая же, как плотность неподвижного отрицательного фона; в результате провод нейтрален. Из системы, имеющей скорость  $V$ , наблюдается сокращенный провод с теми же поперечными размерами. Поэтому плотность отрицательного фона будет  $n'_- = n_0/\sqrt{1 - V^2/c^2}$ , то есть увеличится. Сложнее найти плотность положительных зарядов. Их скорость  $u' = (u - V)/(1 - uV/c^2)$ . Раз плотность была  $n_0$  в системе, где заряды двигались, то «собственная» плотность в сопроceedающей эти заряды системе была  $n_{+0} = n_0\sqrt{1 - u^2/c^2}$  — меньше, чем  $n_0$ . Теперь  $n'_+$  увеличится по сравнению с  $n_{+0}$  соответственно скорости  $u'$ :

$$n'_+ = \frac{n_{+0}}{\sqrt{1 - u'^2/c^2}} = n_0 \sqrt{1 - u^2/c^2} \frac{1 - uV/c^2}{\sqrt{1 - V^2/c^2} \sqrt{1 - u^2/c^2}} = n_0 \frac{1 - uV/c^2}{\sqrt{1 - V^2/c^2}}$$

Видим, что сумма

$$n'_+ - n'_- = -\frac{n_0 u V / c^2}{\sqrt{1 - V^2/c^2}}$$

не равна нулю. В движущейся системе провод заряжен отрицательно.

Заряд на единицу длины  $\lambda = eS(n'_+ - n'_-)$ , где  $S$  – сечение провода. Электрическое поле в движущейся системе

$$E' = \frac{2\lambda}{R} = -\frac{2Sen_0 u V / c^2}{R\sqrt{1 - V^2/c^2}} = \frac{2IV/c^2}{R\sqrt{1 - V^2/c^2}}.$$

направлено к проводу. На положительный заряд будет действовать сила  $qE'$ . Поскольку магнитное поле в лабораторной системе  $H = 2I/cR$ , сила  $qE'$  отличается от лабораторной силы Лоренца только множителем  $\gamma = 1/\sqrt{1 - V^2/c^2}$ . Мы уже видели, что поперечная сила так и преобразуется.

Некоторые не верят, что движущийся провод заряжен. Пусть заряды поставляются из некоторого ящика в провод с одного конца и убираются в подобный же ящик на другом конце. В лабораторной системе провод нейтрален. Заряды ящиков меняются со временем; для простоты пусть они противоположны:  $\pm It$ . В сумме имеем нуль. В движущейся системе провод заряжен. Как же быть с законом сохранения заряда?

Из движущейся системы не будут противоположны заряды ящиков, так как они будут наблюдаться одновременно в движущейся и неодновременно в лабораторной системе. Проверьте, что избыток как раз компенсируется зарядом провода, так что суммарный заряд нулевой.

Мы встречались с преобразованиями полей в электродинамике, но теперь, наконец, удалось их получить хотя и не в полном виде, но уже в релятивистском случае.

## Приложение. Три парадокса

### Пенал и карандаш

Имеем неподвижные пенал и карандаш, причем карандаш как раз помещается в пенале. Разгоняем карандаш, он испытывает лоренцево сокращение. Можно впустить его в пенал с одной стороны, закрыть крышку, и некоторое время карандаш целиком будет внутри пенала.

Теперь разгоним пенал, чтобы он наехал на покоящийся карандаш. Пенал, наоборот, стал короче и явно не уместит в себе весь карандаш. Или, вместо разгона пенала, можно просто перейти в систему карандаша. Может ли результат измениться при переходе в систему отсчета?

Оказывается, вполне может. Нет такого события, как помещение целого карандаша в пенал, так как карандаш нельзя характеризовать единственной координатой: у него два конца. Событиями будут, например, влет в пенал переднего и заднего концов карандаша. В системе пенала длины  $L$  карандаш имеет длину  $L/\gamma$ . Пусть первое событие ( $x = 0, t = 0$ ) будет влет переднего конца в пенал. Влет заднего конца – второе событие ( $x = 0, t = L/\gamma V$ ). Вылет переднего конца из пенала: ( $x = L, t = L/V$ ). В системе карандаша первое событие запишем как ( $x' = 0, t' = 0$ ), второе будет ( $x' = -L, t' = L/V$ ), а третье ( $x' = 0, t' = \gamma(L/V - VL/c^2) = \gamma L/V(1 - V^2/c^2) = L/\gamma V$ ). В системе карандаша меняется последовательность событий: вылет переднего конца происходит раньше влета заднего. Это ясно и из формулировки, но теперь

получилось из преобразований Лоренца. Если относительна одновременность, то и порядок событий во времени может меняться при переходе в другую систему отсчета.

### Видимая форма предметов

Пусть высоко над вами пролетают два звездолета пришельцев в форме куба и шара. Как они выглядят? Обычно отвечают, что предметы испытывают лоренцево сокращение и шар, например, выглядит эллипсоидом, сплюснутым в направлении движения. Это неверно.

Начнем с куба. Его нижняя грань действительно сокращена и видна как прямоугольник со сторонами 1 и  $1/\gamma$  (рис. 4.6). Но из-за движения фотоны, испущенные задней гранью куба точно вниз, не будут задевать об эту грань и тоже попадут в глаз или объектив фотоаппарата.

Чтобы прийти одновременно с «нижними» лучами, верхние должны стартовать на  $1/c$  раньше, и задняя грань будет видна как прямоугольник со сторонами 1 и  $V/c$ . Если ввести угол  $\varphi$  такой, что  $\sin \varphi = V/c$ , то  $\cos \varphi = 1/\gamma$ , и куб выглядит не сжатым, а повернутым на угол  $\varphi$  (рис. 4.6). Точно так же повернутым, то есть не изменившимся, виден шар. Хотя **на самом деле** он сплюснен лоренцевым сокращением, из-за различного времени хода света он **кажется** круглым.

### Парадокс близнецов

Один из двух близнецов летит к звездам с хорошей скоростью, второй же для контроля остается на Земле. Для земного наблюдателя часы путешественника замедлены. В силу принципа относительности так же замедлены все процессы, иначе путешественник мог бы определить абсолютное движение. Вернется еще вполне бодрый космонавт и встретит дряхлого старика.

С точки зрения космонавта, замедлены часы его брата на Земле. Почему же, как мы знаем из фантастики, стареет все же неподвижный брат? Вот это и есть парадокс близнецов.

Обычно принято выдвигать объяснение, состоящее в неинерциальности системы отсчета космонавта. Ему надо вернуться, он испытывает ускорения. Поэтому системы неравноправны. Но некоторые начитанные могут вспомнить уже общую теорию относительности, в которой ускорения – самое обычное дело и вообще все системы равноправны. Желательно разобраться в парадоксе без таких непонятных отсылок, в рамках кинематики.

Попробуем использовать для расчета парадокса близнецов еще интервал, как инвариантную величину. Сумма  $\sum \Delta s = \sum \sqrt{c^2 \Delta t^2 - \Delta x^2}$ , взятая между двумя событиями, не зависит от системы отсчета. В лабораторной системе вдоль пути космонавта  $\Delta x^2 = (V \Delta t)^2$ , и получаем  $c \sum \sqrt{\Delta t^2 (1 - V^2/c^2)} = c \sum \Delta t / \gamma$ . В системе космонавта или, вернее, в совокупности инерциальных сопутствующих систем,  $\Delta x' = 0$ , и та же величина выразится как  $c \sum \sqrt{\Delta t'^2} = c \sum \Delta t'$ . Сразу видно, что полное время в системе космонавта меньше, чем в лабораторной системе, и именно в  $\gamma$  раз. Поскольку интервал – аналог расстояния, существует сравнение: парадокс близнецов ясен примерно в той же степени, как то, что прямая короче кривой. Правда, здесь, наоборот, прямая «длиннее» ломаной (в смысле, что интервал для земного близнеца вдоль прямой мировой линии  $c \sum \Delta t$  больше, чем интервал вдоль ломаной для космонавта  $c \sum \Delta t / \gamma$ ).

Добавим, что если кто так и не понял парадокса близнецов, в этом нет ничего удивительного. До сих пор время от времени возникают споры по этому поводу в научной литературе.

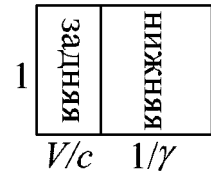


Рис. 4.6.

# ВОЛНОВАЯ ФИЗИКА

Курс лекций для ФМШ

Часть 2. СТРОЕНИЕ ВЕЩЕСТВА

5. КВАНТЫ

А. П. Ершов

12 февраля 2007 г.



# Глава 5

## КВАНТЫ

Мы приступаем к изучению новой, квантовой, физики, развитие которой началось вместе с XX веком. Не так давно (19 декабря 2000 г.<sup>1</sup>) квантовой физике исполнилось 100 лет. Будет правильно начать с Планка и его знаменитой постоянной  $\hbar$ .

### 5.1 Постоянная Планка. Кванты

**Черное излучение.** Нагретые тела светятся. К 1900 г. экспериментаторы хорошо изучили это тепловое излучение и его спектр. Измеряли поток энергии из нагретой полости через малое отверстие наружу (модель абсолютно черного тела). Вышло, что при температуре  $T$  поток энергии  $q = \sigma T^4$ , где  $\sigma = 5,67 \cdot 10^{-5}$  эрг/(см<sup>2</sup>с К<sup>4</sup>) – постоянная Стефана – Больцмана, а частота, при которой больше всего излучается энергии,  $\omega_* \propto T$ .

В старой (классической) теории теплового излучения существовала так называемая **ультрафиолетовая катастрофа**. Полость содержит излучение с различными длинами волн. Каждая из этих волн – это независимое колебание, степень свободы электромагнитного поля. В классической физике на каждую колебательную степень свободы в тепловом равновесии приходится вполне определенная энергия  $kT$ , где  $k = 1,38 \cdot 10^{-16}$  эрг/К – постоянная Больцмана. Не видно было никаких препятствий к существованию сколь угодно коротких волн, так что их «набиралось» бесконечное количество. Тогда бесконечна энергия излучения, а вместе с ней и поток  $q$ , чего не наблюдается.

Если длина волны  $\lambda$  мала, то частота  $\omega = 2\pi c/\lambda$  велика. Чем короче («ультрафиолетовее») мы берем волны, то есть чем больше  $\omega$ , тем больше суммарная энергия излучения с частотами от нуля до  $\omega$ . Это и есть ультрафиолетовая катастрофа.

**Граничная частота. Постоянная Планка.** Раз чудес, вроде бесконечной энергии, не бывает, то для волн большой частоты что-то изменится. Приходится допустить, что при больших частотах энергия быстро спадает; стандартная же энергия  $kT$  может полагаться только волнам достаточно малой частоты. Естественной границей этих режимов

---

<sup>1</sup>В этот день состоялся доклад М. Планка в германском физическом обществе.

будет частота максимума спектра  $\omega_*$ . Ее пропорциональность температуре запишем в «энергетической» форме (так как энергия, конечно, важнее частоты):

$$\hbar\omega_* \sim kT,$$

где  $\hbar$  – коэффициент, получивший название постоянной Планка<sup>2</sup>. Величину его можно найти из потока световой энергии от нагретого тела. По современным данным

$$\hbar = 1,054 \cdot 10^{-27} \text{ эрг} \cdot \text{сек}.$$

**Кванты.** Предыдущая аргументация довольно тривиальна: раз колебаний слишком много, надо их дискриминировать. Основная идея Планка состояла в том, что энергия  $\hbar\omega$  не только «отмечает» граничную частоту, но вообще имеет фундаментальное значение. Именно, излучение состоит из так называемых **квантов**, то есть долей, частичек. Квант с частотой  $\omega$  имеет энергию  $\hbar\omega$ . Таким образом, энергия излучения частоты  $\omega$  меняется не плавно, а дискретно, ступеньками вполне определенной «высоты». Нельзя увеличить эту энергию на  $1,93\hbar\omega$ ,  $0,5\hbar\omega$ ,  $0,017\hbar\omega$ , но можно – на  $\hbar\omega$ ,  $2\hbar\omega$ ,  $2007\hbar\omega$ . Когда энергия кванта превышает тепловую,  $\hbar\omega \gtrsim kT$ , такой квант тепловому движению становится трудно породить (вымораживание). Поэтому квантов больших частот мало, чем и разрешается парадокс. В результате удалось прекрасно описать экспериментальный спектр излучения черного тела. Планк ввел несколько другую постоянную  $h = 2\pi\hbar = 6,626 \cdot 10^{-27}$ , так что энергия записывалась в стиле XIX века как  $h\nu$ . Кванты света называют еще **фотонами**.

Итак, существует новая фундаментальная постоянная  $\hbar$ . Постоянная Стефана – Больцмана  $\sigma$  выражается через нее и, значит, является «менее фундаментальной»; точное выражение, со всеми коэффициентами, имеет вид:  $\sigma = \pi^2 k^4 / (60\hbar^3 c^2)$ .

**Особенности постоянной  $\hbar$ .** Первая особенность – это крайне малая величина. Степень  $(-27)$  в системе единиц, так или иначе определяемой «человеческими» масштабами, есть некоторый рекорд. Практически это значит, что постоянная Планка существенна для микромира. Именно она определяет размеры атомов.

Вторая особенность – необычная размерность (эрг·с). Пока нам не приходилось умножать энергию на время. Впрочем, известная из механики величина – момент импульса – имеет как раз такую размерность. Так что можно ожидать, что в атоме характерный момент импульса будет порядка  $\hbar$ . Мы увидим, что в микромире важную роль играют произведения типа  $px$  и  $Et$ . (В обычной жизни тоже можно перемножить импульс на длину, но это будет настолько большое количество элементарных единиц  $\hbar$ , вроде  $10^{30}$ , что такая дробность не имеет никакого значения. Аналогично мы не замечаем, что человек состоит из такого же числа атомов). Вообще-то в классике есть величина размерности эрг·с, она называется действием. Есть принцип наименьшего действия, подобный принципу Ферма в оптике. Но эти вопросы выходят за рамки данного курса.

<sup>2</sup>Точнее, Планк ввел несколько другую постоянную  $h = 2\pi\hbar = 6,626 \cdot 10^{-27}$ , так что энергия записывалась в стиле XIX века как  $h\nu$ .

**Фотоэффект.** Планк был крайне осторожен, описывая свои результаты, и подчеркивал их формальный характер. Действительно, получалось неудобно. Световые волны, как бы эталон непрерывности, делились на частицы. Неожиданно идея квантов нашла подтверждение в теории фотоэффекта Эйнштейна.

Фотоэффект – процесс, в котором свет выбивает электроны из металла. Сейчас на основе фотоэффекта работает масса приборов (хотя бы пропускники в метро). Оказалось, что электроны вылетают хорошо при освещении коротковолновым светом. Если же светить длинноволновым, то даже при большой интенсивности никакого фотоэффекта нет. Для большинства материалов электроны не выбиваются красным светом (красная граница фотоэффекта). Этот удивительный факт оказался достаточен для Эйнштейна. Он попробовал буквально понять идею Планка и записать закон сохранения энергии:

$$\frac{mv_{max}^2}{2} = \hbar\omega - A.$$

Квант света с энергией  $\hbar\omega$  попадает в металл и передает энергию электрону. Тот часть энергии тратит на преодоление работы выхода  $A$  (для каждого металла своей), а остальное сохраняет в виде кинетической энергии. Если правая часть меньше нуля, то есть частота света мала, никаких электронов не вылетает, как ни свети<sup>3</sup>. Индекс  $max$  обозначает, что кинетическая энергия бывает и меньше, чем  $\hbar\omega - A$ . Энергии электронов в металле различны; работа  $A$  необходима для выбивания самых энергичных «верхних» электронов, а для большинства, энергия которых ниже, нужна и бóльшая работа. Чаще попадают электроны еще меньшей энергии, которых квант вообще не выбивает. Отношение числа электронов к числу фотонов (квантовый выход) меньше единицы, и прибор считается высокочувствительным при выходе 0,1 – 0,4. Милликен поставил точнейшие эксперименты (последние варианты установки напоминали механическую мастерскую в вакууме); все опыты отлично объяснялись этим простым уравнением. По наклону зависимости энергии от частоты можно независимо найти значение постоянной Планка. Оно совпадает с величиной, полученной из теплового излучения.

Следовательно, кванты света – фотоны – существуют **на самом деле**, а не как деталь формального описания излучения нагретых тел.

**Кванты – реальность.** Еще более наглядно квантовый характер света проявился в опытах со слабыми источниками. Оказалось, что при ослаблении светового потока датчики начинают «пицать» не непрерывно, а импульсами, соответствующими единичным фотонам. Человеческий глаз не реагирует на отдельные фотоны видимого света, но может уловить группу из десятка фотонов. Будь глаз человека на порядок чувствительнее, вопрос о природе света был бы ясен уже сотни тысяч лет. Интересно, что дискретность света ближе к нашему порогу ощущений, чем дискретность вещества<sup>4</sup>.

<sup>3</sup>Издредка электрон успевает воспользоваться энергией второго «красного» кванта, пока не растерял энергию первого, и вылететь из металла. Но вероятность таких двухфотонных процессов очень мала.

<sup>4</sup>Некоторые ночные животные, например, лягушки, способны реагировать на отдельные кванты.

Значит, свет – электромагнитная волна – чем-то похож на частицу. Позднее выяснилось, что и «натуральные» частицы, вроде электрона, тоже немного волны. Все это вместе теперь называется квантами. Фотон – квант света, электромагнитного поля. А электрон – квант такого лептонного поля.

Планк получил Нобелевскую премию в 1918 г., Эйнштейн – в 1921 (причем в основном за фотоэффект, а не за теорию относительности).

## 5.2 Строение атома. Атомное ядро

Как мы помним, все состоит из атомов. Остается вопрос, как устроены сами атомы. К началу нашего века некоторые из деталей уже были известны. Электроны – переносчики тока – были обнаружены в свободном состоянии (так называемые катодные лучи, Дж. Дж. Томсон, 1897)<sup>5</sup>. По отклонению в магнитном поле выяснилось, что у них отрицательный заряд и очень малая масса (по сравнению с любым атомом). Поскольку атом нейтрален, должны присутствовать и положительные заряды. Исторический интерес представляет первая разумная модель атома, также изобретенная Дж. Дж. Томсоном (1903). Чтобы с чего-то начать, рассмотрим эту модель.

**Атом Томсона.** В таком атоме электроны сидят в потенциальной яме, образованной положительным зарядом. Скажем, атом водорода – это заряд (+e), размазанный равномерно по шару радиуса  $a$  порядка  $10^{-8}$  см. Внутри помещается точечный электрон. Он предпочитает находиться в центре, но если ему придать очень много энергии, он выскочит наружу и даже может улететь от своего плюса. Для этого требуется энергии порядка  $e^2/a = e\varphi$ , где  $\varphi \simeq e/a$  – характерный потенциал внутри такого атома. В СГС  $e/a \approx 4,8 \cdot 10^{-10}/10^{-8} \approx 4,8 \cdot 10^{-2}$ , и если умножить на 300, получим 14,4 В. Энергия  $e\varphi$  тогда будет около 14 эВ. По измерениям, потенциал ионизации водорода (13,6 эВ) очень близок к этой оценке. Для других атомов получится величина того же порядка. Единица эВ – электронвольт, как видно, характерна для атомного мира.

При меньшей избыточной энергии электрон может колебаться. Как мы помним, внутри равномерно заряженного шара электрическое поле растет линейно с радиусом  $r$ ,  $E = er/a^3$  (как и ускорение силы тяжести внутри однородной планеты). Сила  $eE$  пропорциональна смещению, поэтому колебания электрона будут гармоническими, с частотой  $\omega = \sqrt{e^2/ma^3} \approx (4,8 \cdot 10^{-10}/10^{-8})/\sqrt{10^{-27} \cdot 10^{-8}} \approx 10^{16}$  1/сек. Такой должна быть частота излучаемого атомом света. Это опять похоже на правду – длина волны

---

<sup>5</sup>Во избежание путаницы: Джозеф Джон Томсон (1856–1940, Нобелевская премия 1906) – младший однофамилец Уильяма Томсона (1824–1907, с 1892 за научные заслуги – барон Кельвин, под которым именем он более известен). И Кельвин, и Дж. Дж. Томсон были президентами Лондонского Королевского общества (1890–1895 и 1915–1920 соответственно). Был еще Дж.П. Томсон (1892–1975), сын Дж.Дж., обнаруживший в 1927 г. дифракцию электронов независимо от К. Девиссона и Л. Джермера (Нобелевская премия 1937, совместно с К. Девиссоном).

$\lambda = 2\pi c/\omega \approx 2 \cdot 10^{-5}$  см попадает в область ближнего ультрафиолета. Поскольку оценки довольно грубые, то можно ожидать оптических проявлений и в видимом свете.

Тем не менее, модель Томсона не годится. Один из основных ее недостатков – как раз жестко заданная частота излучения атома. Даже атом водорода излучает вовсе не одну линию. Наоборот, спектр его довольно сложен, хотя его удалось описать так называемыми сериальными формулами. К 1900 году число известных линий (частот) водорода насчитывалось десятками, а для других веществ – тысячами. Кроме того, оставался открытым вопрос, что удерживает положительный заряд от рассыпания.

**Радиоактивность.** В 1896 г. Беккерель открыл радиоактивность. Оказалось, что некоторые вещества испускают «лучи», которые классифицировали в три вида:  $\alpha$ ,  $\beta$ ,  $\gamma$ . Лучи  $\alpha$  – частицы с массой 4 атомных единицы и зарядом 2. Это установлено по отклонению в магнитном поле – в камере Вильсона. Как теперь известно, это ядра гелия. Лучи  $\beta$  отклоняются в другую сторону (то есть их заряд отрицательный) и оказались попросту энергичными электронами. Наконец,  $\gamma$  не отклонялись магнитным полем и напоминали ранее открытые рентгеновские лучи. Как раз они в основном засвечивали фотопластинки Беккереля, даже помещенные в защитную оболочку. Характерная энергия лучей  $E$  оказалась порядка **миллиона** электронвольт. Энергия оценивалась по кривизне траектории: радиус окружности  $R = pc/eH$ . Импульс  $\alpha$  – частицы  $p = \sqrt{2mE}$ , и  $R = \sqrt{2mc^2E}/eH$ . Удобно считать в электронвольтах:  $R = \sqrt{2 \cdot 940 \cdot 1} \cdot 10^6 / (10^3 \cdot 300) \simeq 140$  см при энергии  $E = 1$  МэВ и магнитном поле  $10^3$  эрстед. Для электронов получится порядка сантиметра. Такую кривизну нетрудно измерить, так что сомнений в величине энергии не было.

Энергия в диапазоне МэВ означает, что электрон или  $\alpha$ – частица перед вылетом разогнались разностью потенциалов порядка  $10^6$  вольт. Ясно, что в модели Томсона такая энергия ниоткуда не возьмется.

**Радий.** Невидимые лучи и почерневшие фотопластинки могут казаться далекими от практики. Но радиоактивность можно буквально пощупать. Радий (от лат. radius – лучистый) был открыт супругами Кюри как интенсивно излучающая примесь к урану. Радий и его соли светятся в темноте и заметно теплее окружающей среды. Грамм радия выделяет около 130 калорий в час, что достаточно для расплавления 1,6 г льда<sup>6</sup>. Что его нагревает? Стали сомневаться в законе сохранения энергии. Со временем выяснилось, что радий расходуется, хотя и медленно: период его полураспада около 1600 лет. Поэтому в течение года или десятка лет «на вид» с ним ничего не происходит. Но каждый распад высвобождает так много энергии, что препарат греется без видимых изменений в веществе. Сейчас из некоторых изотопов (как  $^{238}\text{Pu}$ ) делают тепловые источники, например, для питания аппаратуры в многолетнем космическом полете.

---

<sup>6</sup>Лучше, конечно, не трогать радий руками: смертельная доза радиоактивности соответствует поглощению  $\simeq 1,5$  калории на килограмм веса организма. К настоящему времени добыто всего около килограмма радия; актуальность его сильно упала с появлением множества искусственных изотопов.

**Структура атома.** Как «получить» из атома энергию побольше? Либо нужны взаимодействия другой природы, гораздо сильнее электрических, либо, если энергия  $e^2/a$  мала, то надо уменьшить размер. Чтобы получить МэВ вместо 10 эВ, надо взять размер на 5 порядков меньше, то есть около  $10^{-12}$  см. Оставим обозначение  $a$  для размера атома, а этот гипотетический малый размер будем обозначать  $r$ . Если в атоме есть настолько маленькие детали, в которых размещается заряд, то из них вполне могут вылетать излучения нужных параметров.

Такие детали нашлись (это и есть ядро атома), но нашлись и другие взаимодействия (сильные и слабые). Например,  $\beta$ -лучи – продукт слабых взаимодействий, типа распада нейтрона. В начале века об этом думать было рано, и искали «электромагнитное» решение. По современным представлениям,  $\alpha$ -радиоактивность вызвана совместно сильным и электромагнитным взаимодействием, причем конечная энергия  $\alpha$ -частиц – это в основном результат электромагнитного (кулоновского) отталкивания их от ядра.  $\gamma$ -лучи сами имеют электромагнитную природу (энергичные фотоны), излучаются они в ядерных переходах, энергии которых определяются в основном сильным взаимодействием частиц в ядре. Найдено еще несколько (более редких) типов радиоактивности.

**Опыт Резерфорда.** В 1911 г. Э. Резерфорд открыл атомное ядро. Это – важнейшее экспериментальное открытие XX века. Исследовалось отклонение в веществе  $\alpha$ -частиц. Имеющие большую энергию  $\alpha$ -частицы использовались как снаряды, способные проникать в атомы. Поскольку  $\alpha$ -частицы заряжены, они сильно взаимодействуют с веществом и задерживаются в довольно тонких слоях. Поэтому в опыте Резерфорда частицы пропускались через тонкую золотую фольгу.

Ожидаемое отклонение частиц было небольшим. Если в атоме действуют поперечные силы порядка  $F \sim e^2/a^2$ , то за время пролета  $t = \Delta/V$  фольги толщиной  $\Delta$  может набраться поперечная скорость  $V_{\perp} = (F/m)t = e^2\Delta/(a^2Vm)$ . Угол отклонения

$$\theta \sim \frac{V_{\perp}}{V} = \left( \frac{e^2}{amV^2} \right) \left( \frac{\Delta}{a} \right).$$

Первый множитель – отношение характерной атомной единицы энергии  $e^2/a$  к кинетической энергии  $\alpha$ -частицы – это отклонение, набираемое на одном атоме. Второй – количество атомов на пути частицы. Численно получаем  $\theta \sim (10/10^6) \cdot 10^4 = 0,1$  для слоя толщиной  $10^4$  атомов (1 мкм). Это завышенная оценка. Правильнее считать, что не все атомы отклоняют в одну сторону, а скорее их воздействия случайны. Вместо  $(\Delta/a)$  надо ставить  $\sqrt{\Delta/a}$ , как учит статистика, и реальное отклонение будет порядка  $10^{-3}$ .

**Атомное ядро.** Примерно такие отклонения и наблюдались. Однако Резерфорд предложил своим сотрудникам Гейгеру и Марсдену проверить, не будут ли, хотя бы изредка, происходить большие отклонения. И действительно, оказалось, что примерно одна частица из 10000 поворачивает на угол больше  $90^\circ$ . Хотя Резерфорд сравнивал такое событие с отражением артиллерийского снаряда от листка бумаги, догадаться о



возможности больших отклонений можно было хотя бы из большой энергии  $\alpha$ -частиц. Они тоже вылетают из атомов, и должны откуда-то брать свои несколько МэВ. Значит, в атоме есть область, в которой действуют огромные силы. А поскольку большие отклонения редки, эта область очень маленькая. Ее и назвали **атомным ядром**.

Рассеяние на большие углы оказалось «кулоновским», то есть совпадающим с рассеянием на неподвижном точечном заряде  $Ze$ , где  $Z$  – порядковый номер атома мипени (как известно сейчас, это число протонов в ядре). Эта задача решается точно ( $\alpha$ -частица летит по гиперболе), и надо только перевести эти траектории в отсчеты датчика (**резерфордовское сечение взаимодействия**). На больших углах воздействие атомных электронов несущественно из-за его слабости и компенсации вклада попутных атомов. Заряд ядра положительный, он компенсируется отрицательным зарядом  $Z$  электронов.

Мы видим, что  $\alpha$ -частицы дали возможность зондировать атом с хорошим разрешением. Для энергии  $E_\alpha = 1$  МэВ при отскоке назад минимальное сближение с рассеивающим центром, или ядром,  $r$  найдем, сравнивая энергии:  $2Ze^2/r = E_\alpha$ , то есть  $r \sim 2Za \cdot (e^2/a)/E_\alpha = 10^{-5} \cdot 2Za \simeq 10^{-13}$  см при небольших  $Z$ . Многочисленные эксперименты показали, что ядро имеет радиус примерно  $10^{-13} \cdot A^{1/3}$  см, где  $A$  – массовое число атома (атомный вес, то есть на современном языке сумма количеств протонов и нейтронов). Если атом увеличить до размера 1 км (на 13 порядков), то ядро будет по размеру как вишня для легких атомов ( $10^{-13}$  см  $\Rightarrow$  1 см) или как большое яблоко ( $10^{-12}$  см  $\Rightarrow$  10 см) для тяжелых. Сечение взаимодействия ядра с энергичными частицами удобно измерять в барнах ( $10^{-24}$  см<sup>2</sup>).

Таким образом, атом оказался совершенно не похож на томсоновскую модель. Практически вся его масса сосредоточена в очень маленьком ядре. Это объясняет, откуда берется огромная энергия при радиоактивности (хотя бы кулоновское отталкивание почему-либо оторвавшегося фрагмента ядра), но зато ставит новые вопросы.

## 5.3 Атом Бора

Первый из этих вопросов – как может быть устроен неоднородный атом? Мало кто способен вникнуть в данные по рассеянию  $\alpha$ -частиц. А вот если из них удастся получить удобопонятную модель атома, которую потом будут рисовать во всех учебниках, на значках и логотипах, то такой результат уже влияет на все общество.

**Модель Резерфорда–Бора.** Положительное ядро маленькое, значит размер атома обеспечивается отрицательными электронами. Ясно, что электроны не могут висеть неподвижно над притягивающим ядром. Модель Резерфорда – Бора использовала аналогию с Солнечной системой – электроны вроде планет, а ядро – как бы Солнце. Вращение вокруг ядра удерживает электроны на орбитах. Но отсюда возникают два новых вопроса:



1. Раз электрон летает вокруг ядра, он имеет ускорение. В отличие от планет, электрон заряжен и должен излучать энергию. Мощность излучения заряда  $e$ , имеющего ускорение  $w$ , можно оценить по размерности:  $N \sim e^2 w^2 / c^3$ . Потери на излучение приведут к гибели атома – падению электрона на ядро через некоторое время  $t$ . Оцениваем  $N$  как  $e^2 / at$ , а  $w$  как  $v^2 / a = e^2 / ma^2$ , где  $a$  – размер атома и  $m$  – масса электрона:  $e^2 / at \sim e^6 / m^2 a^4 c^3$ , откуда  $t \sim a^3 c^3 m^2 / e^4$ . То же можно записать как  $t \sim (a/c) \cdot (mc^2 / (e^2/a))^2$ . Множитель  $a/c \sim 10^{-18}$  секунд. В скобках – отношение энергии покоя электрона 0,5 МэВ к химической энергии 10 эВ, то есть около  $5 \cdot 10^4$ . Итого выходит  $t \sim 10^{-9}$  с. За такое время все электроны должны упасть на ядро. Тем не менее атомы, из которых состоим и мы, выглядят гораздо более стабильными<sup>7</sup>.
2. Орбиты планет не заданы жестко. Если и найдут планетную систему, похожую на нашу, никто не ждет, что орбиты будут точно такими же. Скорее, будет вполне заметная разница. Для любого радиуса орбиты можно подобрать скорость обращения, и планета будет летать сколько угодно. А вот атомы (данного элемента, данного изотопа) **все одинаковы абсолютно**. Получается, что какая-то орбита выделена, и по ней-то электрон летает.

**Постулаты Бора.** Около 1913 года Нильс Бор довел «планетарную» модель до некоторой завершенности. Упомянутые трудности он разрешил, введя так называемые постулаты Бора:

- В атоме существуют стационарные состояния, в которых электрон не излучает.
- При переходе между состояниями излучается энергия, равная разности энергий состояний.

В какой-то степени можно объяснить, почему дела обстоят таким образом. Падение электрона, сопровождаемое излучением, процесс по атомным масштабам медленный (вспомним оценки частоты излучения –  $10^{16}$  1/с). Излучение мало-помалу невозможно, так как частота излучения какая-то определенная, значит энергия кванта тоже конечна. А половину или десятую часть кванта излучить нельзя. Поэтому электрон способен перескочить в состояние с заметно (на  $\hbar\omega$ ) отличающейся энергией, но **не может** изменять энергию **непрерывно** и даже малыми дискретными порциями.

Вместо полного объяснения Бор скорее констатировал состояние дел. Это был разумный подход, так как до хорошего понимания устройства атома и вообще новой физики прошли десятилетия.

---

<sup>7</sup>Еще гораздо быстрее должен погибнуть сложный атом из-за взаимодействия электронов. В отличие от планетной системы, силы между электронами примерно такие же, как между электроном и ядром. Поэтому устойчивые орбиты (в рамках классической механики) уже в атоме гелия невозможны.

Но остается еще вопрос, как же найти «правильные» состояния? Для этого Бор предложил правило отбора:

- На стационарных орбитах момент импульса электрона равен  $n\hbar$ , где  $n$  – целое число,  $\hbar$  – постоянная Планка.

Так квантовая постоянная впервые была употреблена для выяснения устройства атома.

Мы увидим, что постулаты Бора в общем правильны. Этого нельзя сказать о правиле отбора (в частности, в основном состоянии, при котором энергия минимальна, момент импульса равен нулю, а не  $\hbar$ ; вовсе не момент определяет энергию состояния в атоме водорода). Сейчас можно сказать, что основные заслуги Бора – это концентрация на эксперименте, а не дефектах теории (от него пошла развитая концепция наблюдаемости) и применение квантовой постоянной  $\hbar$  к атому (хотя буквально и неправильное). Удивительно, что, несмотря на такую ошибку, теория Бора позволила рассчитать размеры и свойства (то есть спектр) атома водорода в полном согласии с экспериментом. И сейчас физики для быстрого вычисления состояний применяют уравнения Бора.

**Расчет атома.** Покажем, как работают эти уравнения. На круговой орбите радиуса  $a$

$$mv^2/a = e^2/a^2, \quad mva = n\hbar.$$

Два неизвестных  $a$  и  $v$  отсюда находятся:  $v = e^2/n\hbar$ ,  $a = n^2\hbar^2/me^2$ . Энергия электрона

$$E_n = \frac{mv^2}{2} - \frac{e^2}{a} \equiv -\frac{mv^2}{2} \equiv -\frac{e^2}{2a} = -\frac{me^4}{2\hbar^2} \cdot \frac{1}{n^2}.$$

Минимальная энергия – при  $n = 1$ . Это состояние электрона называется **основным**. Для него  $v = e^2/\hbar$ ,  $a = \hbar^2/me^2$ ,  $E = -me^4/2\hbar^2$ . Подставляя значения, имеем  $a = 0,53 \cdot 10^{-8}$  см,  $E = -2,2 \cdot 10^{-11}$  эрг =  $-13,6$  эВ. Действительно, размер атома получился правильным. То же можно сказать и о энергии электрона, которую можно проверить по затратам на ионизацию атома.

Вспомнив постулаты Бора, найдем спектр водорода. Атом излучает при переходе электрона в состояние с меньшей энергией. Например, при падении с уровня  $n$  на нижний, первый уровень, энергия излучения  $\hbar\omega_{n1} = E_n - E_1$  равна  $13,6 \cdot (1 - 1/n^2) = 10,2; 12,09; 12,75; 13,06; 13,22; 13,32; \dots$  эВ при  $n = 2, 3, 4, 5, 6, 7 \dots$

Действительно, у водорода есть такие линии в спектре, а вся последовательность называется серией Лаймана. Даже минимальная частота серии Лаймана попадает в ультрафиолетовую область (максимальная длина волны 1220 Å), так что эти линии не наблюдаются в видимом свете.

При падении на уровень 2 получим  $\hbar\omega_{n2} = E_n - E_2 : 13,6 \cdot (1/2^2 - 1/n^2) = 1,89; 2,55; 2,86; 3,02; 3,12; \dots \Rightarrow 3,4$  эВ при  $n = 3, 4, 5, 6, 7 \dots$

Это – серия Бальмера. Заметная часть этой серии (простирающейся от 6560 до 3650 ангстрем) соответствует видимому свету, почему она и стала известна раньше. И снова

получилось точное соответствие измерениям. Теория Бора, аккуратно предсказывающая такие замысловатые зависимости, заслужила Нобелевскую премию уже в 1922 г.

Нагретый газ излучает линии при падении электронов с верхних состояний на нижние. Наоборот, из нижнего состояния электрон может подняться на верхнее, поглотив соответствующий квант. Тогда будет наблюдаться спектр поглощения, подобный спектру испускания, но «негативный» – темные линии на светлом фоне. Его можно наблюдать, например, когда свет Солнца проходит через его верхнюю атмосферу, сравнительно холодную.

Остановимся еще на скорости электрона в атоме. В нижнем состоянии  $v = e^2/\hbar = 2,2 \cdot 10^8$  см/с. Полезно записать  $v$  в виде  $v = c \cdot (e^2/\hbar c) = c \cdot (1/137)$ . Видно, что электрон нерелятивистский. Безразмерная величина  $e^2/\hbar c$  называется постоянной тонкой структуры. Удивительно, что ее обратная величина – почти целое число: 137,036...

С ростом  $n$  атом увеличивается в размерах, энергия электрона стремится к нулю (снизу), скорость замедляется. Из космоса к нам иногда приходят кванты «света» от атомов водорода, у которых радиус орбиты чуть ли не в сантиметр. В земных условиях, разумеется, такой атом долго не протянет.

**Условия квантования Бора–Зоммерфельда.** Позднее, в 1914 г., Зоммерфельд рассмотрел эллиптические орбиты. Эллипс задается двумя параметрами, например энергией и моментом импульса. Значит, надо два условия типа боровских – на радиальное и угловое движение.

Вращение, конечно, снова задается моментом  $L$ . Но радиальное движение нельзя задать, скажем, импульсом  $p_r$ , так как он колеблется. Обобщение Зоммерфельда выглядит так:

$$\oint L d\varphi = 2\pi l\hbar, \quad \oint p_r dr = 2\pi n_r \hbar \quad (l, n_r - \text{целые}).$$

Первый интеграл, так как  $L$  сохраняется, эквивалентен равенству  $L = l\hbar$ . Второй уже нетривиален; применение его к атомным орбитам изложено в Приложении 1. Подобные интегралы от импульса по соответствующей координате, взятые по периоду движения, уже были известны физикам – это так называемые адиабатические инварианты. В таком виде условия квантования годятся для любого периодического движения.

**Квантование осциллятора.** Покажем работу условий квантования на примере осциллятора – массы на пружинке. Классическое решение имеет вид  $x = A \sin(\omega t)$ ,  $p = mA\omega \cos(\omega t)$ . Интеграл Зоммерфельда будет

$$\oint p dx = mA^2\omega \cdot \cos(\omega t) d \sin(\omega t) = mA^2\omega^2 \cdot \cos^2(\omega t) dt = mA^2\omega^2 T/2,$$

где  $T$  – период колебаний. Так как  $A\omega$  – это амплитуда скорости, то  $mA^2\omega^2/2$  – это энергия колебаний  $E$ . Получаем

$$\oint p dx = ET = 2\pi n\hbar,$$

так что уровни энергии в параболической потенциальной яме:  $E_n = n\hbar\omega$ . Как говорят, колебательные уровни **эквилидистантны** – соседние отличаются друг от друга на одинаковую ступеньку  $\hbar\omega$ . Колеблющаяся молекула в принципе способна излучать кванты, соответствующие перескоку с данного уровня на любой нижележащий. Но переходы на отдаленные уровни,

для которых энергия кванта была бы  $2\hbar\omega$ ,  $3\hbar\omega$ ..., происходят крайне редко (как говорят, они «запрещены»), а типичны переходы между соседними уровнями. Поэтому в спектре каждое колебание дает одну-единственную линию, а не «гребенку», что и облегчает анализ колебательных спектров<sup>8</sup>.

Точное решение задачи об осцилляторе дает  $E_n = (n + 1/2)\hbar\omega$ . Поправку  $1/2$  мы обсудим ниже. Можно сказать, что условия Бора-Зоммерфельда тем точнее, чем больше номер состояния  $n$ .

## 5.4 Волны де Бройля. Принцип неопределенности Гейзенберга

Схема Бора явно противоречива. Следующий решающий сдвиг в умах произвел Луи де Бройль в 1923 г.

**Волны де Бройля.** Раз Планк и Эйнштейн не побоялись приписать свету дискретность, свойственную частицам, хотя бы пулям или крупинкам соли, то де Бройль решил, что возможно и обратное: частицы имеют волновые свойства. Будем исходить из аналогии. Известно, что для света  $E = \hbar\omega$ , а  $\omega = 2\pi c/\lambda$ . В то же время (из теории относительности)  $E = pc$ . Тогда импульс светового кванта  $p$  можно связать с длиной волны:  $p = 2\pi\hbar/\lambda$ . Но это выражение не содержит никаких следов того, что дело идет о свете: здесь выпала скорость света  $c$ . А почему бы не выполнять тому же и для электрона:  $\lambda = 2\pi\hbar/p = 2\pi\hbar/mv$ ? Оценим характерную длину волны. Кинетическая энергия электрона в атоме водорода  $p^2/2m = 13,6$  эВ. Тогда  $p = \sqrt{2 \cdot 13,6 \cdot 1,6 \cdot 10^{-12} \cdot 9 \cdot 10^{-28}} \approx 2 \cdot 10^{-19}$ ;  $\lambda \approx 3 \cdot 10^{-8}$  см. Выходит близко к атомным размерам. Если еще вспомнить, что сама  $\lambda$  – величина не очень характерная, а более интересна  $\lambda/2\pi \approx 0,5 \cdot 10^{-8}$  см, то это попросту боровский радиус атома!

**Эксперимент Девиссона и Джермера.** Проверить, правда ли электрон – волна, можно в дифракционных экспериментах. В 1927 г. Девиссон и Джермер поставили опыт, в котором пучок электронов отражался от поверхности металла. При энергии в десятки эВ длина волны порядка межатомных расстояний, так что металл работает как отражательная дифракционная решетка. Оказалось, что электроны успешно дифрагируют, и кроме главного «зеркального» максимума наблюдаются боковые в полном соответствии с формулами де Бройля. В том же году Дж.П. Томсон получил дифракцию электронов на тонкой металлической фольге.

Позднее волновые свойства прямо наблюдались и у атомов (главным образом легких – водорода и гелия). Чем тяжелее атом, тем короче его длина волны при данной скорости или энергии, тем труднее ее измерить. Мы с вами тоже имеем волновые свой-

<sup>8</sup>Гребенки в спектрах тоже бывают, но они получаются из-за вращательных переходов (вращательные уровни не эквидистантны).

ства, но на человеческих масштабах их трудно заметить, так как длины волн будут крайне малы ( $10^{-27}$  см при импульсе  $1 \text{ г}\cdot\text{см}/\text{с}$ ).

**Стационарные состояния.** В духе концепции Де Бройля можно вернуться к истолкованию правила квантования Бора. Среди всех возможных орбит особенно важны те, на которых длина волны укладывается целое число раз. Тогда, повернув атом на  $2\pi$ , мы не получим никаких изменений. Можно сказать, что на такой орбите электрон «интерферирует сам с собой».

Это и есть стационарные состояния. Представим себе электрон на круговой орбите радиуса  $a$ . Чтобы на ней хорошо уместилась волна, должно выполняться равенство  $2\pi a = n\lambda$ , или  $pa = n\hbar$ . По форме это не отличается от исходного правила Бора, так что ничего пересчитывать не надо. Но смысл совершенно другой:  $n$  это не величина момента импульса, а число волн, укладывающихся на орбите.

Поэтому электрон и не падает на ядро. Волна не может стянуться в точку. Хотя бы одна длина волны должна размещаться на самой нижней орбите. Волновая природа электрона объясняет также одинаковость атомов. Заметим, что и «пространственность» атома теперь можно понять лучше. Волна имеет протяженность во всех направлениях, и если мы еще рассуждали здесь о плоских орбитах, то больше по инерции. Но правда и то, что эти объяснения все еще слишком приблизительны.

**Принцип неопределенности.** Волновые представления привели Гейзенберга к формулировке принципа неопределенности (1927). Именно, к электрону неприменимо понятие траектории. И действительно, какая может быть траектория у волны? **Соотношение неопределенностей** Гейзенберга выражает эту невозможность количественно. Пусть плоская волна (все равно, свет или электроны), распространяется в направлении  $x$ . Монохроматическая волна вовсе не локализована в пространстве, поэтому рассматриваем волновой пакет, составленный из волн с немного различающимися волновыми числами: в интервале порядка  $\Delta k$  возле среднего значения  $k$ . Импульс монохроматической волны  $p = 2\pi\hbar/\lambda = \hbar k$ . Тогда импульс пакета имеет неопределенность  $\Delta p_x \simeq \hbar\Delta k$ . В п. 2.1 мы видели, что неопределенность  $x$  – компоненты пакета, то есть его ширина  $\Delta x$ , порядка  $1/\Delta k$ . Из очевидного соотношения  $\Delta k \cdot \Delta x \sim 1$ , если де Бройль прав, следует

$$\Delta p_x \cdot \Delta x \sim \hbar. \quad (5.1)$$

Пропустим волну через щель в экране шириной  $d$ . Как и в оптике, после экрана получится пучок, расходящийся под углом  $\lambda/d \sim \hbar/pd$ . Этот же угол можно записать как  $p_y/p$ , то есть в расходящемся пучке появляется поперечный импульс, по порядку величины равный  $p_y = \hbar/d$ . Одни электроны попадут в экран примерно напротив щели, другие отклонятся вверх или вниз. Так как пучок расходится симметрично, значит для него теперь характерна неопределенность поперечного импульса  $\Delta p_y \sim \hbar/d$ .

Пропуская электрон через щель, мы «измерили» его поперечную координату с точностью  $d$ , то есть неопределенность положения стала  $\Delta y = d$ . Отсюда имеем оценку

неопределенности, совпадающую с (5.1):  $\Delta p_y \cdot \Delta y \sim \hbar$ .

Получается, что у электрона нельзя одновременно задать координату и соответствующий ей импульс. Если фиксировать координату, например, пропуская его через узкую щель, неизбежно «разбегается» импульс поперек этой щели. Если фиксировать электрон в точке (другими словами, точно измерить координаты), импульс его станет полностью неопределенным. Это – следствие волновой природы электрона, общее для всех волн. Электрон не имеет траектории (для построения которой необходимо в каждый момент знать и координаты, и все компоненты импульса).

Но ведь в камере Вильсона электрон оставляет явственный след, искривление которого в магнитном поле – стандартный способ измерения скорости и импульса? Опять дело в чрезвычайной малости постоянной Планка. Если толщина следа электрона (цепочки капель воды) 1 мкм =  $10^{-4}$  см, то  $\Delta p \sim 10^{-27}/10^{-4} = 10^{-23}$  в СГС. Импульс же в продольном направлении при энергии 1 МэВ  $p \sim \sqrt{2mE} = \sqrt{2 \cdot 0,9 \cdot 10^{-27} \cdot 10^6 \cdot 1,6 \cdot 10^{-12}} \simeq 5 \cdot 10^{-17}$  СГС на 6–7 порядков больше. Поэтому траектория как бы и есть, пока она достаточно «толстая». Но если переходить к истинно малым (для такой энергии – ядерным) масштабам, от траектории ничего не останется.

Поучительно применение соотношения неопределенностей к оценке размера атома. Поскольку электрон локализован внутри атома, для него искомый размер  $a$  есть неопределенность координаты. Тогда электрон должен иметь неопределенность импульса  $\delta p \sim \hbar/a$ . Энергия атома будет

$$E = \frac{p^2}{2m} - \frac{e^2}{a} = \frac{\hbar^2}{2ma^2} - \frac{e^2}{a}.$$

При малых  $a$  преобладает первое, положительное слагаемое; при больших – второе, отрицательное. Найдем размер  $a$ , отвечающий минимальному значению энергии:  $dE/da = -\hbar^2/ma^3 + e^2/a^2 = 0$ , откуда  $a \sim \hbar^2/me^2$ ! Получается то же самое значение: 0,5 ангстрема. Польза от нового «вывода» та, что здесь ни о каких орбитах не было и речи: сразу видно, что размеры атома одинаковы во всех измерениях. Основное состояние атома водорода сферически симметрично.

**Волны или частицы?** Трудно преодолеть прочно усвоенные догмы, вроде того, что электрон – это маленький твердый шарик. Но это надо сделать тому, кто хочет понимать, как устроена природа. Например, в дифракционном опыте электроны попадают, конечно, в определенные места экрана, то есть на каждый электрон отзывается определенный маленький датчик. В этом смысле электрон – частица, неделимый квант. Но в разные места попадать они будут с разной частотой, точно соответствующей распределению интенсивности дифрагирующей волны. И такое распределение интенсивности сохраняется, даже если электроны заведомо летят поодиночке, например, попадая на экран раз в секунду. Значит, электрон все же волна. То же верно для света (фотонов), нейтронов и пр.

Такое свойство назвали **корпускулярно-волновой дуализм**. Упрощенно говоря,



электроны – и волны, и частицы одновременно. Для облегчения восприятия этой противоречивой (с точки зрения макроскопического опыта) ситуации можно рекомендовать следующую картину мира<sup>9</sup>:

1. Существуют только волны различной природы (свет, электроны и т.п.) Никаких частиц не бывает. Те песчинки, пули и планеты, с которыми мы привыкли иметь дело, тоже волны. Но они размазываются крайне слабо, и мы не замечаем у них волновых свойств. Наша неспособность заметить дифракцию пуль, вылетающих из автомата, и есть основание для возникновения концепции частицы – чего-то, что занимает пространство и летит по определенной траектории.
2. Но все существующие волны – квантовые. Они поглощаются, отражаются, рассеиваются, и т.п. исключительно поштучно, индивидуально. Во вспышке света частоты  $\omega$  может быть пять или семь таких порций – квантов с энергией  $\hbar\omega$  каждый, но никогда не 5 целых и 7 десятых. У некоторых квантов есть и другие отличительные признаки (кроме единичности), как заряд или масса у электрона, которых фотон не имеет, но это не «основа», а как бы отличительная раскраска.

В отличие от «частиц» типа камешка или пылинки, электрон нельзя разрезать на куски. Уже это показывает, что квант и частица не одно и то же. Люди изобрели понятие частицы, отправляясь вовсе не от электронов, но от макроскопических тел, размеры которых не существенны в данной задаче. И хотя пули тоже попадают в мишень по одной целой штуке, вполне возможно это нарушить (скажем, разрезать их в полете лучом лазера). Значит, то, что есть общего у квантов и пуль – это скорее случайные и второстепенные признаки. Поэтому проще признать, что частицы – это антропоморфизм. В микромире их попросту нет. Есть квантованные волны. Из комплексов таких волн (атомов) построено вещество.

И все же электроны и протоны принято называть частицами. Нет смысла плыть против течения, но надо понимать условность терминологии и признать, что это слово в применении к электрону значит совсем другое, чем в бытовой речи.

Кроме того, одно дело – принципиальность, другое – практика. Массу задач можно решить, не задумываясь о волновых свойствах электрона и даже представляя его себе блестящим шариком. И раз задача облегчается, надо этим пользоваться, вспоминая о длине волны там, где она реально нужна.

**Парадокс дифракции.** Дифракция электрона на двух щелях совершенно подобна оптической, то есть на экране электронный пучок даст «забор» интерференционных полос. В минимумах засветки интенсивность нулевая, туда электроны не попадают. Но если закрыть любую из щелей, бывшие минимумы оказываются «освещены», туда начинают прилетать электроны! Это ясно показывает, что электрон взаимодействует с

<sup>9</sup>Конечно, не единственно возможную. См. примечание на стр. 81 и Приложение 2.



обеими щелями, а не ограничивается пролетом через какую-то одну. Открываем снова щель, чтобы дать ему дополнительную возможность достичь экрана, а он отвечает черной неблагодарностью и вообще перестает в какое-то место попадать.

Уже в обычной оптике такое явление удивительно. Но там мы могли сослаться на взаимодействие волн, прошедших различные щели (принцип Гюйгенса). В микрофизике положение еще поразительнее тем, что электроны (как и фотоны и пр.) дискретны и даже отдельный электрон дифрагирует. Можно запускать их по штуке, и они со временем нарисуют ту же самую картину. При этом, конечно, каждый электрон делает не весь забор, а маленькое пятно, например, микронных размеров. Наложение же всех пятен, когда набирается статистика, и будет дифракционной картиной.

Как же может электрон взаимодействовать сам с собой? Никак: по современным представлениям, интерферируют не куски электрона, а пути, возможности его движения (скажем, пути через две щели). Очень грубая аналогия – «разумный» электрон, который «ощупывает» пространство перед собой, прикидывает «экономические» затраты для каждого пути (вроде  $S = \int p dx$ ), записывает амплитуды вида  $\exp(iS/\hbar)$ , суммирует эти экспоненты для всех путей, находит квадрат модуля суммы, разыгрывает точку финиша в соответствии с величиной этого квадрата модуля и приходит в такую случайно выбранную, но зависящую от взаимодействия всех потенциальных затрат точку. То же делает фотон;  $S = \int p dx$  – это что-то вроде  $x/\lambda$ , и мы узнаём принцип Ферма: когда по близким путям фазы одинаковы, эти места и называются «луч света»<sup>10</sup>.

## 5.5 Понятие о квантовой механике. Волновая функция\*

Де Бройль в своей гипотезе исходил не только из формальной симметрии, но и из спектра атомов. Он отмечал, что в физике известен только один класс явлений, для которого характерен дискретный спектр, а именно волны. Простейшие примеры такого рода – колебания струны, закрепленной в двух точках, колебания столба воздуха в духовых инструментах и т.д.

Когда работа де Бройля стала обсуждаться, в Цюрихе П. Дебай поручил своему ассистенту Э. Шрёдингеру рассказать о ней студентам и сотрудникам университета. В результате Шрёдингер, пытаясь понять работу сам, ухитрился придать четкий смысл качественным идеям де Бройля.

**Волновая функция.** Раз электрон – волна, он должен описываться чем-то непрерывным в пространстве. Волны на воде задаются высотой  $h(x, t)$ ; электромагнитные волны – компонентами полей  $E$  и  $H$ . То, что «волнуется» у электрона, Шрёдингер назвал волновой функцией и обозначил буквой  $\Psi$ ; сейчас употребляется также название «амплитуда». Ее физический смысл стал ясен несколько позже.

Волновые процессы подчиняются дифференциальным уравнениям. Подход де Бройля, с

<sup>10</sup>Концепцию разумного электрона (без кавычек, в буквальном смысле) отстаивает Р.С. Нахмансон. См. УФН, 2001. Т. 171, №4, стр. 441. Насколько известно, он – единственный приверженец своей теории.

его длиной волны, явно недостаточен: ведь длина волны  $\lambda = 2\pi\hbar/p$  определена хорошо только для свободного движения, когда импульс  $p$  фиксирован. Если же  $\lambda$  сколько-нибудь меняется с координатой (из-за изменения импульса при движении в поле), то в этой самой степени теряет смысл.

**Импульс и энергия в квантовой механике.** Начнем со свободного движения, попытаюсь получить нечто, пригодное для более общего случая. Естественно связать свободное движение с плоской волной:  $\Psi \sim \exp(ikx - i\omega t) = \exp(2\pi ix/\lambda - i\omega t)$ . Импульс электрона с длиной волны  $\lambda$  равен  $p_x = 2\pi\hbar/\lambda = \hbar k$ . Поэтому  $\Psi \sim \exp(ip_x x/\hbar - i\omega t)$ . «Вытащить» импульс из этой формулы можно, взяв производную:  $p_x \Psi = -i\hbar \cdot \partial\Psi/\partial x$ . Производную можно вычислять уже не только от функции, описывающей плоскую волну. Поэтому можно предположить, что выражения

$$\hat{p}_x = -i\hbar \frac{\partial}{\partial x}, \quad \hat{p}_y = -i\hbar \frac{\partial}{\partial y}, \quad \hat{p}_z = -i\hbar \frac{\partial}{\partial z}, \quad \text{или} \quad \hat{\mathbf{p}} = -i\hbar \nabla \quad (5.2)$$

верны всегда, а не только для плоской волны. (5.2) заменяет дебройлевскую связь импульса и длины волны. Чтобы найти импульс, надо подействовать на волновую функцию **оператором**  $\hat{\mathbf{p}} = -i\hbar \nabla$ . Шляпка над буквой напоминает, что оператор – это нечто более сложное, чем число.

Аналогично полагаем, что энергия кванта  $E = \hbar\omega$  не только для световых волн, а и всех прочих. Тогда плоская волна  $\Psi \sim \exp(ikx - i\omega t) = \exp(2\pi ix/\lambda - iEt/\hbar)$ . Отсюда находим правило извлечения  $E$ :

$$\hat{E}\Psi = i\hbar \frac{\partial\Psi}{\partial t}.$$

**Уравнение Шредингера.** В свободном полете кинетическая энергия электрона постоянна:  $p^2/2m = E$ . Это тождество можно теперь записать в виде уравнения:

$$-\frac{\hbar^2}{2m} \cdot \frac{\partial^2\Psi}{\partial x^2} = E\Psi.$$

Если электрон движется в поле с потенциальной энергией  $U$ , и в трех измерениях, естественным обобщением будет

$$-\frac{\hbar^2}{2m} \Delta\Psi + U\Psi = E\Psi, \quad \text{где} \quad \Delta\Psi = \frac{\partial^2\Psi}{\partial x^2} + \frac{\partial^2\Psi}{\partial y^2} + \frac{\partial^2\Psi}{\partial z^2}. \quad (5.3)$$

Это уже нетривиальное дифференциальное уравнение. По имени автора оно называется **уравнением Шредингера** (в данном случае – стационарным).

Стационарность (постоянство энергии электрона  $E$ ) может не соблюдаться. Нестационарное уравнение Шредингера имеет вид

$$i\hbar \frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m} \Delta\Psi + U\Psi. \quad (5.4)$$

Стационарный случай соответствует специальной зависимости волновой функции от времени:  $\Psi(x) \cdot \exp(-iEt/\hbar)$ . Подстановка этого выражения в (5.4) дает стационарное уравнение (5.3).

**Прямоугольный колодец.** Покажем, как решается стационарное уравнение Шредингера, на простейшем примере. Пусть электрон сидит в прямоугольной потенциальной яме с бесконечно

высокими стенками, так что  $U(x) = 0$  при  $0 < x < a$ , и  $U(x) = \infty$  при  $x < 0$  и при  $x > a$ . Тогда внутри ямы выполняется уравнение

$$-\frac{\hbar^2}{2m} \cdot \frac{\partial^2 \Psi}{\partial x^2} = E\Psi.$$

Его решения – это попросту  $\sin(kx)$  либо  $\cos(kx)$ , с любым  $k$ . Задавая  $k$ , мы тем самым определим энергию  $E$ .

Но не все решения годятся. Вряд ли волновая функция (как-то описывающая электрон) может существовать в областях, где потенциальная энергия бесконечна. Подходят те решения, которые обращаются в нуль на стенках ямы, а именно  $\Psi = \sin(\pi nx/a)$ , где  $n$  – целое число. Подставляя в уравнение, получаем  $E = \pi^2 \hbar^2 n^2 / 2ma^2$ . Только при таких выделенных значениях энергии получается разумное решение. Значит, дискретный спектр – следствие граничных условий. Точно так же закрепленная в двух точках струна имеет дискретный спектр собственных колебаний.

**Осциллятор – основное состояние.** Более сложный пример – уравнение Шредингера для осциллятора:

$$-\frac{\hbar^2}{2m} \cdot \frac{\partial^2 \Psi}{\partial x^2} + \frac{kx^2}{2} \cdot \Psi = E\Psi.$$

Найдем решение для нижнего состояния. Ясно, что волновая функция должна убывать с удалением от начала координат, причем симметрично. Если попробовать решение вида  $\Psi \sim \exp(-\beta x^2)$ , то получим

$$-\frac{\hbar^2}{2m} \cdot 4\beta^2 x^2 \exp(-\beta x^2) + \frac{\hbar^2}{2m} \cdot 2\beta \exp(-\beta x^2) + \frac{kx^2}{2} \cdot \exp(-\beta x^2) = E \exp(-\beta x^2).$$

На  $\exp(-\beta x^2)$  можно сократить. Равенство

$$-\frac{\hbar^2}{2m} \cdot 4\beta^2 x^2 + \frac{\hbar^2}{2m} \cdot 2\beta + \frac{kx^2}{2} = E$$

выполняется всегда (что и нужно от решения) при взаимном сокращении двух постоянных слагаемых и двух, содержащих  $x^2$ , то есть  $\beta = \sqrt{km}/2\hbar$ ,  $E = \hbar\sqrt{k/m}/2 = \hbar\omega/2$ . Этим подтверждается поправка к условиям квантования Зоммерфельда – Бора. Найденное решение называют «нулевым колебанием», его энергия – наименьшая возможная для колебательной системы. Точный нуль запрещен принципом неопределенности: нельзя поместить массу в начало координат и одновременно задать ей нулевой импульс.

Следующие функции имеют вид полиномов, попеременно с нечетными и четными степенями  $x$ , умноженных на ту же  $\exp(-\beta x^2)$ , причем первый следующий уровень соответствует полиному первой степени и т.д. Эквидистантность уровней, хотя довольно сложно, но тоже доказывается (см. Приложение 3).

*Задача.* Найдите уровни энергии для  $\Psi \sim x \cdot \exp(-\beta x^2)$  и  $\Psi \sim (x^2 - b) \cdot \exp(-\beta x^2)$ ; во втором случае еще надо найти подходящее значение  $b$ . Показать, что  $\beta$  для этих уровней то же, что и для основного состояния.

**Суперпозиции состояний.** Можно сложить две волновые функции с разными значениями энергии, при желании умножив их на коэффициенты, например:

$$\Psi = \exp(-\beta x^2/2) \exp(-i\omega t/2) + \sqrt{\beta} x \exp(-\beta x^2/2) \exp(-3i\omega t/2).$$

Такая функция – тоже решение нестационарного уравнения Шредингера, следовательно, осциллятор может находиться и в таком состоянии – суперпозиции двух стационарных. Но в этом состоянии не определена энергия: при измерениях мы будем получать  $\hbar\omega/2$  либо  $3\hbar\omega/2$ , с вероятностями (в этом примере)  $2/3$  и  $1/3$ . Средняя энергия будет равна  $5\hbar\omega/6$ , однако именно такое значение не удастся получить ни разу.

Волновые функции с определенными значениями энергии аналогичны единичным векторам в некоем, так называемом гильбертовом, пространстве (имеющем бесконечное число измерений, соответственно бесконечному числу уровней энергии); тогда суперпозиция похожа на разложение произвольного вектора по компонентам. После измерения вектор состояния всегда «смотрит» точно вдоль какой-то оси: получается энергия какого-то определенного уровня, хотя заранее неизвестно, какого. Вторичное измерение всегда даст уже именно эту энергию. Можно складывать состояния с различными моментами импульса, импульсами и пр., как бы используя разные системы координат. Например, при прохождении плоской волны через щель в экране шириной  $d$  из состояния с нулевым поперечным импульсом готовится сумма (суперпозиция) состояний с различными значениями поперечного импульса, лежащими в интервале около  $\hbar/d$ , или, что то же, выйдет набор плоских волн, отклоненных на углы в пределах  $\lambda/d$ . Хотя это новое состояние известно с максимально возможной определенностью, места попадания в экран будут случайными. Таким образом, источник квантовой случайности – перемешивание нескольких или многих состояний, каждое из которых имеет вполне определенное значение измеряемого параметра.

**Интерпретация волновой функции.** Очевидна трудность согласования размазанной в пространстве волны и точности электрона (попадающего целиком в детектор, даже и маленький). Физический смысл функции  $\Psi$  установил М. Борн. Сначала скажем (не строго), что квадрат модуля  $\Psi$  представляет собой вероятность найти электрон в единице объема в данном месте. Например, для прямоугольного колодца эта вероятность пропорциональна  $\sin^2(\pi nx/a)$  для состояния с номером  $n$ . В атоме вероятность экспоненциально убывает с радиусом. «Размер» атома – это расстояние, на котором заметно убывает вероятность нахождения электрона (как говорят, электронная плотность).

Иногда волновую функцию трактуют слишком примитивно, говоря, что  $\Psi^2\Delta V$  есть вероятность «поймать» электрон в объеме  $\Delta V$ . Это не совсем так. Если мы возьмем кубическую «мышеловку» объема  $L^3$  и запрем там электрон, принадлежащий атому водорода, то при  $L \leq a$  мы совершенно исказим условия, в которых находится электрон. Стенки – это тоже потенциальная яма, которая влияет одновременно с протоном! Например, если электрон пойман далеко от ядра, его потенциальная энергия, очевидно, мала. Кинетическая же энергия всегда положительна. Ясно, что их сумма никак не получится равной большой отрицательной энергии основного состояния ( $-13,6$  эВ).

Ответ в том, что электрон, ограниченный стенками, приобретает добавочную неопределенность импульса и, значит, дополнительную кинетическую энергию. Складываясь с формально отрицательной кинетической энергией вдали от ядра, эта добавка и дает примерно нуль. К подобной поимке электрона сводится по существу любое измерение его координат. Видно, что квантовое измерение – совсем не то, что классическое. При измерении состояние разрушается, и чем точнее измерение, тем разительнее «порча» исходного состояния. Если с нашими гру-

быми инструментами лезть в атом, он наверняка сломается. Особую роль измерения в мире квантов первым осознал Н. Бор.

Правильный смысл  $\Psi^2 \Delta V$  – это не вероятность обнаружить электрон, а вероятность электрону находиться в данном объеме (при условии, что мы не пытаемся это проверить, зачерпывая его каким-то ведром). С помощью распределения вероятности можно считать различные средние, например:

$$\langle U \rangle = \int |\Psi|^2 U \cdot dV.$$

Переход от волновой функции к наблюдаемым величинам обычно и состоит в расчете средних значений.

# ВОЛНОВАЯ ФИЗИКА

Курс лекций для ФМШ

Часть 2. СТРОЕНИЕ ВЕЩЕСТВА

6. АТОМЫ И ЯДРА

А. П. Ершов

12 февраля 2007 г.

# Глава 6

## АТОМЫ И ЯДРА

### 6.1 Атомы. Принцип Паули\*

**Атом водорода.** Для расчета атома водорода надо вставить в уравнение Шредингера потенциальную энергию  $-e^2/r$ . В прямоугольных координатах  $r = \sqrt{x^2 + y^2 + z^2}$ ,

$$-\frac{\hbar^2}{2m} \cdot \left( \frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} + \frac{\partial^2 \Psi}{\partial z^2} \right) - \frac{e^2}{\sqrt{x^2 + y^2 + z^2}} = E\Psi.$$

Шредингер получил решения, и нашел тот же самый спектр водорода, что и Бор. Прямой подстановкой легко проверить, что функция  $\Psi = (1/\sqrt{\pi a^3}) \cdot \exp(-r/a)$  есть решение, если  $a = \hbar^2/m e^2 \approx 0,5$  ангстрема – уже известный нам боровский радиус, при этом энергия  $E = -m^2 e^4 / (2\hbar^2) = -13,6$  эВ. Это состояние называется основным, его энергия соответствует нижнему уровню ( $n = 1$ ), найденному Бором. Однако реальный атом, в отличие от атома Бора, уже не плоский – в основном состоянии он сферически симметричен; кратко это состояние обозначают  $1s$ . Наглядные картинки химиков именно и показывают поверхность, где модуль  $\Psi$  либо его квадрат равны некоторой заданной величине. Для  $s$ -состояния  $\Psi$  не зависит от углов (картинка – сфера).

Только при боровских значениях энергии  $E = -m^2 e^4 / (2n^2 \hbar^2) = -13,6/n^2$  эВ имеются разумные решения, для которых волновая функция спадает на бесконечности, что соответствует попросту тому, что электрон с отрицательной полной энергией нельзя поймать вдали от ядра. Кулоновская яма, хотя ее стенки «мягкие», действует похоже на прямоугольную: дискретный спектр определяется граничным условием. Момент импульса в  $1s$ -состоянии нулевой. Такие же  $s$ -состояния есть для всех уровней энергии, но при  $n > 1$  есть еще и состояния с ненулевым моментом, обозначаемые  $p, d, \dots$ , см. ниже. В рассмотренном приближении момент импульса не влияет на энергию. Однако реально имеется слабое влияние, от чего состояния с  $n > 1$  расщепляются на несколько близко расположенных уровней (так называемая тонкая структура). Для простых атомов, и в особенности для водорода, уточненная теория прекрасно совпадает с экспериментом.

При  $n = 2$  имеются 4 возможных состояния (уравнение Шредингера имеет 4 различных решения). Одно из них – симметричное  $s$ -состояние, с нулевым моментом импульса. Три других – это  $p$ -состояния с моментом  $l = 1$  ( $l$  обозначает момент импульса электрона в единицах



$\hbar$ , то есть  $L = l\hbar$ <sup>1</sup>. Они отличаются проекцией  $m$  момента на ось  $z$ , которая может быть равна  $-1$ ,  $0$  или  $+1$ . Состояние  $(l = 1, m = 0)$  – это хорошо знакомая из учебников химии  $p$  – орбиталь, похожая на «пространственную» восьмерку, вытянутую вдоль оси  $z$ . Состояния с  $m = \pm 1$  напоминают тор, «внутри» которого электрон вращается вокруг оси  $z$  в ту или другую сторону. (Химики вместо таких орбиталей пользуются эквивалентным набором из двух восьмерок, вытянутых вдоль осей  $x$  и  $y$ ). Энергии всех четырех состояний одинаковы (если отвлечься от тонкой структуры).

**Принцип Паули.** Атомы всех элементов, кроме водорода, имеют несколько (и даже много) электронов. Ядро с зарядом  $Ze$  должно удерживать  $Z$  электронов для нейтральности вещества.

Казалось бы, все эти электроны должны «упасть» в основное состояние. С учетом взаимного отталкивания электронов это состояние должно быть мало похоже на «водородное». Однако эксперимент показал, что внешние электроны в атомах щелочных металлов находятся в состояниях, похожих на возбужденные состояния водорода. И вообще, в свойствах элементов прослеживается периодичность, зафиксированная Менделеевым в его таблице. При сваливании электронов в кучу на нижнюю орбиту нет никаких оснований ожидать периодичности.

Для решения этого парадокса В. Паули предложил свой принцип исключения, или принцип запрета (1925). Электроны устроены так, что никакие два из них не могут находиться в одном и том же состоянии. Принцип Паули позволяет разобраться в повторяемости свойств элементов.

Будем мысленно добавлять электроны к исходному ядру. Первый электрон разместится на нижней орбите. Второй уже сюда не пойдет, ему придется летать повыше. Для него ядро с внутренним электроном – это нечто вроде ядра предыдущего элемента, с зарядом  $Z - 1$ . Но тогда гелий должен напоминать по химическим свойствам водород, чего и близко нет. Эту проблему удалось решить, учитывая новую собственную степень свободы электрона. Оказалось, что на внутренней оболочке размещаются все-таки два электрона, различающиеся этой новой переменной (сейчас о ней говорят как о проекции спина). Поэтому гелий содержит два  $s$ -электрона на нижней орбите ( $n = 1$ ). Следующий элемент, литий, еще имеет третий электрон на втором уровне ( $n = 2$ ). Этот электрон он легко отдает, вступая в химическую реакцию. Так последовательно можно описать всю периодическую таблицу. Сама возможность сохранять те же обозначения уровней для многоэлектронных атомов определяется тем, что внутренние электроны экранируют часть заряда ядра в довольно небольшой области пространства (размер нижней орбиты пропорционален  $1/Z$ ). Поэтому для внешних остаются состояния с большими  $n$ , но по-прежнему в поле, похожем на кулоновское.

**Размеры атомов.** Как они зависят от атомного номера? Воспользуемся правилами квантования для электрона в поле заряда  $Ze$ :  $mvR = n\hbar$ ,  $mv^2/R = Ze^2/R^2$ . Отсюда получаем уровни энергии и размер «водородоподобного иона»:  $R = n^2\hbar^2/Zme^2$ ,  $E_n = -me^4/2\hbar^2 \cdot (Z/n^2)$ . При данном  $n$  размер орбиты обратно пропорционален  $Z$ . Поэтому внутренние электроны сидят гораздо ближе к ядру, чем в случае водорода.

---

<sup>1</sup>Точнее, величина момента импульса равна  $\sqrt{l(l+1)}\hbar$ , но для краткости принято называть моментом импульса его максимальную проекцию на выбранную произвольно ось; она равна  $l\hbar$ .

Применяем эти результаты к «среднему» электрону, с характерным значением  $n$ , тогда  $R$  и будет характерным размером атома. Так как внутренние электроны экранируют ядро, на усредненный электрон действует некая доля от заряда  $Z$ , скажем  $Z/2$ .

В таблице Менделеева максимальное значение  $n$  – это номер периода, в который попадает атом; соответствующее  $Z$  будет последним номером элемента в периоде. Всего периодов 7, а элементов помещается около 100; пять периодов содержат 54 элемента, 3 периода – 18 элементов и т.д. Можно предложить корреляцию  $n \sim \sqrt{Z/2}$ . Зависимость  $R \sim n^2/Z$  сокращается: в таком приближении размер атома не зависит от  $Z$ .

Действительно, характерные радиусы всех атомов сильно не отличаются. Хотя есть некоторые колебания размеров внутри периодов таблицы (примерно в 2 раза), систематического роста размера атома с  $Z$  нет. Поэтому плотность веществ в общем растет с атомным весом (хотя влияют еще и детали упаковки в кристаллической решетке).

В рамках боровских представлений следовало бы сказать, что размер атома определяется не средним электроном, а внешним. Если взять щелочной металл с одним внешним электроном, то при экранировке ( $Z - 1$ ) протона остается как бы единственный протон, и размер «по Бору» должен быть примерно как у атома водорода. На самом деле атомные размеры несколько увеличиваются в ряду  $\text{Li} \rightarrow \text{Na} \rightarrow \text{K} \dots$  (например, радиусы по Полингу, используемые в химии: 1,34; 1,54; 1,94 ... Å), и значительно больше, чем размер атома водорода (боровский 0,53 Å, радиус по Полингу 0,34 Å). Здесь и видна разница между планетарной моделью и действительностью. Электрон – волна, и он взаимодействует с внутренними электронами, которые ближе, чем ядро. Выгодно от них несколько оттолкнуться. Для «частицы» такой эффект не работает – на внешней орбите не важно, что там внутри, лишь бы экранировалось.

По этой же причине состояния с одним  $n$ , но различными  $l$  различаются по энергии, особенно во многоэлектронных атомах. На  $s$ -орбите электрон имеет заметную вероятность оказаться во внутренних областях атома, где заряд ядра почти не экранирован, и даже «на ядре». Усиленное в среднем притяжение заставляет его быть ближе к ядру, чем, скажем,  $p$ -электрон, который при ненулевом моменте импульса не может заметно приближаться к центру. Отсюда следует порядок заполнения состояний (т.е. возрастания энергий):  $s, p, d \dots$

**Спин.** О спине следует поговорить подробнее. Постоянная Планка имеет размерность момента импульса и действительно является квантом момента. Опыт показывает, что возможные проекции момента на любую данную ось отличаются на целое кратное  $\hbar$ .

Из симметрии ясно, что максимальная и минимальная проекции равны по величине и противоположны (например, могут быть состояния  $(+\hbar, 0, -\hbar$  для  $p$ -электрона). Но симметрия и квант момента допускают и другой тип состояний, когда проекции не целые. Легко убедиться, что еще возможны только полуцелые значения проекций, например,  $(+3\hbar/2, +\hbar/2, 0, -\hbar/2, -3\hbar/2)$ . Здесь тоже соседние значения отличаются на  $\hbar$ , а крайние – симметричны. Простейший такой случай – это только две возможных проекции  $(+\hbar/2, -\hbar/2)$ . Эти половинки выглядят странно. И в модели Бора, и в действительности орбитальное вращение дает целый момент. Но все-таки вторая возможность осуществляется в природе, только половинка связана не с обращением электрона вокруг ядра, а принадлежит ему самому. Оказывается, что электрон сам по себе имеет свой внутренний момент импульса, равный  $\hbar/2$  (точнее, его проекция  $\pm\hbar/2$ ). Это же выражают, говоря, что спин электрона равен  $1/2$ .

Спин не имеет классического аналога. В частности, неправильно представлять его как результат вращения электрона вокруг оси. Для этого окружная скорость вращения должна превышать скорость света. Действительно, пусть электрон – это шарик размера  $r$ , тогда  $\hbar/2 \sim mvr$ , и  $v \sim \hbar/mr$ . Если подставить вместо  $r$  борковский радиус, получим  $v \sim e^2/\hbar$  – скорость электрона в атоме. Это и не странно, ведь в атоме момент импульса тоже порядка  $\hbar$ . Но электрон заведомо гораздо меньше, чем атом; здесь более подходит классический радиус электрона  $e^2/mc^2 \sim 10^{-13}$  см. Тогда скорость  $v$  на 5 порядков больше, то есть около  $1000c$ .

Исключительно наглядно спин проявился в опыте Штерна – Герлаха (1922). Опыт состоял в пропускании пучка атомов серебра через область неоднородного магнитного поля, в основном поперечного скорости атомов. Магнитный момент атома пропорционален его механическому моменту. Магнитный момент, параллельный полю, втягивается в область сильного поля, антипараллельный – напротив, выталкивается. Пучок атомов с целым моментом расщепился бы на нечетное число компонент (например, для момента 1 – с проекциями на направление магнитного поля  $+1$ ,  $0$  и  $-1$ ). Для проверки этого квантования проекции и предпринимался опыт<sup>2</sup>. Серебро (атомный номер 47) имеет один внешний электрон в s-состоянии. Ожидалось, что его атомы вообще не будут отклоняться: s-состояние, как симметричное, имеет нулевой момент, а моменты внутренних электронов компенсированы.

Однако пучок расщепился на две проекции. Они как раз соответствуют двум возможным направлениям спина внешнего электрона: вдоль и против поля. Для объяснения этого эксперимента и была предложена Гаудсмитом и Уленбеком концепция спина (1925); теоретически она обоснована Паули в 1927 г.

Частицы с полуцелым спином называют еще **фермионами**, так как они подчиняются статистике Ферми – Дирака, в которой нельзя занимать одно состояние более чем одной частице. Это, собственно, и есть принцип Паули (который применим именно для объектов с полуцелым спином). Практически весь наблюдаемый мир состоит из фермионов (нейтрон, протон, электрон). Именно принцип Паули обеспечивает жесткость и непроницаемость предметов. Нельзя проткнуть стол пальцем, потому что для электронов пальца нет места в атомах стола. Все нижние состояния заняты, и приходится размещать электроны выше, на что тратится много энергии.

Для частиц с целым спином (**бозонов**) годится статистика Бозе – Эйнштейна, в которой сколько угодно частиц могут быть в одном состоянии и даже предпочитают такую тесноту (наиболее знакомый из них – фотон, спин 1). Бозоны переносят взаимодействия между фермионами (например, фотон – переносчик электромагнитного взаимодействия). Если четное число фермионов объединяются в тесно связанную систему, то суммарный спин получается целым, и система ведет себя как бозон (пример –  $\alpha$ -частица, спин 0). Конечно, это приближение, оно годится, пока система хорошо изолирована, например,  $\alpha$ -частица не обменивается нуклонами с другими ядрами.

<sup>2</sup>При непрерывном «классическом» распределении наблюдалось бы непрерывное размытие пучка.

## 6.2 Состав ядра. Ядерные силы

Поскольку обнаружилось, что масса атома практически сосредоточена в ядре, и было известно, что массы атомов примерно целые числа, если массу водорода принять за 1, то вначале возникло предположение, что ядра составлены из одинаковых частиц, а именно протонов – ядер водорода. Термин «протон», от греч. *протоос* – первый, ввел Э. Резерфорд в начале 20-х годов. Таким образом, в отличие от электрона и большинства других частиц, протон не был «открыт» внезапно, такого-то числа. Правильнее сказать, что понятие протона зародилось одновременно с планетарной моделью и сформировалось постепенно, по ходу накопления данных.

Однако заряд всех ядер, кроме водорода, при чисто протонном составе был бы слишком велик (например, 16 у кислорода, вместо 8). Думали, что протон можно нейтрализовать электроном, с образованием некоей устойчивой системы. Если часть протонов в ядре «проглотит» по электрону, то с зарядом все будет в порядке. Но ведь заведомо известное связанное состояние протона и электрона – атом водорода. Как электрон может скрываться в ядре, которое на 5 порядков меньше размерами, это было не ясно. Кроме того, спины и электрона, и протона равны  $1/2$ . Составная частица имела бы целый спин (например, нулевой), что противоречило данным о спинах ядер. Например, азот имел бы 14 протонов и 7 электронов в ядре, т.е. нечетное число частиц спина  $1/2$ . Однако спин ядра азота оказался равным 1.

**Нейтрон.** В 1932 г. Д. Чедвик открыл нейтрон. При бомбардировке бериллия  $\alpha$ -частицами возникало необычное излучение (В.Боте, Г.Беккер, 1930). Бериллий – активный и очень ядовитый металл. Возможно, из-за этого до него дошла очередь так поздно. В современных обозначениях происходит реакция



Необычность состояла в проникающей способности: излучение легко проходило через несколько см металла. Для сравнения:  $\alpha$ -излучение задерживается папиросной бумагой,  $\beta$  – миллиметровым слоем металла. И даже  $\gamma$  – лучи от известных источников ослаблялись гораздо сильнее нового излучения.

Отсюда следовало, что новые частицы не имеют электрического заряда. Они не оставляли никаких треков в камере Вильсона или в фотоэмульсии. Заметить же их удалось, потому что излучение выбивало ядра из окружающей среды, например, протоны из парафинового образца (И.Кюри, Ф.Жолио, 1932). Протон заряжен, и за ним уже можно наблюдать. Сравнивая энергии вылетающих протонов и ядер азота, выбиваемых из молекул  $\text{N}_2$  в объеме камеры Вильсона, Чедвик показал, что новое «излучение» – это не  $\gamma$ -лучи, как думали Кюри и Жолио, а нейтральные частицы с массой, примерно такой же, как у протона. Для них уже было готово имя – нейтроны. К этому моменту Чедвик вместе с Резерфордом искали нейтрон в свободном состоянии более 10 лет.

**Состав ядра. Изотопы.** Нейтрон идеально подошел на роль недостающей компо-

ненты ядра. В том же 1932 г. Д.Д. Иваненко и В. Гейзенберг предложили нейтронно-протонную модель ядра, которая принята и сейчас:  $A = Z + N$ , где  $A$  – массовое число,  $Z$  – число протонов (то есть заряд ядра),  $N$  – число нейтронов. Если пренебречь малой разницей масс протона и нейтрона, то же самое уравнение довольно точно годится для суммарной массы ядра. Существо модели, конечно, не в этом уравнении, а в представлении, что нейтрон – не составной агрегат, а элементарная частица, как и протон, со спином  $1/2$ . Конечно, подтверждение гипотезы потребовало массы экспериментов.

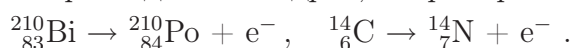
В таблице Менделеева легко найти дробные массы элементов (например, хлор: 35,453). На самом деле почти все элементы – это природные смеси **изотопов**. Изотопы отличаются друг от друга числом нейтронов  $N$  при данном заряде  $Z$ . Природный хлор содержит  $3/4$  изотопа  $^{35}\text{Cl}$  и  $1/4$   $^{37}\text{Cl}$ . Заряд определяет число электронов в атоме и, следовательно, химические свойства элемента. Изменения же  $N$  только немножко влияют на размер ядра. А раз ядро такое маленькое, то заметного влияния его размера на электроны (и, значит, на химические свойства) нет.

**Свойства нейтрона.** Точные измерения массы нейтрона дают 1,008665 атомных единиц (у протона 1,007276). В единицах массы электрона нейтрон «весит» 1838,6, а протон 1836,1. Разность масс около 2,5 электрона позволяет нейтрону распадаться на электрон и протон. Вначале уравнение реакции представляли в виде:



Время жизни свободного нейтрона около 900 секунд. Поэтому в таблице Менделеева нет «нулевого» элемента – нейтрона.

Нейтроны могут также распадаться в ядрах, например



Это и есть  $\beta$  – распад, один из классических видов радиоактивности. Все же большинство естественных ядер (около 260) стабильны: найдено всего около 40 природных радиоактивных изотопов. Почему нейтрон, находящийся в стабильном ядре, не распадается? Физическое объяснение в том, что распад нейтрона в таком ядре энергетически невыгоден из-за действия ядерных сил. Замена нейтрона на протон модифицирует взаимодействия так, что энергия (и масса) ядра возрастают.

Спин нейтрона, как и протона,  $1/2$ . Следовательно, нейтрон не позволяет близким соседям – нейтронам находиться в том же состоянии.

Написанные выше реакции не совсем правильны. Оказалось, что энергия вылетающего электрона не фиксирована – она «гуляет» от нуля до максимального значения  $1,5m_e c^2$ . Кроме того, нарушено сохранение момента – из одного спина  $1/2$  получилось два таких. Для случая  $\beta$  – распада ядер Паули в 1930 г. постулировал существование новой частицы, которая так слабо взаимодействует со всеми остальными, что всегда улетает незамеченной (кстати, под названием «нейтрон»). Этой частице и передается остаток энергии распада, не доставшийся электрону. Спин ее тоже  $1/2$ , так что можно сохранить полный момент. После открытия «настоящего» нейтрона, в 1934 г., Ферми

предложил называть изобретенную Паули частицу «нейтрино» (итал. уменьшительное от «нейтрон»). Правильное уравнение распада нейтрона:



Только в 1953 г. нейтрино были зарегистрированы в эксперименте<sup>3</sup>.

На первый взгляд, распад нейтрона указывает на его составную природу (протон + электрон + нейтрино). Но есть примеры распада протона в нейтрон. Сам по себе протон стабилен, но в некоторых ядрах, где протоны в избытке, один из них распадается:



Здесь  $e^+$  – это позитрон, античастица электрона. (Строго говоря, в этой реакции вылетает то, что называют нейтрино  $\nu$ , при распаде же нейтрона вылетает антинейтрино  $\tilde{\nu}$ ). Более того, можно сделать протоны и нейтроны при столкновении быстрых электронов и даже  $\gamma$  – квантов, внутри которых уж точно нет никаких частиц вещества. Значит, при распаде частицы не «освобождаются» (как осколки снарядов, камней и даже атомов), а рождаются, если для этого хватает энергии.

Как уже говорилось, нейтрон не имеет электрического заряда и поэтому не оставляет за собой ионизованного «хвоста». Только попав в ядро, он проявляет себя – может отскочить упруго, к некоторым ядрам прилипает, переводя их в более тяжелый изотоп. Нейтроны могут и разбивать ядра. Это разнообразие свойств делает нейтроны лучшим инструментом воздействия на ядро и определяет их использование на практике.

Оценим длину свободного пробега  $L$ , предполагая, что сечение столкновения  $\sigma$  нейтрона с ядрами геометрическое – порядка барна ( $10^{-24}$  см<sup>2</sup>). Тогда  $L \sim 1/n\sigma \approx 10$  см при характерной для конденсированного вещества плотности ядер  $10^{23}$  см<sup>-3</sup>. Это объясняет значительную проникающую способность нейтронов.

**Ядерные силы.** Мы видели, что атом в основном – электрическая система. Основную роль играют кулоновские силы. Разумеется, для ядра это не так, ведь заряд ядра не компенсирован. Что-то должно сдерживать ядро от развала.

Близость масс нейтрона и протона не случайна. Эти частицы родственны, и их вместе называют нуклонами, составными элементами ядра (от лат. nucleus – ядро).

Опыт показывает, что между нуклонами существуют значительные силы притяжения. Они действуют только на «ядерных» расстояниях порядка  $r \approx 10^{-13}$  см (что следует хотя бы из способности нейтронов проникать в вещество). Но в этой маленькой области ядерные силы значительно превосходят кулоновские. Различие сил между парами nn, np и pp малó (что и дает основание для использования понятия нуклона).

Например, рассмотрим дейтрон – ядро тяжелого водорода, состава np. Энергия связи протона и нейтрона по измерениям равна 2,2 МэВ. Для сравнения оценим электрическую энергию взаимодействия протонов, удаленных на  $r$ :  $e^2/r = 23 \cdot 10^{-20}/10^{-13} = 23 \cdot 10^{-7}$  эрг = 1,5 МэВ. На первый взгляд, ядерные силы лишь немного больше кулоновских. Но мы не учли еще кинетическую энергию нуклонов. Раз они витают в объеме

<sup>3</sup>Паули предложил приз за обнаружение нейтрино (корзину шампанского) и сдержал слово.



размера  $r$ , им положено иметь кинетическую энергию  $\sim \hbar^2/2mr^2 = 3 \cdot 10^{-5}$  эрг  $\approx 20$  МэВ! Следовательно, ядерные силы запасают не только 2,2 МэВ энергии связи, но и еще около 40 МэВ, сдерживающих нуклоны, которые стремятся разбежаться за счет кинетической энергии «нулевых колебаний». Получаем, что работа ядерных сил примерно в 30 раз больше работы кулоновских. Соотношение сил еще более резкое, так как ядерные силы короткодействующие. В общем принято считать, что ядерные силы в «своем» диапазоне расстояний примерно на два порядка больше электрических.

Два протона на расстоянии  $r$  отталкиваются с электрической силой  $e^2/r^2 = 23 \cdot 10^{-7}/10^{-13} = 2,3 \cdot 10^7$  дин, или около 23 кГ! Наконец мы добрались до ощутимых порядков величины в микромире. Соответственно характерная ядерная сила между двумя нуклонами порядка 1 – 2 тонн. Из-за короткодействия ядерные силы подобны молекулярным. Ядро ведет себя как капля «нуклонной» жидкости, объем капли пропорционален числу нуклонов  $A$ .

Ядерные силы несколько зависят от направления спинов. Например, в дейтроне спины параллельны, так что спин дейтрона в целом 1. Вспомним, что большая часть отрицательной энергии притяжения компенсируется кинетической энергией, и на энергию связи остается всего  $\sim 10\%$  (2,2 МэВ). При антипараллельных спинах взаимодействие чуть слабее, и устойчивого ядра не получается. По этой же причине не бывает динейтрона ( $nn$ ) и ядра  ${}^2\text{He}$  ( $pp$ ), так как в них по принципу Паули спины должны быть противоположны.

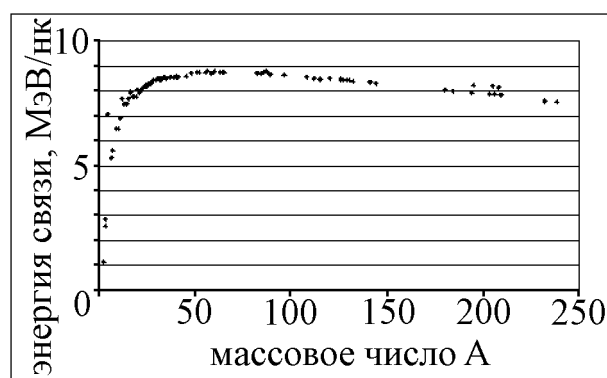


Рис. 6.1.

**Энергия связи.** Масса атома дейтерия 2,014102, водорода 1,007825, нейтрона 1,008665 атомных единиц. Дейтрон чуть легче, чем пара  $n, p$ : дефект массы  $\Delta m = 0,002388$ . Это около 0,12 процента массы, что вполне измеримо на хороших весах.

Дефекту массы соответствует энергия связи  $\Delta mc^2$ . Одна атомная единица – это примерно 940 МэВ, то есть  $\Delta mc^2 = 0,002388 \cdot 940 = 2,2$  МэВ (эта цифра уже упоминалась). Таким же образом, по существу простым взвешиванием, можно найти энергию связи для любого изотопа. График удельной энергии связи  $E$  (рассчитанной на один нуклон) от массового числа  $A$  показан на рис. 6.1. Так как одинаковые  $A$  могут быть у разных элементов, обычно рисуются усредненные данные по стабильным ядрам.

Для протона  $E = 0$ . При малых  $A$  энергия связи быстро растет, так как увеличивается число соседей нуклона (или, другими словами, падает роль поверхности). У элементов, близких к железу ( $A = 56$ )  $E$  имеет максимум 8,7 – 8,8 МэВ/нк. Это наиболее устойчивые ядра. В принципе все вещество должно преобразоваться в железо, но это процесс медленный (или можно сказать, что Вселенная еще молода). Далее с ростом



А устойчивость снова уменьшается (в частности, из-за кулоновского отталкивания).

### 6.3 Реакции деления. Цепная реакция

Среди ядерных реакций вообще реакции деления важны как практически доступный источник энергии. Прежде чем рассматривать деление, остановимся на классических видах радиоактивности.

**Неустойчивость ядер.** Для устойчивости ядра необходим некоторый баланс между числом нейтронов  $N$  и протонов  $Z$ . Карта возможных ядер с выделением более или менее стабильных изотопов показывает, что границы этого баланса довольно узкие. Можно сказать, что существует вытянутая «долина устойчивости» с крутыми берегами (см. рис. 6.2<sup>4</sup>). Ядра с избытком нейтронов стремятся превратить нейтрон в протон путем  $\beta$ -распада. При избытке протонов происходит  $\beta^+$ -распад, то есть протон превращается в нейтрон с испусканием позитрона.

На энергетическом языке это значит, что ядро до распада более тяжелое, чем продукт, так что энергии распада хватает на создание  $\beta$ -частиц, нейтрино и придание им кинетической энергии. Причины существования минимума энергии в направлении «поперек» долины устойчивости можно понять из принципа Паули. Если рассуждать о не очень тяжелых ядрах, примерно до середины таблицы, то у них устойчивая конфигурация  $Z \approx N$ . Как нейтроны, так и протоны занимают свои нижние энергетические уровни. Если добавлять «лишние» нейтроны, им придется занимать уровни повыше, то есть иметь большую кинетическую энергию. То же самое верно для протонов. Можно сказать, что бока долины – это как бы «склоны Паули».

Долина устойчивости поднимается и в продольном направлении. Причина этого – кулоновское отталкивание протонов в ядре. Поскольку электрическая энергия пропорциональна  $Z^2$ , рано или поздно она начинает сказываться даже при сравнительно слабости электрических сил. По этой же причине полоса устойчивости загибается вверх (рис. 6.2): при больших  $Z$  число нейтронов  $N \approx 1,5Z$  (например, для  $^{238}\text{U}$ :  $N/Z = 146/92 = 1,59$ ). Выгодно «разбавить» ядро нейтронами, чтобы ослабить кулоновское отталкивание протонов. Но при этом возрастает кинетическая энергия нейтронов по принципу Паули. Выживают ядра, для которых достигается некоторый компро-

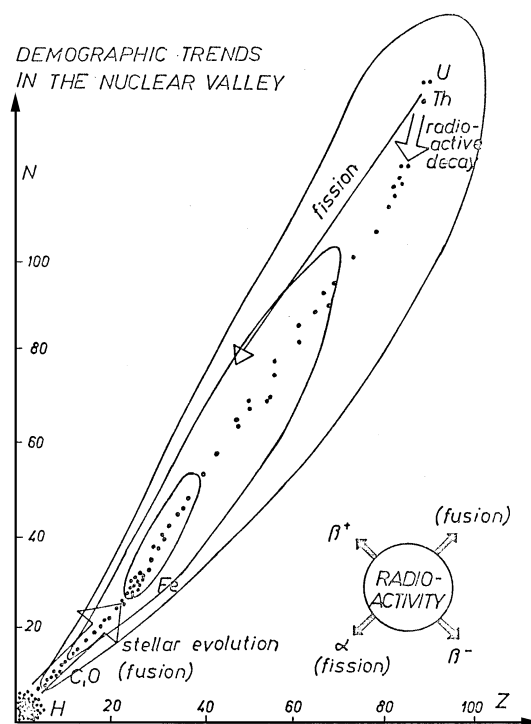


Рис. 6.2.

<sup>4</sup>Взято из доклада венгерского физика Д. Маркса на международной конференции по преподаванию физики: Marx J. Our nuclear history. In: Nuclear energy. Nuclear power. Proc. of seminar on teaching physics in schools (GIREP '81). Balaton, Hungary, 1981. P. 213–233.

мисс между этими эффектами. Конечно, компромисс не идеальный, и энергия связи тяжелых ядер падает.

Поэтому ядро с большим  $Z$  стремится уменьшить свой заряд. Один из способов – это  $\alpha$ -распад. Когда из ядра вылетает  $\alpha$ -частица, заряд уменьшается на 2, а масса на 4. Правда, при этом нейтронно-протонный баланс становится чуть хуже, но дело может поправить последующий  $\beta$ -распад (или несколько).

Альфа-частица, в отличие от электрона, в некотором смысле существует внутри ядра и до распада, как устойчивая комбинация четырех нуклонов. Ясно, что движется она крайне быстро и «стучается» о границы ядра часто. Распад происходит далеко не при каждом таком ударе, так как вылету препятствует потенциальный барьер. Но все же изредка  $\alpha$ -частица производит **туннельный переход** под барьером. Поскольку в квантовой механике все является волной, то существует, хотя и небольшая, волновая функция  $\alpha$ -частицы в классически запрещенной области и тем самым вероятность покинуть ядро. Это подтверждается расчетами периодов полураспада. И наоборот, налетев снаружи на ядро, частица в принципе могла бы изредка проникать внутрь, также пройдя под барьером. На практике же наблюдается только  $\alpha$ -распад, но мы не слышим про обратный ему  $\alpha$ -синтез, потому что концентрации и потоки  $\alpha$ -частиц снаружи ядра несравнимы с «внутренними». На ускорителях такие реакции возможны. С наружной стороны ядра барьер сформирован электрическим отталкиванием, и поэтому его называют кулоновским.

Для примера рассмотрим радиоактивный ряд урана-238 (рис. 6.3). Всего выходит 8  $\alpha$ - и 6  $\beta$ -распадов, с суммарной энергией около 50 МэВ. Поскольку первый распад – самый редкий, все последующие «подстраиваются» под него, и в породах, содержащих уран, всегда есть его «потомки» в сложившемся за миллиарды лет равновесии. Измеряя количества дочерних изотопов, геофизики определяют возраст пород. Подобные ряды дают уран-235, торий-232 и (не существующий в естественном виде) нептуний-237.

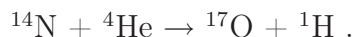
В отличие от  $\alpha$  и  $\beta$  гамма-излучение – это не распад, а переход возбужденного ядра на более низкий энергетический уровень. Никаких перемещений по таблице Менделеева при этом не происходит. Высвободившаяся энергия уносится фотоном ( $\gamma$ -квантом). Позднее были найдены и другие типы радиоактивности – вылет из ядра протонов, нейтронов, ядер углерода (и даже вплоть до магния), спонтанное деление (распад тяжелого ядра на примерно одинаковые осколки). В естественных условиях их роль невелика.

При радиоактивности выделяется огромная энергия. Если 1 МэВ получается при распаде одного ядра, то 1 моль ядер даст  $6 \cdot 1,6 \cdot 10^{23-12+6} \approx 10^{18}$  эрг =  $10^{11}$  Дж. При затратах на отопление дома 10 кВт этой энергии хватит на  $10^7$  секунд, то есть практически на год (а ведь это максимум граммов двести топлива)! Но в своей первобытной форме ядерная энергия дорога («вокруг» грамма радия в начале века строили целые институты) и, еще хуже, неуправляема. Вещество распадается, не обращая на нас никакого внимания.



Рис. 6.3.

В 1919 г. Резерфорд показал, что повлиять на ядро возможно, хотя это и трудно. Энергичные  $\alpha$ -частицы, налетая на ядра азота, «сделали» несколько ядер кислорода:



Была открыта трансмутация, мечта алхимиков<sup>5</sup>. Можно делать и золото, например из  $\alpha$ -частиц и иридия, но это будет гораздо дороже, чем добывать естественное золото.

Для таких реакций нужна энергия. Тяжелые ядра защищены кулоновским барьером, и в них  $\alpha$ -частицам проникнуть трудно. В некоторых лабораториях взялись за постройку циклотронов, чтобы разгонять протоны или другие положительные частицы и преодолевать кулоновское отталкивание. Но затрачивалось энергии всегда больше, чем удавалось получить. Резерфорд до конца жизни не верил в перспективы ядерной энергетики.

**Деление нейтронами.** Сразу после открытия нейтрона его стали использовать для воздействия на ядра. Массу опытов с почти всеми элементами провел Э. Ферми в Риме. Нейтрон не отталкивается от ядра (нет кулоновского барьера) и поэтому может не только стукнуться о ядро, но и проникнуть внутрь, поглотиться. И для этого не надо дорогих ускорителей. Недостаток средств «помог» группе Ферми.

При захвате нейтрона получается более тяжелый изотоп, часто неустойчивый. Распадаясь, он дает либо следующий элемент  $Z + 1$  ( $\beta$ ), либо  $Z - 2$  ( $\alpha$ ). Возникает, как

<sup>5</sup>Собственно,  $\alpha$ - и  $\beta$ -распады – это тоже трансмутация, ведь происходит превращение элементов. Но распад идет сам по себе, без всякой возможности влияния на него. В реакции же Резерфорда кислород получен именно искусственным путем. Интересно, что в естественных радиоактивных рядах по иронии судьбы получается не золото, чего добывались алхимики, а как раз свинец, служивший им сырьем.

говорят, наведенная радиоактивность. Это позволяет делать датчики нейтронного облучения: и сейчас применяются датчики на основе индия, серебра и т.п., активность которых была обнаружена Ферми.

В числе прочего нейтронами просветили и уран. Была получена заметная активность, но идентифицировать продукты реакции не удалось. Надеялись, что получается следующий за ураном элемент 93. Это действительно имеет место, но основная активность была связана с делением урана нейтронами, до чего Ферми и его сотрудники не догадались. Явление было понято спустя несколько лет.

Только в 1938 г. О. Ган и Ф. Штрассман в Германии химическим путем установили, что среди продуктов при облучении урана нейтронами возникает изотоп бария. Барий гораздо легче урана, почти вдвое (номер 56, атомный вес по таблице 137). Трудно представить цепь распадов, ведущую к его образованию (ср.  $^{206}_{80}\text{Pb}$  при обычном распаде урана). Возникла идея, что ядро урана разрывается на две близкие по массе части.

Это – даже не новый вид радиоактивности, а нечто более перспективное. Замечательно выделение огромной энергии при делении ядра. Отталкивание зарядов 56 и 36 (это ядро криптона) дает много больше, чем обычные виды распада. Лиза Мейтнер, начинавшая с Ганом и Штрассманом эту работу, первая оценила энергию реакции; она же предложила термин «деление урана».

Ранее Ирен Кюри<sup>6</sup> выделила продукт, который по всем признакам совпадал с лантаном – соседом бария. В это никто не верил до работ Гана, Штрассмана и Мейтнер. Сейчас известно, что лантан действительно один из продуктов деления, и вообще получается целое распределение ядер. Отсюда видно, что объяснить открытие по крайней мере так же важно, как сделать его.

**Калорийность урана.** Пусть ядро разрывается на две примерно одинаковые части. Раз нейтрон делит ядро, значит, оно и до этого было не слишком прочным. Ядерные силы лишь немного превосходили электрическое отталкивание. И действительно, делился самый последний (на то время) элемент – уран; он и сам может распадаться, даже без такого слабого воздействия, как залетевший нейтрон. В 1940 г. Флеров и Петржак открыли спонтанное деление урана – очень редкий, но возможный процесс. Следовательно, можно записать  $A^{2/3}f \sim (Ze/2)^2/(A^{1/3}r)^2$ , где  $f$  – характерная ядерная сила,  $r$  – ядерная длина  $10^{-13}$  см. Слева – удерживающая сила ( $A^{2/3}$  – число связей между нуклонами в сечении ядра), справа – кулоновское отталкивание. Отсюда  $f \sim (e^2/r^2) \cdot (Z^2/4A^{4/3})$ . Мы уже оценивали  $e^2/r^2 \approx 20$  кГ. Примерно та же величина получается и для  $f$  при  $Z = 92$ ,  $A = 238$ . Но мы знаем, что ядерные силы значительно больше. Как и в случае с дейтроном, мы не учли движения нуклонов в ядре, которое создает добавочное давление и понижает устойчивость. Можно сказать, что  $f$  – это как бы эффективная сила, учитывающая движение нуклонов.

<sup>6</sup>Дочь Марии Кюри, часто упоминается под фамилией Жолио-Кюри, по мужу – Фредерику Жолио.

Если ядро «лопнет» от удара нейтрона, то, как только половинки разойдутся на  $\approx r$ , ядерные силы практически перестанут сдерживать их, а кулоновские заметно не изменятся. Далее осколки деления разлетаются за счет работы кулоновских сил. (Это напоминает выдергивание гвоздя из стены: когда гвоздь уже вылез, человек еще упирается, его сила более «дальнодействующая», от чего можно упасть и что-нибудь себе повредить). Выделившаяся энергия порядка  $Z^2 e^2 / 4A^{1/3} r$ , что можно записать еще так:

$$E \approx \left( \frac{e^2}{2a} \right) \cdot \left( \frac{2a}{A^{1/3} r} \right) \cdot \left( \frac{Z^2}{4} \right).$$

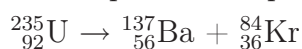
Первый множитель – энергия электрона в атоме (13,6 эВ), второй – отношение размера атома ( $10^{-8}$  см) к размеру ядра ( $10^{-12}$  см), получается  $10^4$ , третий для урана около  $2 \cdot 10^3$ . Итого выделяется около 200 МэВ, или 1 МэВ на нуклон исходного ядра. Ту же оценку можно получить из графика энергии связи, сравнивая тяжелые ядра с ядрами из середины таблицы Менделеева.

Получается, что эта «ядерная» энергия на самом деле электрическая. Работа ядерных сил гораздо меньше и вдобавок имеет другой знак (уменьшает энерговыделение).

При химических реакциях атомы обмениваются электронами, выигрывается какая-то доля от характерной энергии электрона  $e^2/2a$ , то есть порядка электронвольта. Чтобы оценить удельный тепловой эффект – на нуклон, надо поделить на молекулярный вес продуктов реакции. Например, при горении угля или нефти получаются  $\text{CO}_2$  ( $\mu = 44$ ) и  $\text{H}_2\text{O}$  ( $\mu = 18$ ). Достигается температура порядка тысяч градусов, что соответствует десятым долям эВ. Величина энерговыделения при делении велика (в миллионы раз больше по сравнению с «химией») из-за двух факторов: малых размеров ядра (см. вторую дробь в формуле  $E$ ) и его большого заряда (см. третью дробь).

Это уже указывает на перспективы. В отличие от всех предыдущих видов распада (когда тоже выделяются МэВы), деление ядра контролируемо. Оно вызывается нейтронами, и есть надежда им управлять.

**Цепная реакция.** При делении образуются ядра из середины таблицы, для которых характерно гораздо меньшее число нейтронов. Например, для распада



избыток нейтронов в левой части равен  $235 - 137 - 84 = 14$ . Если образуется самый тяжелый природный изотоп криптона – 86, и то имеется 12 лишних нейтронов (барий и так самый тяжелый).

Конечно, при делении скорее должны получаться менее устойчивые изотопы. Например, бывают  ${}^{143}\text{Ba}$  (живет 12 сек) и  ${}^{92}\text{Kr}$  (3 сек), которые в сумме дают баланс. Лишние нейтроны в таких ядрах переходят в протоны цепочкой  $\beta$  – распадов. Но вряд ли только такие распады возможны. С некой частотой возникают и более нормальные изотопы, а часть лишних нейтронов при этом просто вылетает в свободном виде. Этому способствует крайне возбужденное состояние получившихся осколков – от довольно замысловатой формы в момент деления они должны перейти к сферической.

Мысль о свободных нейтронах деления пришла в голову многим, и среди первых был Ферми, а в нашей стране Зельдович и Харитон. Мейтнер после эмиграции из Германии стала работать у Бора. 16 января 1939 г. Бор приехал в США погостить у Эйнштейна в Принстоне. Он сообщил о делении ядра. Эта новость мигом распространилась и дошла до Ферми (тоже к тому времени эмигрант) в Колумбийский университет. Ферми, Даннинг и Пеграм решили начать поиск событий деления (по необычно толстым ионизационным следам от осколков). 26 января началась конференция по теоретической физике в Вашингтоне, где Ферми обсуждал проблему нейтронов с Бором. Еще до окончания конференции в четырех лабораториях США явление деления было экспериментально подтверждено.

Пусть один исходный нейтрон разбивает ядро урана. Если в среднем вылетает, например, 2 свободных нейтрона, они разобьют уже 2 ядра. Четыре полученных нейтрона разбивают 4 ядра, и так далее. Чтобы разбить все ядра, скажем, в килограмме урана ( $< 5$  молей), надо  $N$  таких шагов, причем  $N$  не так велико:  $2^N \sim 30 \cdot 10^{23}$ , или  $N = 24,5 / \lg 2 = 80$ . Нейтроны летают быстро, и вся эта энергия выделится за микросекунды, как будто взорвалось вещество, в миллионы раз мощнее тротила.

Это и есть **цепная реакция**. К счастью, подобный ход событий не так просто осуществить (а буквально такой – и невозможно). Для получения ядерного взрыва (1945, 1949) потребовались усилия самых мощных стран в течение нескольких лет, и эффективность в результате далеко уступала описанному кошмару.

# ВОЛНОВАЯ ФИЗИКА

Курс лекций для ФМШ

Часть 2. СТРОЕНИЕ ВЕЩЕСТВА

7. ЯДЕРНАЯ ТЕХНИКА. ЭЛЕМЕНТАРНЫЕ ЧАСТИЦЫ

А. П. Ершов

12 февраля 2007 г.



## Глава 7

# ЯДЕРНАЯ ТЕХНИКА

Считается, что энергия деления вначале использовалась в бомбах и только позднее были созданы атомные энергетические станции. Это очень неточно. Первые реакторы предшествовали бомбам, только в них не вырабатывалось электричество. А самый первый реактор на Земле возник вообще без участия человека. История освоения деления ядер настолько увлекательна, что было бы глубоко неправильно излагать суть проблемы в каком-то другом порядке.

### 7.1 Атомное оружие

**Замедление нейтронов.** В опытах группы Ферми был открыт поразительный эффект. Оказалось, что активность, наведенная источником нейтронов в образце, зависит еще и от окружающих предметов. Влияли такие вещи, как расположение деревянного стола, материал подставки, и т.д. А когда источник поместили в парафиновый блок, активность серебряного датчика повысилась в сотню раз! В этот момент Ферми произнес известное изречение «Пошли обедать». К концу обеда он уже мог объяснить явление. Парафин (состав  $\text{C}_n\text{H}_2$ ) эффективно замедлял нейтроны, так как массы протона и нейтрона близки. Бильярдный шар, ударяя в лоб другой шар, останавливается, а от бортика стола отскакивает почти без потери энергии. Не каждый удар лобовой, и в среднем теряется половина энергии. Идею тут же проверили, заменив парафин водой. В фонтане (с золотыми рыбками) тоже получился резкий рост активности.

Почему же замедленные нейтроны действуют лучше? Оказалось, что сечение взаимодействия медленного нейтрона с ядром, которое его способно захватить, растет в сотни раз (пропорционально  $1/v$ , где  $v$  – скорость нейтрона). Это квантовый эффект, и грубое его объяснение в том, что взаимодействие зависит не только от размера ядра, но и от длины волны нейтрона. При энергии  $E = 1$  МэВ импульс нейтрона  $p = \sqrt{2mE} \approx 2 \cdot 10^{-15}$  СГС, а длина волны  $\hbar/p \approx 5 \cdot 10^{-13}$  см, то есть меньше размеров ядра. Такой нейтрон «видит» ядро, как и положено, шариком размера  $\sim 10^{-12}$  см. Но если нейтрон охла-

дится даже до 1000 градусов (это 0,1 эВ), длина волны возрастет на три с половиной порядка, до  $\sim 10^{-9}$  см. Это уже в тысячу раз больше, чем размер ядра. Примерно в той же степени возрастает сечение. Представьте себе, что вы проходите мимо столба в 10 метрах от него. Каково же будет ваше удивление, когда вы стукнетесь лбом! Это просто очередное «квантовое чудо». Нейтрон размазан в пространстве и потому может поглотиться «издали».

Раз медленный нейтрон так активно ловится ядром, его пробег в поглощающем веществе (например, серебряном индикаторе) будет уже не единицы см, а доли мм. Тонкая пластинка серебра, через которую быстрые нейтроны в основном пролетали насквозь, поглощает практически все медленные нейтроны. Вот активность и подсакивает. За результаты по активации нейтронами и в особенности за открытие эффекта замедления (и его теоретическое объяснение) Ферми получил Нобелевскую премию (1938).

**Изотопы урана.** Природный уран – в основном изотоп  $^{238}\text{U}$ , с небольшой примесью  $^{235}\text{U}$  (0,72%). Период полураспада  $^{235}\text{U}$  равен 0,7 миллиарда лет (меньше, чем 4,5 млрд для  $^{238}\text{U}$ ). Этот легкий изотоп – родоначальник своего радиоактивного ряда, похожего на ряд  $^{238}$ -го урана. Меньший период полураспада говорит о меньшей стабильности. Уже в 30-е годы было понятно, что тяжелые нечетные изотопы менее устойчивы (так как не все нейтроны либо протоны могут объединяться в пары). Вопрос, какой из изотопов урана делится нейтронами, практически очень важен.

К июню 1940 г. было установлено, что оба изотопа делятся быстрыми нейтронами. Медленные же нейтроны могут делить только  $^{235}\text{U}$  изотоп. Но почему такое может происходить вообще? Как тепловой нейтрон, с энергией, соответствующей комнатной температуре (300 К, или 0,03 эВ) вызывает превращение с выделением 200 МэВ? Ведь надо нарушить баланс сил в ядре, для которого 0,03 эВ совершенно незаметны.

Ответ в том, что кинетическая энергия «на бесконечности» – не главное. Нейтрон притягивается к ядру. И когда он попадает внутрь, то набирает несколько МэВ кинетической энергии – ядерный масштаб, чего и достаточно для менее устойчивого  $^{235}\text{U}$ . По-видимому, именно эта неожиданность эффективности медленных нейтронов мешала исследователям заметить деление  $^{235}\text{U}$  гораздо раньше.

**Организация исследований.** В июле 1939 г., по инициативе Л. Сцилларда и Е. Вигнера, Эйнштейн обратился к президенту США Ф.Д. Рузвельту с письмом, где подчеркивалась важность урановой проблемы и опасность развития этих работ в Германии. Осенью 1939 г. был назначен Урановый консультативный комитет, который подал президенту доклад (1 ноября 1939), где прямо указывалось на реальность создания бомб. 20 февраля 1940 была выдана первая субсидия в \$6000 для закупки материалов.

В июне 1940 г. был организован Исследовательский комитет национальной обороны, а в июне 1942 г. – новый округ инженерных войск (известный как Манхэттенский округ). В сентябре 1942 г. руководителем проекта был назначен генерал Л. Гроувз.

**Реактор Ферми.** 2 декабря 1942 года впервые была получена самоподдерживающаяся

цепная реакция. В Металлургической лаборатории, организованной в Чикаго, Ферми и его сотрудники построили первый ядерный реактор. Это был шар диаметром в несколько метров, сложенный из графитовых кирпичей, разделявших блоки металлического урана и окиси  $\text{UO}_2$ .

Графит (практически чистый углерод) служил замедлителем нейтронов. Нейтроны деления, вылетев из блока урана, попадали в графит и после десятков столкновений снижали свою энергию до «комнатной», соответствующей температуре графита. Рано или поздно нейтрон, диффундируя через графит, мог вновь попасть в урановый блок. Но теперь он, уже медленный, практически сразу ловился ядром  $^{235}\text{U}$ , не замечая ядер  $^{238}\text{U}$  из-за разницы в сечениях на два-три порядка.

Разбитое ядро давало два-три новых нейтрона. Не все они снова разбивали ядро урана, так как могли частично поглотиться в примесях графита, прилипнуть к  $^{238}\text{U}$  (без деления) и, наконец, вылететь наружу. Потери из-за вылета нейтронов минимальны для шара, а при данной форме они тем меньше, чем больше размер реактора. Если освободившиеся нейтроны производят в среднем один акт деления, то «котел» работает в стационарном режиме. Для «разгона» реактора коэффициент размножения нейтронов должен превышать единицу (реально было 1,0006). При первом включении была достигнута мощность 0,5 Вт, а через несколько дней – около 200 Вт. Далее повышать мощность не стали из-за опасности облучения (никакой защиты не было).

Реактор регулировался поглощающими стержнями, содержащими кадмий. Этот элемент очень активно ловит нейтроны (с сечением в сотни барн), и когда стержни введены полностью, размножения не происходит. Постепенное извлечение стержней уменьшает поглощение, и при некотором их положении реактор начинает разгоняться. Зафиксировать заданную мощность или вообще заглушить реакцию можно, снова вдвигая стержни. Природные графит и уран приходилось чистить от кадмия, бора и других естественных поглотителей. К концу 1942 г. были измерены сечения и получены достаточно чистые материалы, что и позволило спроектировать и построить котел.

**Критическая масса.** Цепная реакция, развиваясь по экспоненте, рано или поздно приведет ко взрыву. Ясно, что непрактично сбрасывать на противника ядерный реактор весом в сотни тонн. Делящийся материал следует концентрировать. При заданном составе ядерного взрывчатого вещества и форме заряда существуют критическая масса и критический размер, при превышении которых «образец» взрывается. Например, для  $^{235}\text{U}$  в виде шара критический радиус около 9 см, а масса – около 50 кг.

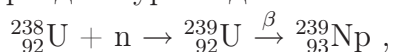
В природном уране 0,72% легкого изотопа  $^{235}\text{U}$ . Оказывается, что основной изотоп  $^{238}\text{U}$  слишком устойчив, чтобы взрываться. Медленные нейтроны вообще не делят ядро  $^{238}\text{U}$ , а попросту захватываются. Быстрые нейтроны делят, но с недостаточной вероятностью. Замедление же нейтронов в природной смеси, с тем чтобы реакция развивалась хотя бы на  $^{235}\text{U}$ , тоже неэффективно, так как нейтроны сильно захватываются тяжелым изотопом в промежуточной области энергий (десятки эВ). (Для этого-то Ферми

и сделал реактор из блоков, чтобы нейтроны замедлялись вне урана, в графите). Ни природная смесь, ни чистый тяжелый изотоп, ни тем более окись  $UO_2$  не размножают нейтронов. Поэтому лишены почвы всяческие слухи о том, что вот на каком-то руднике по недосмотру собралась критическая масса и произошел взрыв. Не могут взорваться и урановые месторождения, так как уран не встречается вместе углеродом, по крайней мере достаточно чистым (реакторным). Хороший замедлитель – тяжелая вода ( $D_2O$ ), как в современных канадских реакторах, но опять таки в природе этого нет. В большинстве современных реакторов замедлитель – обычная вода. Водород тоже слегка поглощает нейтроны, переходя в дейтерий. Поэтому на природной смеси не получается цепная реакция и с водяным замедлителем. Приходится обогащать уран легким изотопом, чтобы реактор с водяным замедлителем все же работал<sup>1</sup>.

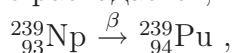
**Разделение изотопов.** Итак, для бомбы и для реактора необходим обогащенный уран. Для обогащения есть несколько способов (газовая диффузия, центрифуги, термодиффузия), некоторые из них мы уже рассматривали в виде задач. Так как массы изотопов различаются мало, разделение – довольно трудный процесс. Например, при газодиффузионном способе за одну ступень отношение концентраций возрастает в 1,0043 раза в самом благоприятном случае, а практически где-то в 1,002. Приходится строить заводы, уже похожие на целые города. Например, только одно здание К-25 Клинтонского завода (Окридж, шт. Теннесси) имело километровые размеры и содержало 16 гектаров диффузионных перегородок, при 3000 ступенях разделения.

Дополнительная очистка делалась на электромагнитных установках. Радиус окружности, по которой движется ион в магнитном поле, тоже зависит от массы. Здесь можно достигать почти полного разделения за один шаг, но зато мала производительность. Например, при токе 1 ампер разделяется  $(3 \cdot 10^9 / 4,8 \cdot 10^{-10}) \cdot 1,6 \cdot 10^{-24} \cdot 235$  г/сек = 2,35 миллиграмма в секунду. Если начать с природной смеси, то для получения 50 кг  $^{235}U$  такая установка должна работать 1000 лет, или надо 1000 установкам работать один год. Практически сначала обогащали природную смесь диффузией до нескольких процентов, а потом применяли электромагнитное разделение. Завод (в том же месте) получился не меньше диффузионного.

**Плутоний.** Одновременно готовился альтернативный вариант. Те нейтроны, которые захватывает в реакторе  $^{238}U$ , не пропадают даром. Получается изотоп  $^{239}U$ , крайне неустойчивый к  $\beta$ -распаду (период полураспада всего 23 мин.):



Образуется 93-й элемент, который практически обнаружил Ферми еще в 1935 г. Так как он следует за ураном, его называли нептунием (по аналогии с порядком планет в Солнечной системе). Нептуний тоже распадается, примерно за двое суток:



уже в 94 элемент – плутоний. Имея четный номер, плутоний стабильнее нептуния –

<sup>1</sup>В далеком прошлом без этого можно было обойтись, см. следующий параграф.

период его полураспада 24000 лет. Плутоний можно накапливать. Он подходит для бомб еще лучше, чем уран – 235 (легче делится нейтронами).

В реакторах значительная доля нейтронов деления поглощается с образованием плутония (выходит порядка одного ядра Pu на два распада урана). На самом деле выход больше, но часть плутония «сгорает», делясь медленными нейтронами, в точности как  $^{235}\text{U}$ . Этим уменьшается добыча, но в то же время поддерживается работа реактора. Таким образом, «негорючий»  $^{238}\text{U}$  перерабатывается в делящийся материал. Плутоний – это другой элемент, и для его отделения от урана есть химические способы (хотя не столь эффективные, потому что оба они – члены группы актинидов, в которой заполняются внутренние электронные оболочки). Все же это легче, чем выделить изотоп, у которого вообще все электроны такие же. Возможный процесс – совместное осаждение с элементом, близким по химическим свойствам, затем растворение осадка, перевод плутония в другое состояние окисления и снова осаждение, после чего плутоний остается в растворе. Несколько таких циклов могут дать достаточно чистое вещество, в том числе и убрать очень опасные осколки деления.

Ферми продемонстрировал возможность работы реактора на природном уране. Его военное значение состояло в перспективе получения плутония. Конечно, масштаб даже первых промышленных реакторов был совершенно другим (например, для охлаждения требовалось примерно такое количество воды, которое потребляет небольшой город). Действительно, один грамм плутония – это порядка  $200 \text{ (МэВ)} \cdot 1,6 \cdot 10^{-6} / (1,6 \cdot 10^{-24} \cdot 239) = 10^{18}$  эрг, то есть грамм плутония в сутки соответствует мощности  $10^6$  Вт. В Хэнфорде на реке Колумбия (шт. Вашингтон) были построены три реактора для наработки плутония (первый заработал в сентябре 1944) и предприятие, где проводилось выделение плутония, под общей вывеской «Хэнфордский механический завод».

## 7.2 Атомная энергетика

Первые реакторы были машинами для накопления плутония. Выделяемое тепло попросту нагревало окружающую среду. Позднее его научились использовать.

**Ядерные реакторы.** Энергетический реактор, как и обычный паровой котел, греет воду и производит пар. В идеале реактор имеет два контура. В первом вода нагревается теплом, выделяющимся при цепной реакции. Нагрев происходит в активной зоне, представляющей из себя решетку **ТВЭЛОВ** – тепловыделяющих элементов, содержащих уран в виде окиси, обогащенной до нескольких процентов. Твэлы имеют длину в несколько метров, толщину порядка сантиметра, и объединяются в тепловыделяющие сборки (ТВС). Загрузка урана обыкновенно порядка сотен тонн. В нормальном режиме твэлы нагреты до нескольких сотен градусов. Вода одновременно служит и теплоносителем (отбирая тепло от твэлов), и замедлителем нейтронов. Активная зона помещена в прочный корпус диаметром в несколько метров. В реакторах с водой под давлением (PWR,

или ВВЭР) поддерживается давление выше линии равновесия вода – пар, чтобы был достаточный теплообмен (с жидкой водой). Типичны условия:  $P = 150$  атм,  $T = 320$  С (при 150 атм температура кипения 342 С, а при 320 С равновесное давление 115 атм). Движение воды обеспечивают циркуляционные насосы (см. рис. 7.1 на стр. 115).

Вода первого контура обычно радиоактивна из-за почти неизбежных дефектов оболочек твэлов. Нагретая вода поступает в парогенератор – теплообменник. Это набор трубок, закрепленных в так называемых трубных досках. Через стенки трубок тепло передается воде второго контура, в норме нерадиоактивной. При этом получается пар, который вращает турбину. Естественно, давление во втором контуре заметно ниже. Остывший пар после турбины охлаждается еще в конденсаторе и уже в жидком виде снова загоняется насосами в парогенератор.

Схема обычной тепловой станции проще, так как нужен только один контур (нет радиоактивности); в ней котел – это и есть парогенератор. Поэтому КПД атомной электростанции несколько ниже. Но зато расход топлива несопоставим: тепловой станции нужен эшелон угля в час, а атомной – примерно вагон твэлов раз в квартал.

**Запаздывающие нейтроны.** Как уже говорилось, активная зона регулируется специальными стержнями – поглотителями нейтронов. Не мешает задуматься, как вообще можно управлять таким быстрым процессом? Для реактора на тепловых нейтронах скорость их порядка километра в секунду, пробег – несколько сантиметров. Для замедления нужно 10 – 100 соударений, так что характерное время меньше миллисекунды. Ясно, что управлять реактором с таким временем разгона нельзя.

Реактор весьма приблизительно описывается уравнением:

$$dN/dt = (K - 1)N/\tau,$$

где  $N$  – число нейтронов,  $K$  – коэффициент их размножения (с учетом того, что часть вылетает, другая часть поглощается в стержнях и т.п.), а время  $\tau$  – доли миллисекунды. Так как каждый нейтрон произошел от реакции, мощность реактора пропорциональна  $N$ . Время разгона будет  $\tau/(K - 1)$ . Миллисекунда будет при большом  $K$ , заметно отличающемся от 1. Для приемлемого значения 100 секунд надо, чтобы  $K$  был равен единице с точностью  $10^{-5}$ . Это тоже трудно выполнимо.

Удобнее вместо  $K$  ввести реактивность  $\rho = (K - 1)/K$ . Поскольку коэффициент размножения всегда близок к 1, уравнение можно упростить:

$$dN/dt = \rho N/\tau,$$

и в таком примитивном приближении время разгона равно  $\tau/\rho$ .

К счастью (или несчастью, как посмотреть) не все нейтроны вылетают мгновенно. Небольшая доля  $\beta$  ( $\approx 1,5\%$  для урана 235) из них запаздывают, вылетая из нестабильных ядер через время порядка секунды (в среднем 2,5 с для  $^{235}\text{U}$ ). Удивительна тонкая компенсация эффектов, которая из характерного времени распада ( $10^{-13}$  с) делает время человеческого масштаба. Ясно, что предвидеть такое нельзя, и поэтому Резерфорд



вполне обоснованно пессимистически оценивал перспективы ядерной энергетики. Если можно так выразиться о явлении природы, это случайность. Именно запаздывание позволяет управлять реактором.

При включении реактора запас реактивности делают довольно большим (десятки процентов). Эта реактивность гасится поглощающими стержнями. Понемногу вынимая их, доводят реактивность до положительной величины (довольно малой доли  $\beta$ ) и позволяют мощности вырасти до желаемого уровня. Номинальная мощность реактора ограничена только возможностями съема тепла, а в принципе можно его разогнать до гораздо большей мощности, что и случается при авариях.

Со временем выгорает  $^{235}\text{U}$  и накапливаются вредные осколки, поглощающие нейтроны. Запас реактивности падает, и приходится вынимать регулирующие стержни. Когда возможности управления исчерпываются, приходится останавливать реактор и перезаряжать его свежим топливом. Это делается два-три раза в год.

**Размножение топлива.** Описанные реакторы **на медленных нейтронах** сейчас во многих странах производят значительную долю электроэнергии. Но они не могут решить глобальную энергетическую проблему, так как активного изотопа ( $^{235}\text{U}$ ) мало в природном уране.

В п. 7.1 упоминалось, что часть нейтронов в реакторе поглощается ядрами  $^{238}\text{U}$ , и в результате нарабатывается плутоний. В типичных реакторах с замедлителем может производиться около 6 ядер плутония на 10 сгоревших ядер  $^{235}\text{U}$  (т.е. **коэффициент воспроизводства** топлива равен 0,6). Этот плутоний частично также делится нейтронами, что приводит к некоторой экономии  $^{235}\text{U}$ . Остальная часть плутония может извлекаться после остановки реактора и использоваться как горючее. Но раз коэффициент воспроизводства заметно меньше 1, такая переработка в плутоний не позволяет использовать сколько-нибудь значительную долю  $^{238}\text{U}$ .

Можно отказаться от замедления и работать на быстрых нейтронах. Такие **быстрые** реакторы сложнее и дороже, но у них есть важное преимущество: коэффициент воспроизводства может быть больше единицы, например, 1,3. Можно говорить уже не о воспроизводстве, а о **размножении** топлива. В принципе быстрые реакторы могут пережечь на энергию и на плутоний весь природный уран, что дает потенциальный выигрыш по запасам в 140 раз. Однако из-за технологических трудностей<sup>2</sup> в мире работает всего одна ядерная электростанция на быстрых нейтронах (Белоярская АЭС в России). В некоторых странах, например, Германии и Франции, уже построенные быстрые реакторы законсервированы из-за опасений аварии и протестов общественности. Некоторую роль здесь сыграло английское название (оказавшееся в нашу эпоху совершенно неудачным) быстрого реактора – размножителя: fast breeder, что вызывает в массовом сознании образ агрегата, с огромной скоростью производящего плутоний.

---

<sup>2</sup>В частности, теплоносителем в быстром реакторе может быть жидкий металл, натрий или свинцовый сплав.



**Безопасность.** Атомная промышленность – одна из наиболее молодых отраслей, и она развивалась уже при современных требованиях безопасности. В идеале при любых условиях радиоактивность не должна выйти наружу. Поэтому современные станции сооружаются внутри герметичного бетонного купола (см. рис. 7.1 на стр. 115), который обязан удерживать в себе все последствия любой аварии.

Существует понятие «барьеров безопасности», последним из которых и является купол. Прочный котел, рассчитанный с запасом на рабочее давление – это тоже барьер. Следующий – оболочка твэлов из циркония или нержавеющей стали. Наконец, первый барьер – это физическое состояние топлива. Сейчас принято использовать окись урана ( $UO_2$ ), то есть химически инертное твердое вещество в виде таблеток, набитых в твэлы. Осколки деления в норме остаются внутри таблетки топлива. В первых реакторах использовался и металлический уран. Он лучше, как топливо (более концентрирован), но крайне опасен при контакте с водой, так как мгновенно воспламеняется.

Кроме физической защиты, важны организационные меры, как грамотное управление реактором, на что существуют строгие правила. Исключительно надежными должны быть все конструкции. Все должно работать как следует – на АЭС нет «пустяков». При любой серьезной угрозе автоматика должна прекратить цепную реакцию.

**Проплавление.** Любой объект должен выдерживать так называемую максимальную проектную аварию (МПА). Для АЭС такой аварией считается разрыв главного трубопровода, питающего котел водой. Тогда вся вода из котла под давлением 150 атм вырывается наружу. По существу, происходит сильный взрыв, каких было много в раннюю эпоху паровых машин. Как уже говорилось, эта вода радиоактивна.

Но опаснее та радиоактивность, что осталась в активной зоне. Конечно, без воды – замедлителя цепная реакция прекращается сама, даже без усилий персонала и автоматики. Это один из факторов внутренней безопасности. Но не только цепная реакция идет в твэлах. Нельзя выключить  $\beta$ -распад осколков деления. Поэтому после глушения реактора в топливе продолжает выделяться тепло с мощностью в несколько процентов от номинальной. Эта активность спадает довольно медленно. Твэлы, оказавшиеся «голыми», без охлаждения водой, начнут разогреваться; топливо расплавится. Получится кипящее озеро на дне котла, которое его проплавит и выльется на бетонный пол реакторного зала. Затем проплавится пол, и сотни тонн топлива вырвутся наружу. Конвекция и пожары разнесут активность на тысячи километров. Это и называют проплавлением, или расплавом активной зоны (core melt down).

Пока полномасштабного проплавления не случилось, хотя не раз до него оставалось совсем немного, а в Чернобыле произошла по меньшей мере равноценная авария. Несколько характерных аварий на атомных станциях рассмотрены в приложении. Эти инциденты, и в особенности чернобыльский, изменили отношение общества к ядерной энергетике. Оценки безопасности, которые давали тысячи лет до первой серьезной аварии, оказались неверны из-за невозможности учесть ошибки персонала. В то же время

последствием аварий стало несомненное повышение уровня безопасности. Вспомним, что в 19 веке часты были взрывы паровых котлов, а сейчас о них как-то забыли. Прогресс технологии часто зависит от масштаба аварий.

**Экология.** При нормальной работе атомная электростанция гораздо экологичнее обыкновенной тепловой. Она потребляет немного топлива, не дает золы, углекислоты, окислов серы и подобных химикатов. Даже чисто радиоактивных выбросов получается меньше, чем у угольной станции (так как уголь всегда содержит радиоактивные примеси). В ядерном цикле единственная радиоактивность, с которой ничего нельзя сделать – это изотопы благородных газов, которые и выпускают в трубу. Все остальное химически связывается и выдерживается в хранилищах, пока не спадет радиоактивность, для последующей переработки. Считается, что в устойчивых геологических формациях отходы, при тщательной упаковке, могут лежать тысячи лет даже без присмотра<sup>3</sup>.

**Естественный ядерный реактор Окло.** Как отмечалось в предыдущем разделе, при современном составе урана цепная реакция возможна только в специальных условиях, достижимых при развитой технологии. Напомним, что легкий изотоп менее устойчив (период полураспада около 0,7 млрд. лет против 4,5 млрд. для  $^{238}\text{U}$ ). Процент  $^{235}\text{U}$  со временем падает. И, значит, в прошлом он был больше. Два миллиарда лет назад доля  $^{235}\text{U}$  была приблизительно 3,7%, т.е. состав урана был примерно тот же, как сейчас в твэлах атомных электростанций. В 1956 г. П. Курода предсказал возможность самопроизвольных цепных реакций в богатых урановых месторождениях в далеком прошлом, в присутствии воды, замедляющей нейтроны. Как водится, никто не поверил.

Однако в 1972 г. в Габоне (Западная Африка) был открыт естественный ядерный реактор. Добытая там урановая руда перерабатывалась во Франции (Габон – это бывшая французская колония). Поскольку уран добывают с целью концентрации  $^{235}\text{U}$ , содержание этого изотопа постоянно проверяется, в том числе это мера против воровства и терроризма. И вот заметили, что в некоторых пробах легкого изотопа нехватает (0,717% вместо 0,72). Вначале предполагалось загрязнение руды обедненным ураном (из «хвостов» переработки). Когда эта возможность не подтвердилась, рассматривались в том числе и такие гипотезы, как занесение отработанного урана из реакторов, построенных древними цивилизациями либо инопланетянами. Наконец, расследование показало, что «дефектный» уран прибывал с месторождения Окло. Месторождение состоит из ряда «линз» руды (каждая поперечником порядка 10 м и толщиной до метра), причем некоторые представляют собой почти чистую  $\text{UO}_2$ . Примерно таковы размеры активной зоны реакторов, созданных человеком. Около 2 млрд лет назад урановая руда собралась на дне древней реки в достаточной концентрации. Вода, проникая в трещины и поры руды, замедляла нейтроны и обеспечивала цепную реакцию. При чрезмерном росте мощности вода выкипала, замедление ослабевало и реакция гасилась. Таким образом

---

<sup>3</sup>Проект хранилища Юкка-Маунтин в США предусматривает безопасное размещение отработанного ядерного топлива на период 10000 лет; в первые 300 лет предусматривается мониторинг.

реакторы саморегулировались; всего в месторождении обнаружено 15 реакторных зон.

Работал реактор порядка полумиллиона лет при очень малой средней мощности (десятки кВт). В некоторых линзах выгорание  $^{235}\text{U}$  достигало 40%. Сейчас осколки деления почти полностью распались. Следов вредного влияния радиации в округе не обнаружено. Правда, есть мнение, что современная жизнь зародилась именно тогда и что реактор мог этому способствовать. К концу 1998 года месторождение выработано, и только небольшая часть реакторной зоны № 2 сохранена в виде природного памятника. Известен еще один естественный реактор в Бангомбе (30 км от Окло, одна реакторная зона). Его судьба (на 1997 г.) зависела от решения французского правительства: надо было выкупить месторождение за 3,5 млн долларов у владельца.

Недавно сотрудник Института гидродинамики СО РАН В.Ф. Анисичкин выдвинул идею о возможности концентрации урана в недрах планет (например, на границе твердого ядра, в  $\approx 1200$  км от центра), где мог образоваться метровый слой окиси. Там в прошлом могла идти реакция, а при внезапном сжатии (скажем, ударом большого метеорита) был возможен и взрыв. Этим можно объяснить формирование необычно большого спутника – Луны. Сейчас, к счастью, взрыв невозможен, так как  $^{235}\text{U}$  осталось мало<sup>4</sup>. Но не исключены «спокойно» работающие реакторы на быстрых нейтронах, в которых критичность поддерживается долгое время за счет переработки урана-238 в плутоний. Глубинные реакторы могут влиять на конвекцию в недрах, движение континентов, генерацию магнитного поля и пр. Похожие идеи развивает независимый американский геофизик J.M. Herndon, только у него реактор поперечником около 8 км находится в центре Земли.

Не на слонах стоит планета,  
не на слонах и черепахах,  
она висит в пустом пространстве,  
усердно бегая по кругу.  
А рядом с ней планеты-сестры,  
а в середине жарко солнце,  
большой костер из водорода  
и прочих разных элементов,  
Кто запалил его? ...

## 7.3 Термоядерные реакции

Если в ядерных электростанциях и атомных бомбах извлекается по сути электрическая энергия (спуск с «кулоновского склона» долины устойчивости при распаде тяжелого ядра), то в термоядерных реакциях уже дело идет о настоящей ядерной энергии, то есть работе ядерных сил притяжения. Топографически это соответствует очень крутому обрыву в долину устойчивости со стороны легких ядер. Происходит синтез ядер из более легких.

---

<sup>4</sup>Впрочем, удар крупного метеорита и сам по себе достаточно опасен.

Простейшая реакция синтеза – это образование дейтона:



( ${}^2\text{H}$  также обозначается D). Выделяется 2,2 МэВ (энергия связи) и, значит, выход на единицу массы примерно тот же, что в реакции деления ( $\approx 1$  МэВ/нуклон). Энергия связи выделяется в виде  $\gamma$ -кванта.

**Слияние легких ядер.** Строго говоря, нейтрон – это не ядро. Настоящая реакция синтеза из ядер, например, такая:



${}^3\text{H}$  – это тритий (Т). Здесь выделяется  $17,6/5 = 3,5$  МэВ/нуклон. Эта реакция идет сравнительно легко и сейчас выглядит перспективной для управляемого синтеза.

Слиянию ядер препятствует электрическое отталкивание. Ядра надо столкнуть достаточно энергично, чтобы преодолеть кулоновский барьер и сблизить их на  $10^{-13}$  см, куда уже достают ядерные силы притяжения. Не обязательно для этого разгонять их до энергии в МэВ, так как ядра – волны и могут проникать под барьер. Считается достаточной температура в 10 кэВ, или 100 млн. градусов. По нашим масштабам это тоже горячо, поэтому реакции синтеза и называют термоядерными.

С ростом заряда исходных ядер растет кулоновский барьер, а выигрыш в энергии связи падает. Поэтому эффективен может быть только синтез из ядер с минимальными зарядами – изотопов водорода и, может быть, He3. По той же причине термоядерные бомбы называют еще водородными.

В реакции  ${}^1\text{H} + \text{n} \rightarrow {}^2\text{H}$  кулоновского барьера нет вообще, так как нейтрон не заряжен. Захват нейтрона происходит, например, в воде с образованием HDO и D<sub>2</sub>O на ядерных реакторах (отчего несколько снижается реактивность). Но нейтроны нестабильны, поэтому их нельзя накапливать в виде некоего «вещества» для переделки в дейтоны. Только в работающих ядерных реакторах нейтроны имеются в заметных количествах; уже говорилось и о других «синтезах» с захватом нейтрона, которые вносят известный вклад в общее энерговыделение реактора.

**Водородная бомба.** Еще в апреле 1943 г. в Лос-Аламосе велась проработка термоядерного оружия. Правда, работа в то время имела низкий приоритет, и это направление отдали Э. Теллеру, как самому неуправляемому сотруднику теоретического отдела (которым руководил Г. Бете).

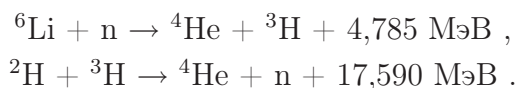
После удачного создания атомных бомб началась гонка вооружений. 29 августа 1949 г. прошло успешное испытание атомной бомбы в СССР. Почти сразу воздушная разведка США обнаружила в атмосфере следы этого ядерного взрыва. Бомба, как теперь известно, была практически копией американской (в Лос-Аламосе работало несколько информаторов). Однако потенциал СССР был достаточен и для изготовления оригинальных конструкций.

Мощность атомной бомбы ограничена (с ростом массы заряда падает КПД) на уровне сотен килотонн. Наоборот, синтез идет тем лучше, чем крупнее бомба. Явилась идея

использовать атомный взрыв для создания высоких температур, необходимых для термоядерного взрыва. Активнее других в США действовал Теллер, в СССР основной вклад принадлежит Сахарову, Зельдовичу и Тамму.

В 1951 г. в США проведено первое испытание с заметным выделением термоядерной энергии. Дейтерий и тритий были в жидком состоянии, и из-за малой вообще плотности жидкого водорода изделие имело размер железнодорожного вагона. Во втором испытании – «Майк», атолл Эниветок, 1 ноября 1952 г.) выделилась энергия 3 Мт (по другим данным 20 Мт). Была создана параллельная Лос-Аламосской структура (Ливерморская лаборатория им. Лоуренса) для ускорения работ по водородным бомбам, и началась работа по транспортабельной конструкции.

12 августа 1953 г. произведено советское испытание водородной бомбы («слойка» Сахарова). По предложению И. Гинзбурга, взрывчатым материалом был дейтерид лития-6 (LiD) – твердое вещество, не требующее криогенных операций. Идет цепь реакций:



Тритий делается прямо в бомбе из лития и нейтрона. Первые нейтроны дает атомный взрыв, а остальные берутся из второй реакции синтеза. Слои LiD чередовались со слоями урана-238. При высоких температурах взрыва тепло распространяется очень быстро (излучением). Поэтому температура выравнивается раньше, чем давление (в противоположность нашему житейскому опыту). Уран имеет 92 электрона и, когда полностью ионизируется, дает 93 частицы. Все они «давят» и сжимают слои термоядерного топлива, облегчая реакцию. К тому же часть 14–МэВ нейтронов расшибает ядра урана, с выделением энергии распада.

«Сухая» бомба опередила американцев, хотя основные идеи в США были высказаны раньше. Частично опережение можно объяснить тем, что американцы больше сил израсходовали на совершенствование ядерных бомб и на криогенные конструкции.

Современные водородные бомбы основаны на «третьей» идее. Как сейчас известно, это идея имплозии (сжатия внутрь) термоядерного заряда. Атомная бомба взрывается не внутри основного заряда, а рядом. Излучение взрыва нагревает и испаряет поверхность заряда. Возникает реактивная сила, сжимающая вещество. При остановке сжатия происходит сильный нагрев, и зажигается реакция.

Бомба может иметь также внешнюю урановую оболочку для более полной реакции, дополнительного энергосодержания и эффективного заражения местности. 1 марта 1954 г. на атолле Бикини была испытана такая трехступенчатая бомба. Из-за непредвиденной смены направления ветра под радиоактивный след попали японские рыбаки, один из которых вскоре умер. Сейчас на вооружении стоят в основном «чистые» бомбы, так как точность средств доставки стала очень высокой и поражать большие территории не нужно (радиация опасна и для нападающей стороны).

Советская бомба, основанная на «третьей идее», испытана осенью 1955 г. под Семи-

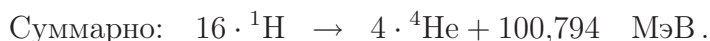
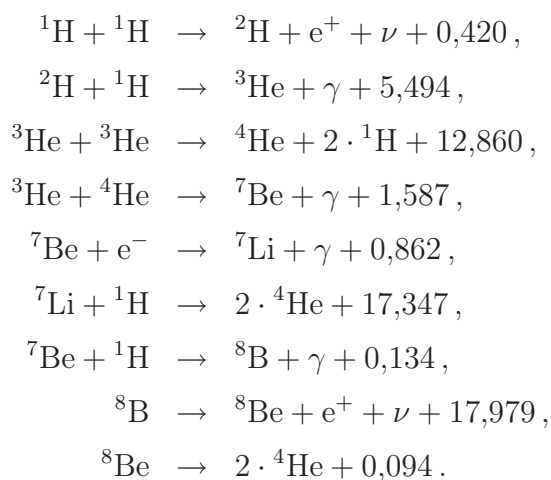
палатинском. 30 октября 1961 г. на Новой Земле взорвана 50-Мт бомба.

**Реакции в звездах.** Термоядерные реакции обеспечивают энерговыделение и светимость звезд. Все знают, что водород выгорает и образуется гелий. Но реакция типа



крайне маловероятна, так как должны встретиться в одной точке четыре ядра. Кроме того, в такой реакции не сохраняется заряд (хотя можно прибавить в правую часть два позитрона).

По современным представлениям, в Солнце протекают следующие реакции (pp-цикл, каждый раз встречается два ядра либо распадается одно):



Выходит 6,3 МэВ на нуклон, то есть близко к максимуму, достижимому по энергии связи. Когда водород выгорит, начинается образование более тяжелых ядер, с меньшим энерговыделением. До этого Солнцу еще далеко. Но на небе известны выгоревшие звезды. В зависимости от начальной массы может выйти белый карлик, сверхновая (остаток от нее – нейтронная звезда) либо черная дыра. Солнце – это звезда второго поколения, собравшаяся из остатков взрыва сверхновых.

Наблюдения потока нейтрино от Солнца давали заметный дискрепанс – недостаток нейтрино. Это могло быть связано с недостатками теории либо близким ( $\sim 10^6$  лет) угасанием Солнца. Однако в последние годы показано, что часть солнечных нейтрино «по дороге» переходит в другой вид, чем и объясняется несоответствие. Эти опыты доказали, что нейтрино имеет ненулевую массу (хотя и меньшую 2 эВ).

**Управляемый термоядерный синтез.** Эта крупная проблема уже разбиралась по частям ранее. Напомним, что есть два основных подхода – магнитные ловушки и инерциальный синтез. Сейчас уже условия приблизились к нулевому выходу энергии (breakeven – затраты равны результату). Будущее зависит от финансирования, а основное обоснование проектов вытекает из экологии.



А.Д. Сахаровым высказана идея использовать для контролируемого производства энергии взрывной термоядерный синтез – подземные взрывы настоящих термоядерных зарядов. При этом снимается проблема неустойчивости, основная для микрочарядов, подвергаемых обжатию лазерами и пучками. Известно, что термоядерные «изделия» взрываются вполне надежно.

Сейчас это направление развивается сотрудниками РФЯЦ–ВНИИТФ. Предлагаются котлы взрывного сгорания (КВС), на 10 – 25 кт. Взрывы должны происходить с интервалами порядка 1 часа<sup>5</sup>. Расход металла на подземную взрывную камеру должен быть порядка  $8 \cdot 10^6$  тонн, диаметр камеры 130 м, высота 250 м. Предполагается чисто дейтериевое топливо, так что из дефицитных материалов расходоваться будет только небольшое количество урана в атомных «детонаторах». Не вызывает сомнений работоспособность такой системы, но очевидны и барьеры (в большой степени психологические), которые мешают эксперименту такого масштаба. Впрочем, размеры сооружения не так велики – меньше типичной атомной станции.

## 7.4 Элементарные частицы

Напомним, что слово «частицы» условное, и правильнее было бы говорить о квантах полей. Из таковых сейчас известны:

**Фотон.** Квант электромагнитного поля. Совпадает со своей же античастицей. Масса нулевая, спин 1.

**Лептоны.** Электрон, мюон и  $\tau$ -лептон, плюс к каждому из них свое нейтрино – электронное, мюонное и  $\tau$ -лептонное. Это кванты лептонного поля. Заряды  $e$ ,  $\mu$  и  $\tau$  единичные, нейтрино не заряжены. Масса мюона 207 электронных масс,  $\tau$ -лептон еще тяжелее, массы нейтрино пока не измерены (возможно, нулевые). Спины у всех  $1/2$ . Еще вводят лептонные заряды (электронный, мюонный и пр.) как свойства, которым обладают лептоны и только они. Античастицы истинные, то есть отличаются от частиц.

**Кварки.** Есть шесть видов кварков:  $u, d; s, c; t, b$  плюс античастицы. Эти три пары симметричны тройкам лептонов.

Кварк	u	d	s	c	b	t
$q/e$	2/3	-1/3	-1/3	2/3	-1/3	2/3
$m/m_e$	$\sim 300$	300	500	1500	5000	45000

Из кварков в некотором смысле построены **адроны** – тяжелые сильно взаимодействующие (через ядерные силы) частицы. Кварки имеют дробный заряд ( $1/3$  либо  $2/3$ ). Например, состав протона  $uud$ , а нейтрона  $udd$ . Если кварков три, это называется **барион** (спин полуцелый), а если два, то получается **мезон** (целый спин). Мезон состоит

<sup>5</sup>10 килотонн тротилового эквивалента примерно соответствуют сжиганию 1 килотонны угля или мазута, то есть одному железнодорожному эшелону.



из кварка и антикварка (не обязательно такого же). Например,  $\pi^+ = u\bar{d}$ ,  $\pi^- = \bar{u}d$ ,  $\pi^0$  составлен из  $u\bar{u}$  и  $d\bar{d}$  в равных долях. Заряд адрона всегда выходит целый.

Кварки найдены только внутри адронов в опытах по рассеянию (аналог опытов Резерфорда при высоких энергиях – до ТэВ). Не наблюдалось кварков в свободном состоянии. Это связано с огромными силами взаимодействия. Когда, скажем, мезон пытаются разделить, совершаемая работа достаточна для рождения новых кварков, которые соединяются с исходными уже в два мезона. Можно сказать, что вакуум «лопается» и в месте разрыва появляются новые частицы. До некоторой степени это напоминает разрезание магнита – получается не два полюса, а два новых диполя.

Из этого экспериментального факта можно оценить силу взаимодействия кварков:  $FR \sim m_p c^2$ , где  $R \approx 10^{-13}$  см (ядерный масштаб). Получаем  $F \sim 1,6 \cdot 10^{-24} \cdot 9 \cdot 10^{20} / 10^{-13} \approx 16 \cdot 10^9$  дин = 16 тонн силы! На этом уровне уже теряет смысл понятие «состоять из». Пока мы не сломали протон, внутри него есть три кварка. Но когда пытаемся их пересчитать, вынимая поодиночке, мы этим самым делаем еще несколько. Такие огромные силы переносятся так называемыми глюонами (как электромагнитные – фотонами).

Поскольку в некоторых адронах оказались одинаковые кварки, что противоречит принципу Паули, пришлось приписать кваркам еще и «цвет», красный, синий либо зеленый. Любая свободная комбинация бесцветна, то есть содержит кварки всех трех цветов (барион) либо пару цвет-антицвет (мезон). Соответственно глюонов бывает 8 видов, глюоны маркируются парой цветов (парное взаимодействие кварков). Например, может быть красно/анти-синий глюон, который из синего кварка делает красный.

Осталось добавить переносчиков слабых взаимодействий (**Z и W – бозоны**) – это что-то вроде тяжелых фотонов. Некоторых частиц в теперешней систематике недостает, и их пока что ищут.

## Приложение. Ядерные инциденты

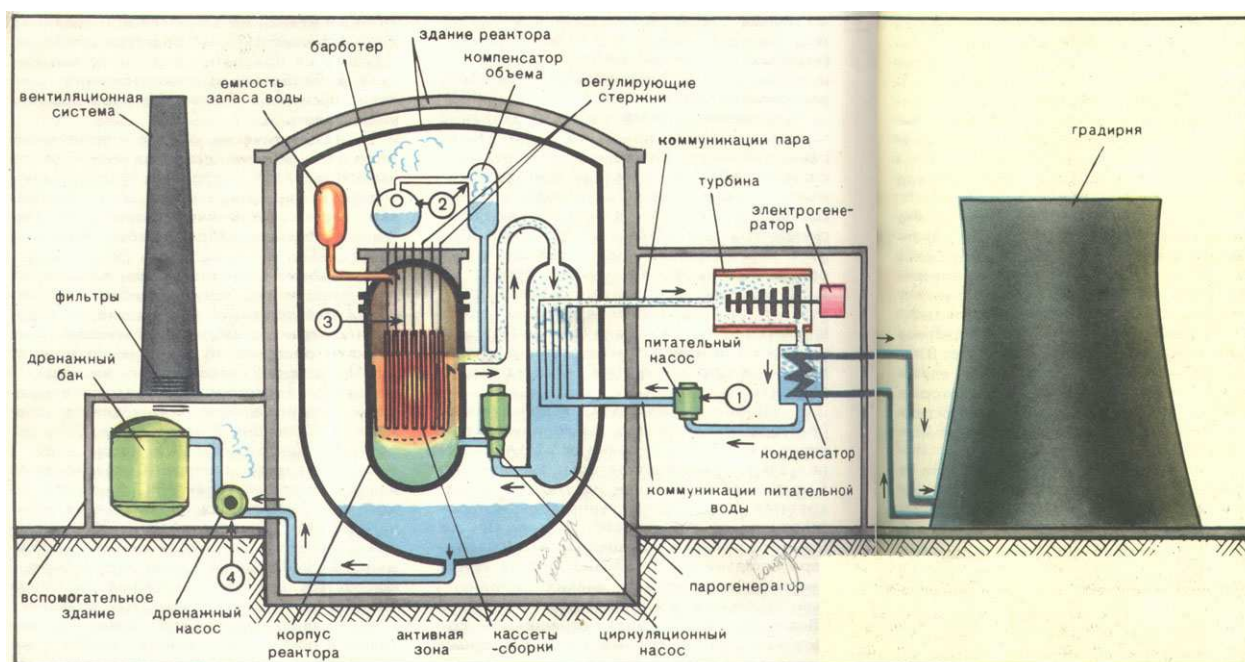
Ясно, что полной гарантии от катастрофы дать невозможно. Вдруг случится землетрясение, или купол разобьет метеорит. А иногда может испортиться любое оборудование. Редкие, но опасные неприятности оценивались с помощью вероятностного анализа. Выходило, что станции должны работать тысячи лет до первой серьезной катастрофы.

В теории это звучит хорошо. Но теперь мы знаем, что эти оценки не оправдались. Несколько аварий оказались неожиданно тяжелыми.

**Browns Ferry.** 22 марта 1975 произошел пожар на АЭС Browns Ferry (шт. Алабама). Причиной была свечка, с которой рабочий пытался (по сквозняку) найти дыру в стене, через которую проходили связки кабелей. Вблизи одной из связок поток воздуха направил пламя на пенопластовую затычку, которая тут же вспыхнула. Пламя распространилось по кабелям, от чего не сработала автоматика тушения пожара.

Персонал пытался потушить пожар своими силами. После безуспешных усилий вызвали пожарных и сообщили на пульт управления. Операторы остановили реакторы. Что же до пожарной команды, ее телефон оказался перепутан. Пока нашли правильный телефон и пока прибыли пожарные, огонь разгорелся как следует. Потом оказалось, что нехватает респираторов и что неизвестно, как тушить пожар на атомной станции, да еще среди электропроводки. Наконец, стали тушить попросту водой, и только через семь часов после возгорания пожар был ликвидирован. К счастью, помещение реактора не было затронуто, хотя дальнейшее промедление было уже крайне опасно. Жертв и утечки радиоактивности не было.

**ТМІ.** Ночью 28 марта 1979 произошла серьезная авария на втором блоке АЭС Three Mile Island возле Гаррисбурга, шт. Пенсильвания. Схема станции показана на рис. 7.1<sup>6</sup>. Это типичная АЭС с водо-водяным реактором.



Перегорел предохранитель, и неожиданно остановились питательные насосы (① на рис. 7.1), гонящие воду во втором (нерадиоактивном) контуре. Парогенератор стал перегреваться. На этот случай имелось запасное водоснабжение. Автоматически включились резервные насосы, но вода не пошла. Потом выяснилось, что во время недавних учений по технике безопасности закрыли задвижку на питающей трубе, да так и оставили. Когда наконец это заметили и пустили воду, парогенератор уже высох, перегрелся, и холодная вода вызвала температурные напряжения, от чего в нем появились трещины.

Тем временем активная зона перестала отдавать тепло в парогенератор и также перегрелась. Автоматика опустила регулирующие стержни и остановила цепную реак-

<sup>6</sup>Рисунок и описание аварии взяты из статьи: М.С. Бабаев, И.И. Кузьмин, В.А. Легасов, В.А. Сидоренко. Проблемы безопасности на атомных электростанциях. Природа, 1980, №6, стр. 30–43.

цию. Но остаточное тепло продолжало выделяться, давление в реакторе росло. Как положено, открылся клапан, чтобы стравить избыток давления. И здесь случилась неприятность – клапан не закрылся сам, и давление продолжало падать. Неисправность клапана отобразилась на мнемосхеме реактора. Но именно это место оказалось закрыто плакатом по технике безопасности, так что операторы не заметили неисправности. Мало того, манометр показывал опасное давление, так как стрелку зашкалило и она так и залипла в этом положении. Тогда операторы выпустили еще часть воды, от чего температура еще повысилась и началась опасная вибрация главных насосов первого контура, не приспособленных для перекачки пароводяной смеси.

Все еще не понимая, что происходит, операторы выключили насосы. В результате падения давления и остаточного нагрева в верхней части активной зоны образовался паровой пузырь. В норме оставшаяся в реакторе вода обеспечивает достаточный теплообмен с твэлами за счет естественной конвекции. Но в пузыре теплоотдача, конечно, гораздо ниже.

Только утренняя смена разобралась в ситуации, закрыла клапан, повысила давление и ликвидировала газовый пузырь. Своевременное охлаждение локализовало аварию, но часть активной зоны расплавилась. На полу реакторного зала оказалось  $\approx 3000$  тонн радиоактивной воды. Блок вышел из строя навсегда, и только через десять лет стала возможна разборка активной зоны с помощью роботов.

Заражения местности не произошло (только некоторое количество изотопов благородных газов, которые не поглощаются никакими фильтрами, вышло в атмосферу через трубу). Никаких последствий для персонала, населения и окружающей среды не удалось заметить. Тем не менее, имела место некоторая паника в городе, отчасти вызванная избытком информации плохого качества.

**Чернобыльская авария.** Реакторы РБМК (большой мощности канальный) развились из реакторов, производящих плутоний. Сначала военные реакторы приспособили для производства электричества, а потом разработали уже гражданские РБМК. Замедление нейтронов происходит в графитовой кладке, пронизанной топливными каналами (которых около 1600). В каналы вставлены трубы, по которым циркулирует вода – теплоноситель. Прямо в воду опущены пучки твэлов (тепловыделяющие сборки). Реактор одноконтурный (полученный пар прямо идет в турбину) и не имеет внешнего прочного корпуса. Предполагалось, что трубы обеспечивают прочность, а при повреждении оболочки твэла данная труба выключается и производится ремонт.

26 апреля 1986 г. произошла самая крупная (пока что) авария. 4 блок АЭС останавливался на профилактику, и по этому случаю планировался эксперимент по использованию инерции турбогенератора после отключения пара для питания аварийных потребностей станции.

Вначале предполагалось провести отключение от уровня 1000 МВт (тепловых). Из-за задержки отключения на несколько часов (до 23.00 25.04) и последующей ошибки

оператора мощность внезапно упала до 30 МВт, и только к 1.00 25.04 удалось довести мощность до 200 МВт. Уместно отметить, что при резком снижении мощности реактор стремится сам выключиться из-за ксенонового отравления. Один из осколков деления  $^{135}\text{J}$  распадается в  $^{135}\text{Xe}$ , имеющий рекордное сечение поглощения нейтронов (2,7 миллиона барн, то есть порядка квадрата длины волны нейтрона). В норме этот ксенон выжигается нейтронами деления. При падении же мощности нейтронов становится меньше, а запасенный иод продолжает распадаться (характерное время – несколько часов), и концентрация ксенона растет. Из-за этого падает реактивность, еще сильнее спадает мощность, и такой shut down прогрессирует (говорят, что реактор проваливается в иодную яму). Для поддержания мощности операторам пришлось вытащить почти все регулирующие стержни.

При этом управляемость реактора резко уменьшилась. Предусмотрено несколько автоматических блокировок, которые в норме должны были бы остановить реактор. Стремясь продолжить злосчастный эксперимент, операторы отключили эти блокировки, разумеется, нарушив все правила и инструкции. На момент начала опыта характерное время реактора уменьшилось до 20 секунд, и приходилось регулировать мощность каждые несколько секунд оставшимися 6 стержнями (в норме надо не менее 30).

Наконец в 01.23.04 эксперимент начался (был отключен турбогенератор; при этом еще отключили аварийную защиту по остановке всех генераторов, чтобы иметь возможность повторить эксперимент в случае неудачи). Замедляя вращение, генератор стал давать меньше энергии, в том числе на циркуляционные насосы. Меньше стал поток охлаждающей воды, стало быть повысилась ее температура. Нагрев повысил содержание паровых пузырьков и снизил плотность жидкости.

Для дальнейшего важно, что реактор РБМК имеет положительный паровой коэффициент реактивности. Вода в каналах приводит к некоторому поглощению нейтронов. Поэтому, когда при кипении ее плотность падает из-за пузырей, падает и поглощение, и мощность растет. (В водо-водяном реакторе та же вода служит и замедлителем. Поэтому при ее выкипании реактор, наоборот, останавливается. В РБМК основной замедлитель – графит, а эффект замедления водой слабее эффекта поглощения). В результате развился лавинообразный процесс роста мощности (примерно до 100 номинальных значений). Давление пара порвало трубы и разрушило корпус (около 01.24.00).

В 01.23.40 оператор нажал кнопку аварийной защиты. Стержни пошли вниз, но застряли из-за тепловых деформаций. Надо добавить, что нижние отрезки стержней не поглощают, но зато вытесняют воду из каналов. Поэтому на начальном этапе движения стержни еще добавляли реактивности (в норме этот эффект безопасен).

Далее начался пожар, потушенный без всяких предосторожностей. Затем реактор засыпали с вертолетов песком, свинцом, доломитом и т.п. Под реактором залили толстую бетонную плиту, а вокруг возвели объект «Укрытие». (Размеры здания – сотни метров, что сопоставимо с пирамидой Хеопса). Несколько процентов топлива рассея-

лось по всей Европе. Поскольку загрузка урана около 200 тонн, радиоактивный выброс гораздо сильнее, чем при взрыве атомной бомбы.

Причина катастрофы – недостатки конструкции реактора (отсутствие внутренней стабильности) в сочетании с лихостью персонала и неверной организацией. Добавим, что незадолго до события атомные электростанции, как доказавшие свою надежность, из министерства среднего машиностроения были переданы местным энергетикам (в данном случае украинскому министерству энергетики). Это с очевидностью снизило компетентность персонала.