ИЗДАТЕЛЬСТВО «НАУКА»
МОСКВА 1966

Я. П. Терлецкий

Парадоксы теории относительности

УДК 530.12(088.3)

Книга посвящается подробному рассмотрению тех вопросов специальной теории относительности, которые при обычном ее изложении наиболее трудны для усвоения.

Для выяснения спорных и неясных вопросов теории относительности в книге используется эффективный метод парадоксов. Противоречия, возникающие в теории, формулируются в виде парадоксов, а затем эти парадоксы анализируются и разъясняются.

Автор стремится показать, что глубокий теоретический анализ теории относительности способен вскрыть в ней новые, ранее не замечавшиеся стороны, развитие которых может привести к возникновению новых, перспективных направлений теоретической физики. Одним из таких направлений является теория частиц отрицательной и мнимой собственной массы. В книге показано, что частицы мнимой собственной массы, движущиеся с сверхсветовой скоростью, могут быть непротиворечиво введены в теорию, если «принцип причинности» рассматривать лишь как следствие закона возрастания энтропии.

В книге исследуются возможные свойства частиц отрицательной и мнимой массы и те макроскопические следствия, к которым приводит допущение существования таких частиц.

Книга рассчитана на широкие круги научных работников, аспирантов и студентов физических факультетов, а также может быть полезна преподавателям физики в высшей и средней школе.

2-3-2
477-66

ПРЕДИСЛОВИЕ

Известно, что представления о пространстве и времени, диктуемые теорией относительности, сильно отличаются от наших привычных, обиходных представлений об этих всеобщих формах существования материи. С точки зрения привычных представлений многие следствия теории относительности выглядят как «парадоксы». Парадоксальными представляются следствия об изменении длин движущихся масштабов и хода движущихся часов. Непривычно ограничение, накладываемое на скорость распространения сигналов. Много странного и даже парадоксального можно усмотреть в релятивистских законах движения, например в соотношениях, связывающих энергию, импульс и массу.

Книга посвящается изложению, анализу и разъяснению подобных «парадоксов» с точки зрения релятивистских представлений о пространстве и времени. Путем такого разрешения парадоксов удается наиболее наглядно уяснить особенности и самую суть релятивистских представлений.

Прежде всего рассматриваются парадоксы (главным образом философского порядка), связанные с противоречием, существующим между названием и истинным содержанием теории относительности. Выясняется, что теория относительности, понимаемая как современная физическая теория взаимосвязанных пространства и времени, не дает никаких оснований для позитивистских выводов.

Анализ кинематических парадоксов (масштабов и часов, законов сложения скоростей и их четырехмерной интерпретации), проведенный в книге, приводит к более отчетливому пониманию четырехмерной геометрической сущности представлений о пространстве и времени теории относительности. Анализ парадоксов релятивистской механики также показывает, что только четырехмерный подход к законам природы позволяет правильно понять возникающее в теории относительности объединение законов сохранения энергии импульса и собственной массы в единый всеобщий закон.

К существенно новым теоретическим результатам приводит анализ парадоксов, связанных с утверждением о предельности скорости света. Выясняется, что принцип причинности, правильно понимаемый как следствие или как выражение закона возрастания энтропии, не запрещает абсолютно сверхсветовые скорости переноса физических воздействий. Запрещаются лишь макроскопические сверхсветовые сигналы, но не микроскопически обратимые процессы. Выяснение этого обстоятельства позволяет сделать ряд теоретических предсказаний о возможности существенно новых физических объектов и процессов, которые можно пытаться обнаружить экспериментально.

К числу таких объектов относятся частицы отрицательной массы. В книге анализируются те следствия, к которым привело бы открытие подобных частиц. В частности, показано, что в космосе могут реализоваться принципиально новые чрезвычайно мощные энергетические источники, если частицы отрицательной массы реально существуют в природе.

Вопросы, рассмотренные в этой книге, были изложены в факультативном курсе лекций, прочитанном для студентов III курса физического факультета МГУ в весеннем семестре 1962 г. Этот же курс с некоторыми изменениями был повторен в 1964 г. Большая помощь при составлении конспекта курса 1962 г. была оказана мне аспирантом Ю. П. Рыбаковым, которому я пользуюсь случаем выразить благодарность. Этот конспект, изданный ЛФОП физического факультета МГУ в 1962 г., был использован при написании данной книги.

Приношу также искреннюю благодарность профессору К. П. Станюковичу, просмотревшему рукопись и сделавшему ценные критические замечания.

Профессор МГУ Я. П. Терлецкий

ВВЕДЕНИЕ

Парадоксы, т. е. неожиданные следствия или выводы теории, противоречащие сложившимся ранее представлениям, играют особую роль в процессе развития науки. При разрешении того или иного теоретического парадокса приходится обращаться к наиболее принципиальным положениям теории и иногда пересматривать или уточнять связанные с ней представления. Таким образом, теоретические парадоксы в процессе их разрешения представляют некоторую внутреннюю причину развития теории, способствующую ее логическому совершенствованию, а иногда даже выяснению границ применимости и путей дальнейшего обобщения.

Конечно, основным для развития всякой теории являются факты, получаемые из экспериментов и наблюдений. Однако одни только факты не могут сами по себе подтвердить, уточнить или изменить теорию, если они не приводят к подтверждению и уточнению или пересмотру логической структуры теории. Поэтому для развития теории имеет большое значение раскрытие внутренних противоречий и их разрешение. Противоречия же в теории отчетливее всего обнаруживаются, когда они возникают в форме тех или иных парадоксов. Таким образом, анализ теоретических парадоксов не является самоцелью, а представляет лишь средство для выяснения истинного содержания теории, уточнения отдельных ее положений и отыскания путей ее дальнейшего развития.

В этой книге нас будут интересовать парадоксы теории относительности. При этом мы не будем касаться проблем, связанных с теорией гравитационного поля, хотя и рассмотрим некоторые парадоксы, связанные с переходом в неинерциальные системы отсчета, которые обычно принято относить к области общей теории относительности, поскольку под последней принято подразумевать теорию относительности, распространенную на неинерциальные системы отсчета при наличии гравитационных полей.

Многие противоречия возникают в теории относительности из-за стандартного способа ее изложения по тому классическому образцу, который был дан еще Эйнштейном. Со времени первой работы Эйнштейна теория относительности пополнилась большим количеством новых представлений. В результате многочисленных приложений выяснилось главное содержание теории. Выяснилось, что некоторые представления, считавшиеся основными в период зарождения теории, оказались в действительности лишь вспомогательными средствами, использованными для построения теории. Оказалось также, что теория может быть построена на базе различных постулатов. Выяснилось, иначе говоря, что постулаты Эйнштейна не могут отождествляться с самим содержанием теории относительности.

Глубокий анализ содержания теории относительности важен именно сейчас, когда намечается новый этап крутой ломки теоретических представлений в связи с проникновением внутрь самих элементарных частиц и открытием принципиально новых физических процессов в космосе, протекающих в радиогалактиках и сверхзвездах или квазарах.

Мы увидим, что анализ вопроса о предельности скорости сигналов в теории относительности приведет нас к пересмотру содержания так называемого принципа причинности и к общему выводу о принципиальной возможности существования частиц, имеющих отрицательные и даже мнимые собственные массы. Но если такие частицы действительно существуют в природе, то их открытие приведет к радикальной перестройке всей существующей физической картины мира. А это в свою очередь приведет к новым открытиям, умножающим власть человека над природой.

I

НАЗВАНИЕ И СОДЕРЖАНИЕ ТЕОРИИ ОТНОСИТЕЛЬНОСТИ

§ 1. Противоречие между названием и содержанием

Многие парадоксы и противоречивые суждения о смысле тех или иных выводов теории относительности возникают из противоречия, существующего между названием и содержанием теории относительности.

Название «теория относительности» как бы означает, что содержанием теории является «относительность». Относительность же, положенная в основу, трудно отличима от релятивизма, т. е. учения об относительности наших знаний, относительности в смысле субъективности. Такая трактовка физической теории весьма импонирует позитивистам и всячески рекламируется философами-идеалистами. Они видят в теории относительности пример физической теории, противоречащей материализму. Отсюда делается обобщающий вывод о якобы несовместимости современной физики с диалектическим материализмом.

С другой стороны, философы-материалисты, поверившие в то, что содержанием теории относительности действительно является «относительность», пытаются либо вообще отвергать теорию относительности как якобы противоречащую материализму, либо обосновывают понятие какой-то «физической относительности», отличной от релятивизма и якобы лежащей в основе современной физики. Пытаясь провести последнюю точку зрения, стремятся изъять из теории субъекта, заменив его измерительными приборами, не замечая при этом, что любой измерительный прибор лишь вместе с познающим субъектом приобретает принципиально отличные черты от всех остальных изучаемых объектов материального мира.

Эти споры философов становятся совершенно беспредметными, когда к теории относительности подходят как к физической теории с определенным конкретным содержанием. Выясняется, что содержанием теории относительности является физическая теория пространства и времени, учитывающая существующую между ними взаимосвязь геометрического характера. При этом «относительность» оказывается имеющей подчиненный характер (иногда даже чисто иллюстративный), не отличающейся от «относительности» в классической механике и других разделах теоретической физики.

Таким образом, для устранения ошибочных представлений, навязываемых названием, необходимо наиболее полно вскрыть содержание теории.

§ 2. Происхождение названия «теория относительности»

Название «теория относительности» возникло из наименования основного принципа, или постулата, положенного Пуанкаре и Эйнштейном в основу всех теоретических построений новой теории пространства и времени.

Название же «принцип относительности», или «постулат относительности», возникло как отрицание представления об абсолютной неподвижной системе отсчета, связанной с *неподвижным эфиром*, вводившимся для объяснения оптических и электродинамических явлений.

Дело в том, что к началу XX в. у физиков, строивших теорию оптических и электромагнитных явлений по аналогии с теорией упругости, сложилось ложное представление о необходимости существования абсолютной неподвижной системы отсчета, связанной с электромагнитным эфиром. Зародилось, таким образом, представление об абсолютном движении относительно системы, связанной с эфиром, представление, противоречащее более ранним воззрениям классической механики (принцип относительности Галилея). Опыты Майкельсона и других физиков опровергли, однако, эту теорию «неподвижного эфира» и дали основание для формулировки противоположного утверждения, которое и получило название «принципа относительности». Так это название вводится и обосновывается в первых работах Пуанкаре и Эйнштейна.

В своей основополагающей работе «К электродинамике движущихся тел» [1] Эйнштейн пишет: «...неудавшиеся попытки обнаружить движение Земли относительно «светоносной среды» ведут к предположению, что не только в механике, но и в электродинамике никакие свойства явлений не соответствуют понятию абсолютного покоя, и даже более того,— к предположению, что для всех координатных систем, для которых справедливы уравнения механики, имеют место те же самые электродинамические и оптические законы, как это уже доказано для величин первого порядка. Мы намерены это положение (содержание которого в дальнейшем будет называться «принципом относительности») превратить в предпосылку...» (см. сборник «Принцип относительности». Лоренц, Пуанкаре, Эйнштейн, Минковский. Изд. ОНТИ, 1935, стр. 134).

Аналогично формулирует «постулат относительности», независимо от Эйнштейна, и Пуанкаре [2]: «Эта невозможность показать опытным путем абсолютное движение Земли представляет по-видимому закон природы; мы естественно приходим к тому, чтобы принять этот закон, который мы назовем постулатом относительности, и принять его без оговорок» (там же, стр. 51).

Итак, название «принцип *относительности*» возникло как *отрицание* представления об абсолютной системе отсчета и *абсолютного* движения относительно этой системы, возникшего в конце XIX в. в связи с попытками механического объяснения электромагнитных явлений.

То, что это название не отражает существа исходного положения теории и всего ее содержания, неоднократно отмечалось учеными, разрабатывавшими теорию относительности.

Так, крупнейший советский теоретик Л. И. Мандельштам ([4, стр. 172]) в своих лекциях по теории относительности разъяснял: «Название «принцип относительности» одно из самых неудачных. Утверждается независимость явлений от неускоренного движения замкнутой системы. То, что это называется «принципом относительности», вводит, как увидим потом, в заблуждение».

На неудачность названия указывал и один из творцов теории относительности, раскрывший ее содержание в четырехмерной геометрической форме, Герман Минковский. Еще в 1908 г. он утверждал [3]: «...термин «постулат относительности», для требования инвариантности по отношению к группе G_c, кажется мне слишком бледным. Так как смысл постулата сводится к тому, что в явлениях нам дается только четырехмерный в пространстве и времени мир, но что проекции этого мира на пространство и на время могут быть взяты с некоторым произволом, мне хотелось бы этому утверждению скорее дать название: п о с т у л а т а б с о л ю т н о г о м и р а (или коротко: мировой постулат)» (см. сборник «Принцип относительности», стр. 192).

Таким образом, мы видим, что названия «принцип относительности» и «теория относительности» не отражают истинного содержания теории.

§ 3. Теория относительности как современная теория пространства-времени

Содержание теории относительности как четырехмерной физической теории пространства и времени впервые отчетливо было вскрыто Германом Минковским в 1908 г. Лишь опираясь на эти представления, Эйнштейн сумел в 1916 г. построить общую теорию пространства-времени, включающую явление гравитации (так называемую общую теорию относительности). В этом параграфе дается краткий обзор тех основных положений теории относительности, которые характеризуют ее как четырехмерную теорию пространства и времени.

Основным отличием представлений о пространстве и времени теории относительности от представлений ньютоновской физики является органическая в з а и м о с в я з ь п р о с т р а н с т в а и в р е м е н и. Эта взаимосвязь раскрывается в формулах Лоренца преобразования координат и времени при переходе от одной системы отсчета к другой.

Каждое физическое явление протекает в пространстве и времени и не может быть изображено в нашем сознании иначе как в пространстве и времени. Пространство и время суть формы существования материи. Они являются в с е о б щ и м и формами, в которых существует материя. Никакой материи не существует вне пространства и времени.

Конкретным способом изображения пространства и времени является *система отсчета*, т. е. координатно-временное многообразие чисел x, y, z, t, составляющее воображаемую координатную сетку и временную последовательность всех возможных пространственных и временных точек.

Одно и то же пространство и время могут изображаться различными координатно-временными сетками, т. е. различными системами отсчета. Вместо чисел x, y, z, t пространство-время может изображаться числами, x', y', z', t', причем эти числа не произвольны, а связаны с x, y, z, t совершенно определенного вида формулами преобразования, которые и выражают свойства пространства-времени.

Итак, каждое возможное изображение пространства и времени связывается с определенной системой отсчета Σ. Система же отсчета может быть связана с реальным *телом отсчета*, каковым может быть любое твердое тело. Координаты x, y, z могут быть связаны с конкретными точками тела отсчета или определены другими операциями, проделываемыми по отношению к телу отсчета. Моменты времени t также могут быть связаны с показаниями конкретных часов, расставленных в различных точках тела отсчета. Тело отсчета необходимо, очевидно, для проведения конкретных измерений пространственных и временных отношений.

Не следует, однако, отождествлять систему отсчета с телом отсчета, как это предлагают физики, придерживающиеся операционалистической философии. Физики при изображении явлений пользуются *любыми* системами отсчета, в том числе и такими, с которыми невозможно связать какое-либо реальное тяжелое твердое тело, могущее играть роль тела отсчета. Основанием для такого произвольного выбора служит представление о полном равноправии всех мыслимых систем отсчета. Следовательно, выбор системы отсчета является лишь выбором способа изображения пространства и времени, удобным для отображения исследуемого явления.

Итак, система отсчета выбирается познающим субъектом для изображения физических явлений. Она выбирается в зависимости от постановки задачи, а не задается как некоторое якобы обязательно присутствующее при всяком материальном процессе большое тяжелое тело.

Если избраны две системы отсчета Σ и Σ', каждая из которых изображает одно и то же пространство-время, то, как это установлено в теории относительности, координаты в системах Σ и Σ' связаны так, что *интервал* s_{12}, определяемый для двух разобщенных событий как

$$-s_{12}^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 - \\ - c^2 (t_1 - t_2)^2, \qquad (3.1)$$

остается одинаковым при переходе от Σ к Σ', т. е.

$$-s_{12}^2 = (x_2' - x_2')^2 + (y_1' - y_2')^2 + (z_1' - z_2')^2 - \\ - c^2 (t_1' - t_2')^2. \qquad (3.2)$$

Иначе говоря, s_{12} является инвариантом преобразований Лоренца, связывающих координаты и время в Σ и Σ':

$$x' = \frac{x - vt}{\sqrt{1 - v^2/c^2}}, \quad y' = y, \quad z' = z, \quad t' = \frac{t - \frac{v}{c^2}x}{\sqrt{1 - v^2/c^2}}. \qquad (3.3)$$

Из (3.3), так же как из (3.1), (3.2), следует, в частности, о т н о с и т е л ь н о с т ь о д н о в р е м е н н о с т и пространственно разобщенных событий, т. е. если для двух событий в системе Σ имеем $t_2 = t_1$, $x_2 \neq x_1$, $y_2 = y_1$, $z_2 = z_1$, то в системе Σ', движущейся со скоростью v, будем иметь

$$t_2' - t_1' = -\frac{v}{c^2}(x_2' - x_1'), \quad x_2' \neq x_1', \quad y_2' = y_1', \quad z_2' = z_1',$$

т. е. (3.4)

$$t_2' \neq t_1'.$$

Эта относительность одновременности приводит к сокращению движущихся масштабов и изменению хода движущихся часов, как это мы подробно разберем в § 10 и 11.

В этих свойствах пространственно-временных координат и отражается существо новых представлений о пространстве и времени, связанных в единое геометрического типа многообразие, многообразие с особой, определяемой (3.1) и (3.2), четырехмерной, псевдоевклидовой геометрией, геометрией, в которой время тесно связано с пространством и не может рассматриваться независимо от последнего, как это видно из (3.4).

Из этих же представлений вытекают важнейшие следствия для законов природы, выражаемые в свойстве ковариантности (т. е. неизменяемости формы) любых физических процессов по отношению к преобразованиям четырехмерных пространственно-временных координат.

В требовании ковариантности также отражается представление о пространстве-времени как о едином четырехмерном многообразии.

Так представляют себе физики, конкретно применяющие теорию относительности, ее реальное содержание. При этом понятие «относительность» приобретает лишь смысл возможной множественности пространственно-временных изображений явлений при абсолютности содержания, т. е. законов природы.

§ 4. Бесплодность попыток связать теорию относительности с философским релятивизмом

В значительной степени вследствие неправильного названия теории относительности истинное ее содержание как новой теории взаимосвязанных пространства и времени часто подменяется представлением об относительности всякого физического знания. Возможность произвольного выбора системы отсчета и изображения явлений толкуется как относительность содержания физических теорий, относительность в смысле отношения к *наблюдателю*, связанному с системой отсчета.

Отсюда подводится фундамент под представление, что вообще наблюдаемый физический объект не существует без наблюдателя, так как его проявление, фиксируемое измерительными приборами, зависит от точки зрения наблюдателя.

Эта идея может быть сколько-нибудь последовательно проведена только в том случае, если стать на точку зрения, что содержанием любой физической теории является установление связи одних измерений с другими. Эта точка зрения никогда, однако, не была творческой, ибо всякая теория движется вперед посредством гипотез о сущности самих явлений, а не о способах их измерений.

В области теории относительности упор на относительность в субъективистском смысле задерживал развитие новых представлений и понятий о пространстве и времени. Так, например, безуспешные поиски общей относительности в смысле Маха не способствовали развитию теории гравитационного поля как теории четырехмерного риманова континуума неоднородной кривизны.

II

ПОСТУЛАТЫ ЭЙНШТЕЙНА И ПРЕОБРАЗОВАНИЯ ЛОРЕНЦА

§ 5. Роль постулатов Эйнштейна и их кажущаяся противоречивость

Преобразования Лоренца, отражающие свойства пространства-времени, первоначально были получены Лоренцом посредством анализа уравнений электромагнитного поля, а затем выведены Эйнштейном исходя из двух постулатов — принципа относительности и принципа постоянства скорости света. Эйнштейн сформулировал эти постулаты следующим образом:

1. Законы, по которым изменяются состояния физических систем, не зависят от того, к которой из двух координатных систем, находящихся одна относительно другой в равномерном поступательном движении, эти изменения состояния относятся.

2. Каждый луч света движется в «покоящейся» системе координат с определенной скоростью c независимо от того, испускается ли этот луч света покоящимся или движущимся телом.

Значение этих постулатов для дальнейшего развития теории пространства-времени состояло в том, что их принятие прежде всего означало отказ от старых представлений о пространстве и времени, как о многообразиях, не связанных органически одно с другим.

Принцип относительности сам по себе не представлял чего-либо абсолютно нового, так как он содержался и в ньютоновской физике, построенной на базе классической механики. Распространение его на любые физические процессы вызывало возражения только с позиций механической теории электромагнитного и светового эфира.

Принцип постоянства скорости света также предлагал лишь возврат к доэфирным представлениям и в области электромагнитных явлений, т. е. не был чем-то абсолютно неприемлемым с точки зрения ньютоновских представлений о пространстве и времени.

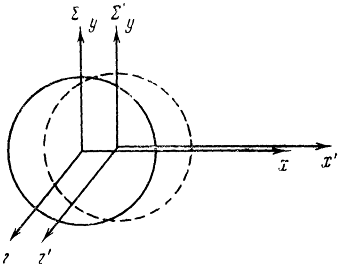

Рис. 1

Однако эти два принципа, взятые вместе, привели к противоречию с конкретными представлениями о пространстве и времени, связанными с механикой Ньютона. Это противоречие обычно иллюстрируется в форме следующего парадокса (рис. 1).

Пусть в системе отсчета Σ в начальный момент $t = 0$ в точке, совпадающей с началом координат $x = y = z = 0$, произошла вспышка источника света. В последующий момент времени t фронт световой волны в силу закона постоянства скорости света распространился до сферы радиуса $R = ct$ с центром в начале координат системы Σ. Однако в соответствии с постулатами Эйнштейна это же явление мы можем рассмотреть и с точки зрения системы отсчета Σ', движущейся равномерно и прямолинейно вдоль оси x, так что ее начало координат и направления всех осей совпадали в момент времени $t = 0$ с началом координат и направлениями осей первоначальной системы Σ. В этой движущейся системе соответственно постулатам Эйнштейна за время t свет также распространится до сферы радиуса $R = ct$, однако в отличие от предыдущей сферы центр новой сферы должен лежать в начале координат системы Σ', а не Σ. Несовпадение этих сфер, т. е. одного и того же физического явления, представляется чем-то совершенно парадоксальным и неприемлемым с точки зрения существующих представлений. Кажется, что для разрешения парадокса надо отказаться либо от принципа относительности, либо от принципа постоянства скорости света.

Теория относительности предлагает, однако, совершенно иное разрешение парадокса, состоящее в том, что события, одновременные в одной системе отсчета Σ, неодновременны в другой движущейся системе Σ', и наоборот. Тогда одновременные события, состоящие в достижении световым фронтом сферы, определяемой уравнением

$$x^2 + y^2 + z^2 = c^2t^2, \tag{5.1}$$

не являются одновременными с точки зрения системы Σ', где одновременны другие события, состоящие в достижении тем же световым фронтом точек сферы, определяемой уравнением

$$x'^2 + y'^2 + z'^2 = c^2t'^2. \tag{5.2}$$

Таким образом, одновременность пространственно разобщенных событий перестает быть чем-то абсолютным, как это принято считать в повседневном макроскопическом опыте, а становится зависящей от выбора системы отсчета и расстояния между точками, в которых происходят события. Эта *относительность одновременности* пространственно разобщенных событий свидетельствует о том, что пространство и время тесно связаны одно с другим, так как при переходе от одной системы отсчета к другой, физически эквивалентной, промежутки времени между событиями становятся зависящими от расстояний (нулевой промежуток становится конечным, и наоборот).

Итак, постулаты Эйнштейна помогли нам прийти к новому фундаментальному положению в физической теории пространства и времени, положению о тесной взаимосвязи пространства и времени и об их нераздельности. В этом, пожалуй, состоит главное значение постулатов Эйнштейна, ибо сам вывод преобразований Лоренца, как мы увидим ниже, может быть сделан различными путями с различной степенью использования первого и второго постулатов.

Некоторые авторы отводят особую роль постулату о постоянстве скорости света и формулируют его таким образом, чтобы изобразить этот постулат как основное содержание теории относительности. Основным аргументом в пользу такого мнения, по-видимому, является та особая роль, которая отводится Эйнштейном световым сигналам, при помощи которых устанавливается (т. е.

определяется) одновременность пространственно разобщенных событий. Световой сигнал, распространяющийся всегда только со скоростью света, приравнивается, таким образом, к некоторому инструменту, устанавливающему связь между временными отношениями в различных системах отсчета, без которого якобы понятия одновременности разобщенных событий и времени вообще теряют смысл.

Необоснованность такого истолкования содержания теории относительности легко доказывается, если обратиться к одному из возможных выводов преобразований Лоренца, опирающемуся только на постулат относительности и вместо постулата о постоянстве скорости света использующему лишь допущение о зависимости массы тела от скорости.

§ 6. Механическое определение одновременности пространственно разобщенных событий

Прежде чем доказать, что преобразования Лоренца можно вывести, опираясь исключительно на механику и на содержащийся в ней принцип относительности, необходимо показать, что одновременность событий и синхронизация часов могут быть установлены и без световых сигналов, чисто механическими экспериментами.

Один из способов синхронизации при помощи световых сигналов тождественных часов, находящихся в пространственно разобщенных точках, состоит в регистрации световых импульсов, пришедших в эти точки от импульсного источника света, расположенного на равных расстояниях от указанных точек, например, в центре соединяющего их отрезка прямой. В силу постулируемого постоянства скорости света оба импульса должны одновременно достигнуть равноудаленных от источника точек.

Нетрудно видеть, что в данном способе синхронизации совершенно несущественно то, что сигналы движутся именно со скоростью света c. Здесь важно лишь, что они движутся с одинаковой скоростью. Следовательно, световые импульсы могут быть заменены, например, точечными массами, движущимися из центра отрезка прямой к его концам с одинаковой скоростью u. Важно лишь осуществить такой процесс отталкивания этих масс, чтобы без дополнительных измерений была бы гарантирована одинаковость их скоростей.

Закон сохранения количества движения позволяет осуществить такой процесс, если разлетающиеся массы одинаковы и в процессе расталкивания не участвует какое-либо третье тело, забирающее на себя часть количества движения. Очевидно, таких процессов симметричного расталкивания одинаковых тел можно придумать очень много. Таким образом, в принципе возможно осуществить синхронизацию пространственно разобщенных часов при помощи чисто механического устройства.

Подчеркнем, что здесь речь идет лишь о принципиальной возможности, ибо практически световая сигнализация, по-видимому, во всех случаях будет более удобной и надежной.

Может показаться, что предложенный механический способ синхронизации часов способен указать истинную, безотносительную одновременность пространственно разобщенных событий, одинаковую во всех системах отсчета. Однако так обстояло бы дело, если бы был справедлив закон классической механики о неизменности массы тел. На самом же деле опыт показывает, что масса (определяемая как отношение импульса к скорости) зависит от скорости, монотонно возрастая с увеличением последней. Следовательно, и закон сохранения количества движения приобретает вид

$$\sum_k m_k \mathbf{u}_k = \mathbf{P} = \mathbf{const} \tag{6.1}$$

в «неподвижной» системе отсчета и

$$\sum_k m'_k \mathbf{u}'_k = \mathbf{P}' = \mathbf{const}' \tag{6.2}$$

в «движущейся» со скоростью $\mathbf{v}$ системе, причем

$$m'_k \neq m_k, \tag{6.3}$$

так как входящая в (6.1) масса является монотонно возрастающей функцией скорости, т. е. $m(\mathbf{u})$.

Нетрудно показать, что в силу (6.1—6.3) одновременный в «неподвижной» системе отсчета приход в равноудаленные точки симметрично оттолкнувшихся одно от другого одинаковых по массе тел окажется неодновременным с точки зрения движущейся системы отсчета.

Пусть два шарика одинаковой массы отталкиваются один от другого в момент $t = 0$ по часам, расположенным в начальной точке x_0, и движутся со скоростями $-u$ и $+u$ к точкам x_1 и x_2, расположенным на одинаковых расстояниях l от точки x_0 (рис. 2, *а*). В момент времени $t = l/u$ оба шарика достигнут точек x_1 и x_2, в которых расположены преграды I и II (рис. 2, *б*). Таким образом,

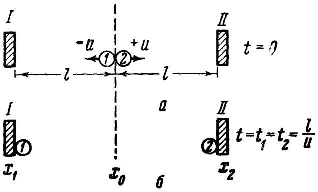

Рис. 2

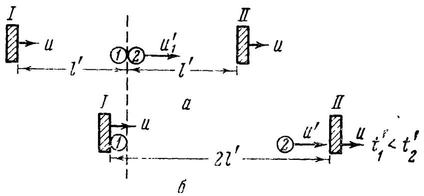

Рис. 3

в избранной системе отсчета шарики одновременно достигнут равноудаленных от исходной точки преград.

Рассмотрим этот же процесс с точки зрения другой системы отсчета Σ', движущейся влево со скоростью $-u$ (рис. 3). Относительно этой системы отсчета преграды I и II движутся со скоростью u вправо. До момента расталкивания оба шарика также движутся со скоростью u вправо и имеют полный импульс $2m(u)\,u$. В момент $t' = 0$ по часам, расположенным в системе Σ' и проходящим в этот момент мимо точки x_0, шарики расталкиваются. При этом левый приобретает скорость $-u$, т. е. останавливается, а правый приобретает некоторую общую скорость u' которую можно определить из закона сохранения полного импульса. Согласно этому закону, очевидно:

$$P' = 2m(u)\,u = m(u')\,u', \tag{6.4}$$

откуда

$$u' = 2u\,\frac{m(u)}{m(u')}, \tag{6.5}$$

следовательно, скорость u' не равна $2u$, как это было бы по классической механике с неизменной массой.

Пусть расстояние между преградами I и II в системе Σ' равно $2l'$, тогда по часам этой системы левый шарик достигнет преграды I (вернее преграда I достигнет неподвижного в данном случае шарика I) в момент времени

$$t_1' = l'/u. \tag{6.6}$$

Правый же шарик достигнет преграды II, пройдя путь $2l'$, в момент времени

$$t_2' = \frac{2l'}{u'} = \frac{2l'}{2u} \cdot \frac{m(u')}{m(u)} = t_1' \frac{m(u')}{m(u)}. \tag{6.7}$$

Поскольку в силу монотонного возрастания массы со скоростью

$$m(u') > m(u) \text{ при } u' > u, \tag{6.8}$$

постольку согласно (6.7)

$$t_2' > t_1'. \tag{6.9}$$

Следовательно, принцип относительности в механике может быть совмещен с фактом возрастания массы со скоростью только в том случае, если признать о т н о с и т е л ь н о с т ь одновременности разобщенных событий, т. е. принять, что пространственно разобщенные события, одновременные в одной системе отсчета ($t_2 = t_1$), являются неодновременными в другой движущейся системе отсчета ($t_2' > t_1'$).

§ 7. Вывод преобразований Лоренца без постулата о постоянстве скорости света

Выведем преобразования Лоренца, опираясь лишь на «естественные» допущения о свойствах пространства и времени, содержавшиеся еще в классической (доэфирной) физике, опиравшейся на общие представления, связанные с классической механикой. Примем как аксиомы следующие требования:

1) *изотропность пространства*, т. е. все пространственные направления равноправны;

2) *однородность пространства и времени*, т. е. независимость свойств пространства и времени от выбора начальных точек отсчета (начала координат и начала отсчета времени);

3) *принцип относительности*, т. е. полная равноправность всех инерциальных систем отсчета.

Различные системы отсчета лишь по-разному изображают одно и то же пространство и время как всеобщие формы существования материи. Каждое из этих изображений обладает одинаковыми свойствами. Следовательно, связь между координатами и временем в одной «неподвижной» системе (x, y, z, t) с координатами и временем в другой «движущейся» системе (x', y', z', t'), т. е. формулы преобразования координат и времени не могут быть произвольными. Установим те ограничения, которые накладывают «естественные» требования на вид функций преобразования

$$
\begin{aligned}
x' &= f_1(x, y, z, t),\\
y' &= f_2(x, y, z, t),\\
z' &= f_3(x, y, z, t),\\
t' &= f_4(x, y, z, t).
\end{aligned}
\tag{7.1}
$$

I. Вследствие однородности пространства и времени преобразования должны быть линейными.

Действительно, если бы производные функций f_1, f_2, f_3, f_4 по x, y, z и t не были бы константами, а зависели от x, y, z и t, то и разности $x_2' - x_1'$, $y_2' - y_1'$, $z_2' - z_1'$, $t_2' - t_1'$, выражающие проекции расстояний между точками *1* и *2* в «движущейся» системе, зависели бы не только от соответствующих проекций $x_2 - x_1$, $y_2 - y_1$, $z_2 - z_1$, $t_2 - t_1$ в «неподвижной» системе, но и от значений самих координат x, y, z, t, что противоречило бы требованию независимости свойств пространства от выбора начальных точек отсчета.

Если положить, что проекции расстояний вида

$$\xi' = x_2' - x_1' = f_1(x_2, \ldots) - f_1(x_1, \ldots) \tag{7.2}$$

зависят только от проекций расстояний в неподвижной системе, т. е. от

$$\xi = x_2 - x_1, \tag{7.3}$$

но не зависят от x_1, то

$$\partial\xi'/\partial x_1 = 0 \text{ при } \xi = \text{const}, \tag{7.4}$$

т. е.

$$\frac{\partial f_1(x_1 + \xi, \ldots)}{\partial x_1} - \frac{\partial f_1(x_1, \ldots)}{\partial x_1} = 0, \tag{7.5}$$

или

$$\partial f_1/\partial x_1 = \text{const}. \tag{7.6}$$

Аналогично можно доказать, что производные f_1 по всем другим координатам y_1, z_1, t_1 также равны константам, а следовательно, и вообще все производные f_1, f_2, f_3, f_4 по x, y, z, t суть константы. Итак, преобразования (7.1) линейны.

II. Выберем «движущуюся» систему Σ' таким образом, чтобы в начальный момент $t = 0$ точка, изображающая ее начало координат, т. е. $x' = y' = z' = 0$, совпадала с точкой, изображающей начало координат «неподвижной» системы, т. е. $x = y = z = 0$ (рис. 4), а скорость движения системы Σ' была бы направлена только по x. Если мы также учтем требование изотропности пространства, то линейные преобразования, для системы отсчета Σ', выбранной указанным образом, запишутся в виде

$$
\begin{aligned}
x' &= k(v)(x - vt), \\
y' &= \lambda(v)y, \qquad (7.7) \\
z' &= \lambda(v)z, \\
t' &= \mu(v)t + \alpha(v)x.
\end{aligned}
$$

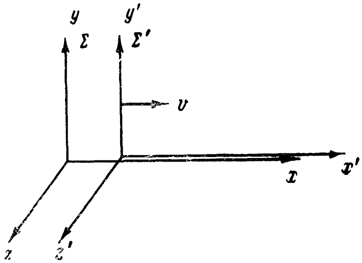

Рис. 4

Здесь отсутствуют члены, содержание y и z в выражениях x' и t' в силу изотропности пространства и наличия единственного выделенного направления вдоль оси x соответственно постановке задачи. На этом же основании в выражениях для y' и z' отсутствуют члены, пропорциональные соответственно z и y, а коэффициенты λ при y и z одинаковы. Члены, содержащие x и t, отсутствуют в выражениях для y' и z' в силу того, что ось x' (т. е. линия $y' = z' = 0$) все время совпадает с осью x (т. е. линией $y = z = 0$). Последнее было бы невозможно, если бы y' и z' зависели от x и t.

III. Изотропность предполагает также симметричность пространства. В силу же симметрии ничто не должно измениться в формулах преобразования, если изменить знаки у v и x, т. е. одновременно изменить направление оси x и направление движения системы Σ'.

Слецовательно,

$$\begin{aligned} -x' &= k(-v)(-x+vt), \\ y' &= \lambda(-v)y, \\ z' &= \lambda(-v)z, \\ t' &= \mu(-v)t - \alpha(-v)x. \end{aligned} \tag{7.8}$$

Сравнивая (7.7) с (7.8), получаем

$$\begin{aligned} &k(-v) = k(v), \quad \alpha(-v) = -\alpha(v), \\ &\mu(-v) = \mu(v), \quad \lambda(-v) = \lambda(v). \end{aligned} \tag{7.9}$$

Вместо $\alpha(v)$ удобно ввести другую функцию $\eta(v)$ так, чтобы α выражалась через η и μ посредством соотношения

$$\alpha(v) = -\frac{v}{\eta(v)}\mu(v). \tag{7.10}$$

Очевидно, согласно (7.10), $\eta(v)$ — симметричная функция, т. е. $\eta(-v) = \eta(v)$.

Используя (7 10), преобразования (7.8) можно записать в виде

$$\begin{aligned} x' &= k(v)(x - vt), \\ y' &= \lambda(v)y, \\ z' &= \lambda(v)z, \\ t' &= \mu(v)\left[t - \frac{v}{\eta(v)}x\right], \end{aligned} \tag{7.11}$$

причем все входящие в эти формулы коэффициенты $k(v)$ $\lambda(v)$, $\mu(v)$ и $\eta(v)$ суть симметричные функции v.

IV. В силу принципа относительности обе системы «движущаяся» и «неподвижная» абсолютно эквивалентны, а поэтому обратные преобразования от системы Σ' к Σ должны быть тождественны прямым (от Σ к Σ'). Обратные преобразования должны отличаться лишь знаком скорости v, так как система Σ движется относительно системы Σ вправо со скоростью v, а система Σ движется относительно системы Σ' (если последнюю считать неподвижной) влево со скоростью $-v$. Следовательно, обратные преобразования должны иметь вид

$$\begin{aligned} x &= k(-v)[x' - (-v)t'], \\ y &= \lambda(-v)y', \\ z &= \lambda(-v)z', \\ t &= \mu(-v)\left[t' - \frac{(-v)}{\eta(-v)}x'\right]. \end{aligned} \tag{7.12}$$

Сравнивая (7.12) с (7.11), получаем

$$\lambda(v)\lambda(-v) = 1. \tag{7.13}$$

Но в силу симметрии $\lambda(-v) = \lambda(v)$ и, следовательно, $\lambda^2 = 1$, т. е. $\lambda = \pm 1$. Очевидно, имеет смысл лишь знак плюс, так как знак минус давал бы при $v = 0$ перевернутую по y и z систему. Следовательно,

$$\lambda = 1. \tag{7.14}$$

Замечая, что коэффициенты k, μ и η тоже симметричные функции v, первое и последнее из уравнений (7.11) и (7.12) можно записать в виде

$$x' = k(x - vt) \ldots (A), \quad x = k(x' + vt') \ldots (a),$$

$$t' = \mu\left(t - \frac{v}{\eta}x\right) \ldots (B), \quad t = \mu\left(t' + \frac{v}{\eta}x'\right) \ldots (b).$$

Умножая (A) на μ, (B) на vk и складывая, получим

$$\mu x' + vkt' = \mu k\left(1 - \frac{v^2}{\eta}\right)x,$$

$$x = \frac{x'}{k(1 - v^2/\eta)} + \frac{vt'}{\mu(1 - v^2/\eta)}.$$

Сравнивая это выражение с (a), получаем

$$k = \frac{1}{k(1 - v^2/\eta)}, \quad k = \frac{1}{\mu(1 - v^2/\eta)};$$

откуда имеем

$$\mu = k, \quad k^2 = \frac{1}{1 - v^2/\eta}. \tag{7.15}$$

Следовательно, извлекая квадратный корень и замечая, что знак минус так же, как и для λ, не имеет смысла, получаем

$$\mu(v) = k(v) = \frac{1}{\sqrt{1 - v^2/\eta(v)}}. \tag{7.16}$$

Итак, преобразования приобретают вид

$$x' = k(v)(x - vt), \quad y' = y, \quad z' = z,$$

$$t' = k(v)\left(t - \frac{v}{\eta(v)}x\right), \tag{7.17}$$

или подробнее

$$x' = \frac{x - vt}{\sqrt{1 - v^2/\eta}}, \quad y' = y, \quad z' = z, \quad t' = \frac{t - vx/\eta}{\sqrt{1 - v^2/\eta}}, \tag{7.18}$$

где $\eta(v)$ — неизвестная пока функция v.

V. Для определения вида $\eta(v)$ обратимся вновь к *принципу относительности* Очевидно, преобразования (7.17) должны быть универсальными и применимыми при любых переходах от одних систем отсчета к другим. Таким образом, если мы дважды перейдем от системы Σ к Σ' и от Σ' к Σ'', то полученные формулы, связывающие координаты и время в системе Σ'' с координатами и временем в Σ, должны также иметь вид преобразований (7.17). Это вытекающее из принципа относительности требование в совокупности с предыдущими требованиями обратимости, симметрии и т. д. означает, что преобразования должны составлять группу.

Воспользуемся этим требованием групповости преобразований. Пусть v_1 — скорость системы Σ' относительно системы Σ и v_2 — скорость системы Σ'' относительно системы Σ'. Тогда согласно (7.17)

$$x' = k(v_1)(x - v_1 t), \qquad x'' = k(v_2)(x' - v_2 t'),$$
$$t' = k(v_1)\left(t - \frac{v_1 x}{\eta(v_1)}\right), \qquad t'' = k(v_2)\left(t' - \frac{v_2 x'}{\eta(v_2)}\right).$$

Выражая x'' и t'' через x и t, получаем

$$\begin{aligned} x'' &= k(v_2)\,k(v_1)\left[x - v_1 t - v_2\left(t - \frac{v_1}{\eta(v_1)}\,x\right)\right], \\ t'' &= k(v_2)\,k(v_1)\left[t - \frac{v_1}{\eta(v_1)}\,x - \frac{v_2}{\eta(v_2)}\,(x - vt)\right]. \end{aligned} \tag{7.19}$$

Согласно сформулированному выше требованию эти же преобразования должны записываться в виде (7.17), т. е.

$$\begin{aligned} x'' &= k(v_3)(x - v_3 t), \\ t'' &= k(v_3)\left(t - \frac{v_3}{\eta(v_3)}\,x\right). \end{aligned} \tag{7.20}$$

Коэффициенты, стоящие при x в первой из формул (7.20) и при t во второй из формул (7.20), одинаковы. Следовательно, в силу тождественности (7.19) и (7.20) должны быть одинаковы и коэффициенты, стоящие при x в первой из формул (7.19) и при t во второй из формул (7.19), т. е.

$$k(v_2)\,k(v_1)\left[1 + \frac{v_1 v_2}{\eta(v_1)}\right] = k(v_2)\,k(v_1)\left[1 + \frac{v_2 v_1}{\eta(v_2)}\right]. \tag{7.21}$$

Последнее равенство может быть удовлетворено только при

$$\eta(v_1) = \eta(v_2) = \text{const}. \qquad (7.22)$$

VI. Итак, в преобразованиях (7.18) η является константой, имеющей размерность квадрата скорости. Величина и даже знак этой константы не могут быть определены без привлечения каких-либо новых допущений, опирающихся на опытные факты.

Если положить $\eta = \infty$, то преобразования (7.18) превращаются в известные преобразования Галилея

$$x' = x - vt,\ y' = y,\ z' = z,\ t' = t. \qquad (7.23)$$

Эти преобразования, справедливые в механике малых скоростей, не могут быть, однако, приняты как точные преобразования, справедливые при любых скоростях тел, когда становится заметным изменение массы тел со скоростью.

Действительно, как мы видели в § 6, учет изменения массы со скоростью приводит к необходимости принять положение об относительности одновременности разобщенных событий. Последнее же несовместимо с преобразованиями Галилея (7.23). Таким образом, константа η должна быть выбрана конечной.

Из опыта известно, что при больших скоростях уравнения механики точки имеют вид

$$\frac{d}{dt}(m\mathbf{u}) = \mathbf{f}, \text{ где } m = \frac{m_0}{\sqrt{1 - u^2/c^2}}; \qquad (7.24)$$

m_0 — собственная масса, совпадающая с массой частицы при малых скоростях ($u \ll c$); c — константа, имеющая размерность скорости и численно равная $3 \cdot 10^{10}$ *см/сек*, т. е. совпадающая со скоростью света в пустоте. Этот опытный факт трактуется как зависимость массы от скорости, если массу определять как отношение импульса тела к его скорости.

Константа c^2, входящая в (7.24), имеет такую же размерность, какую имеет константа η, входящая в формулы преобразования координат и времени (7.18). Естественно поэтому положить

$$\eta = c^2, \qquad (7.25)$$

поскольку в экспериментально полученную зависимость массы от скорости не входит никакая иная константа, имеющая размерность квадрата скорости. Принимая (7.25), преобразования (7.18) записывается в виде

$$x' = \frac{x - vt}{\sqrt{1 - v^2 / c^2}}, \quad y' = y, \quad z' = z, \quad t' = \frac{t - \frac{v}{c^2} x}{\sqrt{1 - v^2 / c^2}}. \tag{7.26}$$

Пуанкаре назвал эти преобразования координат и времени **преобразованиями Лоренца***.

В силу обратимости обратные преобразования Лоренца, очевидно, должны быть записаны в виде

$$x = \frac{x' + vt'}{\sqrt{1 - v^2 / c^2}}, \quad y = y', \quad z = z', \quad t = \frac{t' + \frac{v}{c^2} x'}{\sqrt{1 - v^2 / c^2}}. \tag{7.27}$$

Примененные нами соображения размерности для выбора константы η не вполне, однако, однозначны, так как вместо соотношения (7.25) с таким же правом можно было бы выбрать

$$\eta = -c^2. \tag{7.28}$$

Оказывается, однако, что совпадающие с опытом уравнения механики (7.24) могут быть получены лишь как следствия преобразований Лоренца (7.26) и не могут быть совмещены с преобразованиями, получающимися из допущения (7.28). Действительно, известно, что уравнениями механики, опирающимися на преобразования Лоренца, являются уравнения Минковского **, согласно которым масса увеличивается со скоростью по формуле (7.24). Если же в качестве преобразований координат выбрать

$$x' = \frac{x - vt}{\sqrt{1 + v^2 / c^2}}, \quad y' = y, \quad z' = z, \quad t' = \frac{t + \frac{v}{c^2} x}{\sqrt{1 + v^2 / c^2}}, \tag{7.29}$$

---

* Г. А. Лоренц получил эти преобразования до Пуанкаре и Эйнштейна в 1904 г. [5] (см. сб. Лоренц, Пуанкаре, Эйнштейн, Минковский. «Принцип относительности». ОНТИ, 1935, стр. 21). Однако впервые в виде (7.26) они были записаны Пуанкаре, назвавшими их преобразованиями Лоренца.

** Подробнее см. § 14.

то соответствующие уравнения Минковского дадут убывающую со скоростью массу m, что противоречит опыту.

Итак, не обращаясь к постулату о постоянстве скорости света в пустоте, не ссылаясь на электродинамику и не используя свойств световых сигналов, для определения одновременности, мы вывели преобразования Лоренца, используя лишь представление об однородности и изотропности пространства и времени, принцип относительности и формулу зависимости массы от скорости.

Обычно, следуя пути, намеченному еще в первой работе Эйнштейна, вместо формулы зависимости массы от скорости используют постулат о постоянстве скорости света в пустоте. Согласно этому постулату при переходе от системы Σ к системе Σ' должно оставаться инвариантным уравнение

$$x^2 + y^2 + z^2 - c^2t^2 = 0, \tag{7.30}$$

описывающее фронт световой волны, распространяющейся из начала координатной системы Σ. Легко убедиться в том, что уравнение

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = 0. \tag{7.31}$$

после подстановки формул преобразования (7.30) не изменяет своего вида, т. е. (7.31) переходит в (7.30) лишь в том случае, если $\eta = c^2$.

Мы применили, однако, иной вывод, не использующий постулат о постоянстве скорости света с тем, чтобы показать, что преобразования Лоренца могут быть получены независимо от способа сигнализации, избранного для синхронизации часов, измеряющих время. Физики могли бы вообще ничего не знать о скорости света и о законах электродинамики, однако могли бы получить преобразования Лоренца, анализируя факт зависимости массы от скорости и исходя из механического принципа относительности. Таким образом, преобразования Лоренца (7.26) выражают общие свойства пространства и времени для любых физических процессов. Эти преобразования, как это выяснилось в процессе доказательства, составляют непрерывную группу, называемую г р у п п о й Л о р е н ц а. В этом факте в наиболее общем виде и отображаются свойства пространства и времени, раскрытые теорией относительности.

§ 8. Вывод преобразований Лоренца, основывающийся на постулате о постоянстве скорости света

Возможен и иной вывод преобразования Лоренца, основывающийся на принципе постоянства скорости света и «естественных» допущениях об однородности и изотропности пространства и времени. Этот вывод предпочитают приводить тогда, когда стремятся доказать, что якобы основным содержанием теории относительности является предельность скорости распространения световых сигналов, используемых для синхронизации часов.

Начало этого вывода совпадает с предыдущим. Из «естественных» условий изотропности и однородности получаются линейные преобразования (7.11) с симметричными коэффициентами $k(v)$, $\lambda(v)$, $\mu(v)$ и $\eta(v)$. Далее формулируется требование постоянства скорости света в обеих системах отсчета.

Полагая, что фронт светового сигнала распространяется из начала координат, имеем для точек фронта волны в обеих системах отсчета

$$x^2 + y^2 + z^2 - c^2t^2 = 0, \tag{8.1}$$

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = 0. \tag{8.2}$$

Подставляя в (8.2) выражения (7.11), получаем

$$(x - vt)^2k^2 + \lambda^2y^2 + \lambda^2z^2 - c^2\mu^2\left(t - \frac{v}{\eta}\,x\right)^2 = 0. \tag{8.3}$$

Для того чтобы (8.3) не противоречило (8.1), необходимо потребовать:

$$k^2 = \frac{\mu^2c^2}{\eta}, \quad k^2 - \frac{\mu^2c^2v^2}{\eta^2} = \lambda^2,$$
$$(v^2k^2 - \mu^2c^2) = -\lambda^2c^2. \tag{8.4}$$

Исключая из этих уравнений λ и μ, получаем

$$1 - \frac{v^2}{\eta} = \frac{\eta}{c^2}\left(1 - \frac{v^2}{\eta}\right). \tag{8.5}$$

Отсюда $\eta = c^2$ или $\eta = v^2$. Второе решение согласно первому и последнему уравнению (8.4) приводит к $\lambda = 0$, что не имеет физического смысла. Поэтому необходимо положить

$$\eta = c^2. \tag{8.6}$$

Используя первое и последнее уравнения из (8.4), получаем

$$k^2 = \mu^2 = \frac{\lambda^2}{1 - v^2/c^2}\,. \tag{8.7}$$

Следовательно, **преобразования** (7.11) приобретают вид

$$x' = \lambda(v)\,\frac{x - vt}{\sqrt{1 - v^2/c^2}}\,,$$

$$y' = \lambda(v)y, \qquad z' = \lambda(v)z,$$

$$t' = \lambda(v)\,\frac{t - v/c^2\,x}{\sqrt{1 - v^2/c^2}}\,. \tag{8.8}$$

Эти преобразования отличаются от преобразований Лоренца (7.26) постоянным, масштабным множителем $\lambda(v)$. Если $\lambda(v) \neq 1$, то согласно (8.7) среди инерциальных систем существуют некоторые выделенные системы. Например, «неподвижная» система Σ выделена тем, что при переходе из нее в «движущуюся» систему Σ' координаты y и z *увеличиваются* в λ раз, в то время как при переходе из Σ' в Σ y' и z' должны *уменьшаться* в λ раз. Отсутствие выделенности системы Σ означает, очевидно,

$$\lambda(v)\lambda(-v) = 1, \text{ или } \lambda^2 = 1. \tag{8.9}$$

Но требование отсутствия выделенности системы фактически означает требование относительности. Таким образом, преобразования Лоренца в окончательном виде (т. е. при $\lambda = 1$) получаются лишь после применения требования, согласующегося с принципом относительности. В этом выводе, однако, принцип относительности играет как бы второстепенную роль, так как преобразования Лоренца почти в окончательном виде получаются без применения этого принципа.

Сопоставление приведенных нами двух выводов преобразований Лоренца подтверждает наше утверждение, что содержанием теории относительности является не принцип относительности и не предельность скорости световых сигналов, а наличие группы Лоренца, из которой и получаются все практически важные следствия теории относительности.

III

КИНЕМАТИЧЕСКИЕ ПАРАДОКСЫ

§ 9. Изображение преобразований Лоренца на плоскости Минковского

Первыми наиболее поражающими следствиями преобразований Лоренца являются сокращение движущихся масштабов в направлении движения и замедление хода движущихся часов. С точки зрения повседневных представлений о пространстве и времени эти следствия представляются парадоксальными.

Исчерпывающее, но всегда кажущееся несколько формальным разъяснение этих кинематических явлений дается на плоскости x, ct, если в соответствии с правилами четырехмерной геометрии Минковского изобразить на ней сетку координат «неподвижной» и «движущейся» систем.

Преобразования Лоренца оставляют инвариантным (неизменным) интервал s_{12} между любыми двумя событиями, определяемый согласно (3.1), как в этом легко убедиться подстановкой (7.26) в (3.2).

Совмещая первое событие с моментом $t = 0$ и началом отсчета системы Σ и вводя симметричные обозначения координат и времени

$$x_0 = ct,\ x_1 = x,\ x_2 = y,\ x_3 = z, \tag{9.1}$$

интервал между вторым и первым событием записывается в виде

$$x_0^2 - x_1^2 - x_2^2 - x_3^2 = s^2. \tag{9.2}$$

Четырехмерная геометрия, определяемая инвариантностью интервала (9.2), качественно отличается от обычной евклидовой геометрии, определяемой инвариантностью расстояния, т. е.

$$x_1^2 + x_2^2 + x_3^2 = R^2, \tag{9.3}$$

или от простого четырехмерного обобщения евклидовой геометрии, где инвариантом считается

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = \rho^2. \tag{9.4}$$

В евклидовых геометриях, определяемых (9.3) или (9.4), квадрат «расстояния» всегда положителен и, следовательно, «расстояние» является действительной величиной. Но в четырехмерной геометрии, определяемой интервалом (9.2), являющимся аналогом «расстояния», квадрат интервала может быть положительным, отрицательным или равным нулю. Соответственно в этой псевдоевклидовой геометрии интервал может быть действительной или мнимой величиной. В частном случае он может быть равен нулю для несовпадающих событий.

Иногда кажется, что качественное различие между четырехмерной евклидовой геометрией и четырехмерной псевдоевклидовой геометрией стирается, если, воспользовавшись предложением Минковского, считать время пропорциональным некоторой мнимой четвертой координате, т. е. положить

$$x_4 = ix_0 = ict. \tag{9.5}$$

В этом случае квадрат интервала запишется как

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = -s^2, \tag{9.6}$$

т. е. с точностью до знака совпадает с (9.4). Однако в силу мнимости x_4 выражение (9.6), так же как и (9.2), может иметь различные знаки и, таким образом, качественно отличается от выражения (9.4).

В силу инвариантности интервала качественное различие связи между событиями не зависит от выбора системы отсчета и действительный, или времениподобный, интервал ($s^2 > 0$) остается действительным во всех системах отсчета, мнимый же, или пространственноподобный, интервал ($s^2 < 0$) также остается мнимым во всех системах отсчета.

Все эти особенности псевдоевклидовой геометрии могут быть наглядно проиллюстрированы на плоскости Минковского x_1, x_0 (рис. 5).

Отрезками Oa и Ob на этой плоскости изображены соответственно единичные масштабы временной оси x_0 и пространственной оси x_1. Кривая, выходящая вправо из точки a, является гиперболой, описываемой уравнением

$$x_0^2 - x_1^2 = +1, \quad (9.7)$$

а кривая, выходящая вверх из точки b, является гиперболой, описываемой уравнением

$$x_0^2 - x_1^2 = -1. \quad (9.8)$$

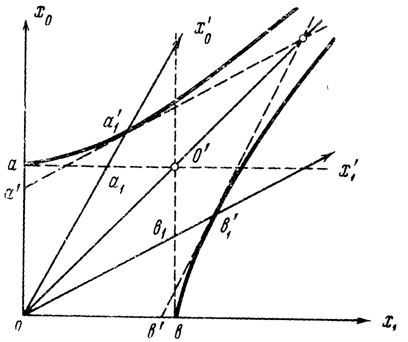

Рис. 5

Таким образом, точка начала координат и все точки, лежащие на гиперболе, выходящей из точки a, разделены единичным времениподобным интервалом. Точки же, лежащие на гиперболе, выходящей из точки b, отделены от начала координат пространственноподобным интервалом. Штриховая линия, выходящая параллельно оси x_1 из точки a, изображает точки с координатами $x_0 = 1$, а линия, выходящая из точки b параллельно оси x_0, изображает точки с координатами $x_1 = 1$.

На этой же плоскости нанесены линии Oa_1' и Ob_1', изображающие соответственно точки с координатами $x_1' = 0$ и $x_0' = 0$, а также линии, проходящие через a', a'_1 и b', b_1' и соответственно изображающие точки с координатами $x_0' = 1$ и $x_1' = 1$. Эти линии изображают координатную сетку системы Σ'.

Из рисунка видно, что переход от системы Σ к системе Σ' соответствует переходу от прямоугольных координат к косоугольным на плоскости Минковского. Последнее следует также непосредственно из преобразований Лоренца (7.26), которые можно записать также в виде

$$x_1' = \frac{x_1 - \beta x_0}{\sqrt{1-\beta^2}}, \quad x_0' = \frac{x_0 - \beta x_1}{\sqrt{1-\beta^2}}, \quad (9.9)$$

где $\beta = v/c$, или в виде

$$x_1' = x_1 \operatorname{ch}\psi - x_0 \operatorname{sh}\psi,$$
$$x_0' = -x_1 \operatorname{sh}\psi + x_0 \operatorname{ch}\psi, \qquad (9.10)$$

где

$$\operatorname{ch}\psi = \frac{1}{\sqrt{1-\beta^2}}, \quad \operatorname{sh}\psi = \frac{\beta}{\sqrt{1-\beta^2}},$$

и очевидно

$$\operatorname{ch}^2\psi - \operatorname{sh}^2\psi = 1.$$

Но преобразования (9.10) тождественны преобразованиям перехода от декартовых координат к косоугольным. При этих преобразованиях времениподобные векторы, т. е. векторы, направленные из начала отсчета в точки, лежащие выше линии OO', в любой системе кординат также останутся времениподобными, так как концы векторов лежат на гиперболах. Следовательно, и пространственноподобные векторы во всех системах координат останутся пространственноподобными.

Из рис. 5 видно, что «пространственная» проекция единичного вектора Ob_1' на ось Ox_1' равна единице, а на ось Ox_1' равна Ob', т. е. меньше единицы. Следовательно, масштаб, покоящийся в системе Σ', при измерении из системы Σ окажется укороченным. Но это утверждение обратимо, ибо «пространственная» проекция вектора Ob на ось Ox_1' равна Ob_1, т. е. в системе Σ' меньше, чем Ob_1', являющийся единичным вектором.

Аналогично дело обстоит и с «временными» проекциями на оси x_0 и x_0'. Отрезок Oa_1', изображающий в системе Σ' процесс, длящийся единицу времени, в системе Σ будет проектироваться как Oa', т. е. как процесс, длящийся меньшее время, чем $Oa = 1$. Следовательно, ход часов, покоящихся в системе Σ', при измерении из системы Σ окажется замедленным. Легко проверить, что это явление также обратимо, т. е. ход часов, покоящихся в системе Σ, оказывается замедленным в системе Σ'.

Рассмотрим эти кинематические явления сокращения пространственных и временных четырехмерных проекций более детально, не обращаясь к рис. 5, а опираясь только на преобразования Лоренца и привычные представления о пространстве, измеряемом масштабами, и времени, измеряемом часами.

§ 10. Сокращение движущихся масштабов

Если длина неподвижного масштаба может быть измерена путем прикладывания к нему эталонных масштабов, без использования каких-либо часов, то длину движущегося масштаба невозможно измерить из неподвижной системы отсчета без использования часов или сигналов, отмечающих одновременность прохождения концов измеряемого масштаба относительно точек эталона. Таким образом, под длиной движущегося масштаба надо понимать расстояние между его концами,

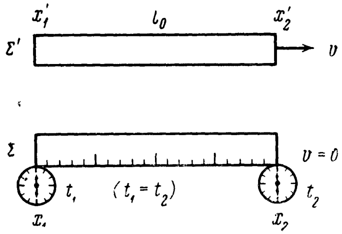

Рис. 6

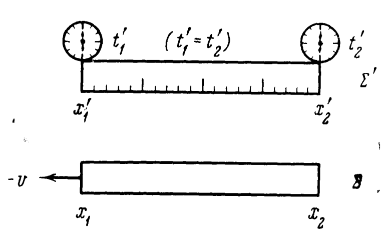

Рис. 7

измеренное при помощи неподвижного эталона в один и тот же момент времени для каждого конца. **Одновременность** измерения положений концов является существенно необходимым условием опыта. Легко видеть, что нарушение этого условия может привести к тому, что измеренная длина может оказаться любой, в том числе отрицательной или равной нулю.

Схема опыта по измерению длины движущегося масштаба приведена на рис. 6.

Пусть $l_0 = x_2' - x_1' = l'$ — длина движущегося масштаба, предварительно измеренная путем непосредственного приложения к эталону, помещавшемуся в любой системе координат *. Тогда, если моменты t_1 и t_2 прохождения концов масштаба мимо точек x_1 и x_2 неподвижного эталона одинаковы (т. е. $t_1 = t_2$), то $x_2 - x_1 = l$ является, по определению, длиной движущегося мас-

* Таким образом, мы считаем, что измерение длины путем приложения к эталону является операцией, дающей один и тот же результат независимо от дальнейших перемещений масштаба из одной системы отсчета в другую. При этом масштаб, конечно, считается «жестким».

штаба. Согласно (7.26) имеем

$$l' = x_2' - x_1' = \frac{(x_2 - x_1) - v\,(t_2 - t_1)}{\sqrt{1 - v^2/c^2}}, \tag{10.1}$$

откуда в силу $t_1 = t_2$ получаем

$$l = l_0 \sqrt{1 - v^2/c^2}, \quad (l' = l_0). \tag{10.2}$$

Парадоксальность этого вывода состоит в том, что в силу принципа относительности точно такая же формула должна получиться для длины масштаба, находящегося в системе Σ и измеряемого из системы Σ'. Иначе говоря, представляется необходимым удовлетворение обратного соотношения

$$l_0 = l \sqrt{1 - v^2/c^2}, \tag{10.3}$$

которое находится в явном противоречии с (10.2), если под l_0 и l понимать также измеряемые величины.

Противоречие, однако, снимается, если учесть, что относительность предполагает совершенно симметричное изменение всей схемы измерения, т. е. переход от схемы, изображенной на рис. 6, к схеме, изображенной на рис. 7.

В этой схеме уже $t_1 \neq t_2$, но $t_1' = t_2'$, т. е. концы нижнего масштаба засекаются не в один и тот же момент времени по часам, помещенным в системе Σ, но в один и тот же момент по часам, находящимся в системе Σ'. Тогда, применяя формулы обратных преобразований Лоренца, получаем

$$x_2 - x_1 = \frac{(x_2' - x_1') + v\,(t_2' - t_1')}{\sqrt{1 - v^2/c^2}}, \tag{10.4}$$

откуда в силу $t_2' = t_1'$ имеем

$$l' = l \sqrt{1 - v^2/c^2} = l_0. \tag{10.5}$$

Эта формула действительно означает, что уменьшается длина масштаба l, измеренного из системы Σ'. Но эта формула уже не находится в противоречии с формулой (10.2), ибо входящие в нее l и l_0 измеряются иначе, чем l и l_0, входящие в (10.2).

Следовательно, укорочение или удлинение измеряемых масштабов зависит лишь от того, в какой системе отсчета производятся одновременные измерения положений концов масштабов, ибо события, одновременные в одной системе отсчета, неодновременны в другой.

§ 11. Замедление движущихся часов

Замедление движущихся часов может быть обнаружено в опыте, схема которого изображена на рис. 8. Движущиеся со скоростью v часы, измеряющие время t', проходят последовательно мимо точки x_1 в момент t_1

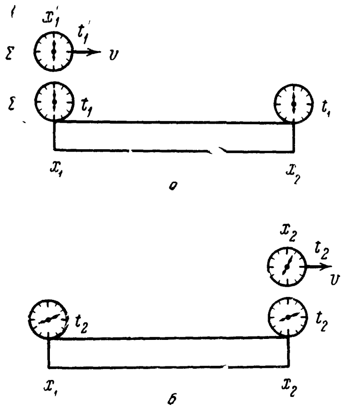

Рис. 8

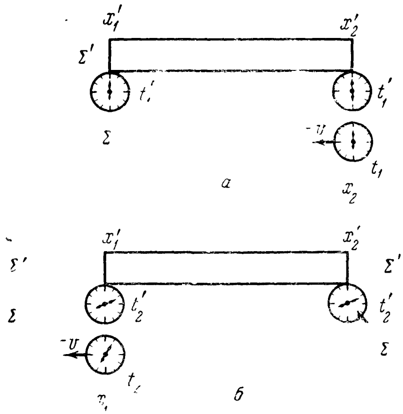

Рис. 9

и мимо точки x_2 в момент t_2. В эти моменты производится сравнение положений стрелок движущихся часов и соответствующих неподвижных, находящихся рядом с ними *.

Пусть за время движения от точки x_1 до точки x_2 стрелки движущихся часов отмерят промежуток времени τ_0, а стрелки часов *1* и *2*, предварительно синхронизованных в неподвижной системе Σ, отмерят промежуток времени τ. Таким образом,

$$\tau' = \tau_0 = t'_2 - t'_1, \qquad \tau = t_2 - t_1. \tag{11.1}$$

Но согласно обратным преобразованиям Лоренца (7.27) имеем

$$t_2 - t_1 = \frac{(t'_2 - t'_1) + \frac{v}{c^2}(x'_2 - x'_1)}{\sqrt{1 - v^2/c^2}}, \tag{11.2}$$

---

* Предполагается, что одновременность одноместных событий устанавливается непосредственно, без каких-либо сигналов или иных особых процессов. Одновременность же часов, находящихся

Подставляя (11.1) в (11.2) и замечая, что движущиеся часы все время находятся в одной и той же точке движущейся системы отсчета Σ', т. е. что

$$x_1' = x_2', \tag{11.3}$$

получаем

$$\tau = \frac{\tau_0}{\sqrt{1 - v^2/c^2}}, \quad (\tau_0 = \tau'). \tag{11.4}$$

Эта формула означает, что промежуток времени, отмеченный неподвижными часами, оказывается большим, чем промежуток времени, отмеренный движущимися часами. Но это и означает, что движущиеся часы отстают от неподвижных, т. е. их ход замедляется.

Формула (11.4) так же обратима, как и соответствующая формула для масштабов (10.2). Однако, написав обратную формулу в виде

$$\tau_0 = \frac{\tau}{\sqrt{1 - v^2/c^2}}, \tag{11.5}$$

мы должны подразумевать, что $\tau_0 = t_2' - t_1'$ и $\tau = t_2 - t_1$ измеряются уже не в опыте, изображенном на рис. 8, а в опыте, изображенном на рис. 9. В этом случае, действительно, согласно преобразованиям Лоренца

$$t_2' - t_1' = \frac{(t_2 - t_1) - \frac{v}{c^2}(x_2 - x_1)}{\sqrt{1 - v^2/c^2}} \tag{11.6}$$

при условии

$$x_2 = x_1 \tag{11.7}$$

получаем формулу (11.5).

Полученное замедление является вполне реальным *, однако оно имеет, так сказать, чисто кинематическую природу. Например, в схеме опыта, изображенной на

в точках x_1 и x_2 системы Σ, предполагается предварительно установленной посредством световых сигналов или процедуры, рассмотренной в § 6.

* Известно, что быстро движущиеся радиоактивные частицы, например μ-мезоны, распадаются с бо́льшим полупериодом, чем неподвижные.

рис. 8, тот результат, что часы *2* оказались впереди движущихся часов, с точки зрения движущейся системы Σ' объясняется тем, что часы *2* с самого начала шли не синхронно с часами *1* и опережали их (в силу неодновременности разобщенных событий, одновременных в другой движущейся системе отсчета).

Таким образом, как замедление движущихся часов, так и сокращение движущихся масштабов не являются парадоксальными, если освоиться с представлением об относительности одновременности пространственно разобщенных событий.

§ 12. Парадокс часов

Более поразительным и вызывающим большое число споров и недоразумений является так называемый парадокс часов. Пусть часы A находятся в точке I в неподвижной инерциальной системе отсчета Σ, а одинаковые

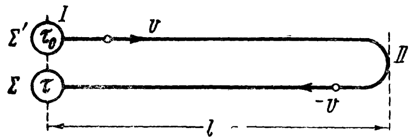

Рис. 10

с ними часы B, находившиеся в начальный момент также в точке I, движутся к точке II со скоростью v. Затем, пройдя путь l до точки II, часы B замедляются и, приобретая противоположную скорость $-v$. возвращаются в точку I (рис. 10).

Если время, требуемое на изменение скорости часов B на обратную, достаточно мало по сравнению с временем прямолинейного и равномерного движения от точки I до точки II, то время τ, отмеренное часами A, и время τ_0, отмеренное часами B, можно вычислить согласно (11.4) по формулам

$$\tau = 2l / v, \ \tau_0 = \tau \sqrt{1 - v^2 / c^2} + \delta, \tag{12.1}$$

где δ — возможная малая поправка на время ускоренного движения часов B. Следовательно, часы B, вернувшись в точку I, реально отстанут от часов A на время

$$\Delta\tau = \tau - \tau_0 = 2\frac{l}{v}(1 - \sqrt{1 - v^2/c^2}) + \delta. \qquad (12.2)$$

Поскольку расстояние l может быть сделано сколь угодно большим, постольку поправка δ может не приниматься во внимание вообще.

Особенность этого кинематического следствия преобразований Лоренца состоит в том, что здесь отставание хода движущихся часов является вполне реальным эффектом, а не результатом избранной процедуры измерения, как это имело место в предыдущем параграфе. Реально должны отстать все процессы, связанные с системой Σ', от процессов, идущих в системе Σ. В том числе должны отстать и биологические процессы организмов, находящихся вместе с часами B. Должны замедлиться физиологические процессы в организме человека, путешествующего в системе Σ', в результате чего организм, находившийся в системе Σ', в момент ее возврата в точку I окажется менее постаревшим, чем организм, остававшийся в системе Σ.

Парадоксальным представляется здесь то, что одни из часов реально отстают от других. Ведь это кажется противоречащим самому принципу относительности, так как согласно последнему любую из систем Σ и Σ' можно считать неподвижной. Но тогда представляется, что лишь в зависимости от нашего выбора реально отстающими могут стать любые из часов A и B. Но последнее явно абсурдно, так как реально отстают часы B от часов A.

Ошибочность последнего рассуждения состоит в том, что системы Σ и Σ' физически не равноправны, так как система Σ все время инерциальна, система же Σ' некоторый промежуток времени, когда производится изменение ее скорости на обратную, неинерциальна. Следовательно, вторая из формул (12.1) для системы Σ' неправильна, так как во время ускорения ход удаленных часов может сильно изменяться за счет инерциального гравитационного поля.

Однако и это совершенно правильное объяснение представляется весьма поразительным. Ведь в течение большого промежутка времени обе системы движутся друг относительно друга прямолинейно и равномерно. Поэтому с точки зрения системы Σ' часы A, находящиеся в Σ, отстают (а не уходят вперед) в полном соответствии с формулой (12.1). И лишь за малый промежуток времени, когда в системе Σ' действуют инерциальные силы, часы A быстро уходят вперед на промежуток времени, вдвое больший, чем $\Delta\tau$, вычисляемый по формуле (12.2). При этом чем большее ускорение испытывает система Σ', тем быстрее бежит время на часах A.

Наглядно суть полученных выводов может быть разъяснена на плоскости Минковского (рис. 11).

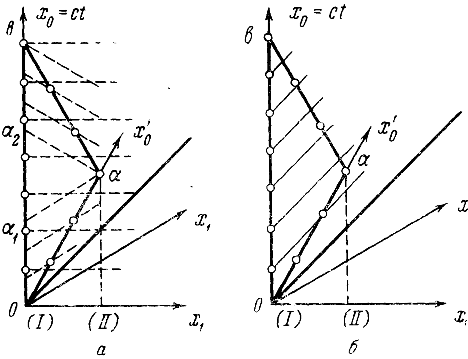

Рис. 11

Отрезок Ob на рис. 11, a изображает покоящиеся часы A, ломаная линия Oab — движущиеся часы B. В точке a действуют силы, ускоряющие систему часов B и изменяющие ее скорость на обратную. Точки, расставленные на оси Ob, разделяют единичные промежутки времени в неподвижной системе Σ, связанной с часами A.

Точки на ломаной Oab отмечают равные единичные промежутки времени, измеряемые часами B, находящимися в системе Σ'. Из рисунка видно, что число единичных отрезков, укладывающихся на линии Ob, больше, чем число таких же, но относящихся к системе Σ' отрезков, укладывающихся на ломаной Oab. Следовательно, часы B отстают от часов A.

Согласно рисунку «неподвижные» часы A также отстают от часов B вплоть до момента, изображаемого точкой a. Одновременным с этим моментом является момент α_1, однако до тех пор, пока часы B еще движутся со скоростью v. Но через малый промежуток времени, требуемый для замедления часов B и сообщения им скорости $-v$, на часах B практически останется тот же момент A, но одновременным с ним моментом в системе Σ станет момент α_2, т. е. почти мгновенно время системы Σ как бы перескочит на конечный интервал $\alpha_1\alpha_2$.

Этот перескок времени не является, однако, реально наблюдаемым эффектом. Действительно, если из системы Σ регулярно, через единичные интервалы посылать в систему Σ' световые сигналы, то они совершенно регулярно будут приниматься системой Σ сперва более редко, а затем, после изменения скорости на обратную, более часто. Никакого разрыва в показаниях часов A в системе Σ' наблюдаться не будет, как это видно из рис. 11 *б*,

Таким образом, «парадокс часов» также является лишь непривычным для обычных представлений о пространстве и времени следствием псевдоевклидовой геометрии четырехмерного пространственно-временного многообразия.

§ 13. Четырехмерная и трехмерная скорости

Прямолинейно и равномерно движущиеся материальные точки в четырехмерном пространстве Минковского, очевидно, изображаются прямыми линиями. Согласно терминологии Минковского линии, изображающие движение точек, называются **мировыми линиями**. Следовательно, точки, движущиеся неравномерно и непрямолинейно, изобразятся кривыми линиями Минковского. В каждой данной точке такое движение характеризуется направлением касательной к мировой линии. Направление касательной можно определить как четырехмерный вектор с компонентами

$$dx_k/ds \quad (k = 0, 1, 2, 3), \tag{13.1}$$

где

$$(ds)^2 = (dx_0)^2 - (dx_1)^2 - (dx_2)^2 - (dx_3)^2 \tag{13.2}$$

— квадрат элемента интервала, являющегося четырехмерным инвариантом. В соответствии с терминологией Минковского этот вектор можно назвать вектором **мировой касательной**.

Компоненты вектора мировой касательной можно изобразить также в виде

$$\frac{dx_k}{ds} = \frac{dx_k}{dx_0} \frac{1}{\sqrt{1-(dx_1/dx_0)^2-(dx_2/dx_0)^2-(dx_3/dx_0)^2}} =$$

$$= \frac{1}{c} \frac{dx_k/dt}{\sqrt{1-u^2/c^2}} . \qquad (13.3)$$

Иначе говоря, этот вектор имеет четыре компоненты, соответственно равные

$$\frac{1}{\sqrt{1-u^2/c^2}}, \quad \frac{u_x/c}{\sqrt{1-u^2/c^2}}, \quad \frac{u_y/c}{\sqrt{1-u^2/c^2}}, \quad \frac{u_z/c}{\sqrt{1-u^2/c^2}} . \qquad (13.4)$$

При $u \ll c$ нулевая компонента этого вектора превратится в единицу, а первая, вторая и третья компоненты образуют трехмерный вектор $\mathbf{u}/c$. Следовательно, умноженный на c четырехмерный вектор (13.3) может рассматриваться как четырехмерное обобщение обычной трехмерной скорости $\mathbf{u}$, т. е.

$$U_k = c \frac{dx_k}{ds} = \frac{dx_k}{dt\sqrt{1-u^2/c^2}} = \frac{dx_k}{d\tau} , \qquad (13.5)$$

где

$$dt\sqrt{1-u^2/c^2} = d\tau \qquad (13.6)$$

— элемент «собственного времени», отсчитываемого согласно (11.4) по часам, движущимся вместе с материальной точкой.

Четырехмерная длина вектора U_k, очевидно, равна * согласно (13.2) и (13.5):

$$U_k U_k = U_0^2 - U_1^2 - U_2^2 - U_3^2 = c^2. \qquad (13.7)$$

В четырехмерной псевдоевклидовой геометрии только вектор (13.5) может считаться четырехмерной скоростью, т. е. ковариантно введенной физической

---

* Здесь и далее применяется условная запись суммирования по индексу, встречающемуся дважды; например, скалярное произведение двух векторов записывается в виде

$$A_k B_k = \sum_{k=0}^{3} \varepsilon_k A_k B_k = A_0 B_0 - A_1 B_1 - A_2 B_2 - A_3 B_3,$$

где

$$\varepsilon_0 = 1, \ \varepsilon_1 = -1, \ \varepsilon_2 = -1, \ \varepsilon_3 = -1.$$

величиной, являющейся обобщением трехмерной скорости. Обычная же скорость $\mathbf{u}$ не является такой ковариантной физической величиной, ибо не составляет величины, преобразующейся как четырехмерный вектор при преобразованиях Лоренца.

Только при малых скоростях ($u \ll c$), когда преобразования Лоренца в пределе переходят в преобразования Галилея, вектор $\mathbf{u}$ становится физически корректно введенной величиной, ковариантной по отношению к преобразованиям Галилея и поворотам осей координат в пространстве.

Компоненты вектора U_k, так же как и любого другого четырехмерного вектора, преобразуются при переходе к другой системе отсчета, движущейся со скоростью v, так же как компоненты dx_k^*, т. е. согласно преобразованиям Лоренца [см. (9.9)]:

$$U_1' = \frac{U_1 - \beta U_0}{\sqrt{1-\beta^2}}, \quad U_2' = U_2, \quad U_3' = U_3, \quad U_0' = \frac{U_0 - \beta U_1}{\sqrt{1-\beta^2}}, \tag{13.8}$$

где $\beta = v/c$. При этом «длина» вектора скорости остается постоянной и равной ic.

Аналогично введению четырехмерной скорости можно ввести и четырехмерное ускорение с компонентами

$$\frac{d^2x_k}{d\tau^2} = \frac{dU_k}{d\tau}. \tag{13.9}$$

Легко убедиться в том, что при $v \ll c$ первые три компоненты четырехмерного вектора ускорения совпадают с компонентами трехмерного ускорения.

Согласно (13.7)

$$U_k \frac{dU_k}{d\tau} = 0, \tag{13.10}$$

т. е. вектор четырехмерного ускорения всегда перпендикулярен вектору четырехмерной скорости. Компоненты вектора четырехмерного ускорения также преобразуются по формулам (13.8).

Для произвольного вектора A_k эти формулы можно

---

* Различие преобразований ковариантных и контравариантных векторов в рассматриваемых случаях отсутствует, поэтому мы здесь не будем его вводить.

записать в виде

$$A_k' = a_{k\alpha} A_\alpha, \tag{13.11}$$

где матрица $\|a_{ik}\|$ записывается как

$$\|a_{ik}\| = \begin{Vmatrix} \frac{1}{\sqrt{1-\beta^2}} & \frac{-\beta}{\sqrt{1-\beta^2}} & 0 & 0 \\ \frac{-\beta}{\sqrt{1-\beta^2}} & \frac{1}{\sqrt{1-\beta^2}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{Vmatrix}. \tag{13.12}$$

Преобразование вектора четырехмерной скорости и других четырехмерных векторов по формулам (13.8), (13.12) не приводит к каким-либо парадоксальным выводам, более непривычным, чем те следствия преобразований Лоренца, которые нами были получены для преобразований координат и времени. Некоторые новые неожиданности появляются лишь тогда, когда от четырехмерных векторов мы возвращаемся к трехмерным векторам, например к трехмерному вектору скорости $\mathbf{u}$.

Любой четырехмерный вектор A_k можно изобразить как состоящий из трехмерной части $\mathbf{A}$ с компонентами A_1, A_2, A_3 и нулевой компоненты A_0, т. е. как

$$(A_0, \mathbf{A}) = (A_0, A_1, A_2, A_3). \tag{13.13}$$

Согласно этому обозначению и формулам (13.4), (13.5) четырехмерный вектор скорости можно записать как

$$(U_0, \mathbf{U}) = \left(\frac{c}{\sqrt{1 - u^2/c^2}}, \frac{\mathbf{u}}{\sqrt{1 - u^2/c^2}} \right). \tag{13.14}$$

Следовательно, обычный вектор трехмерной скорости $\mathbf{u}$, измеряемый известными в классической физике операциями, выражается через компоненты четырехмерного вектора скорости как

$$\mathbf{u} = c \frac{\mathbf{U}}{U_0}. \tag{13.15}$$

Из этой формулы и преобразований (13.8) легко получить формулы для компонент вектора $\mathbf{u}$ в движущейся системе отсчета Σ'. Очевидно

$$u'_x = c\frac{U'_1}{U'_0} = c\frac{U_1 - v/cU_0}{U_0 - v/cU_1} = \frac{u_x - v}{1 - vu_x/c^2},$$

$$u'_y = c\frac{U'_2}{U'_0} = c\frac{U_2\sqrt{1 - v^2/c^2}}{U_0 - v/cU_1} = \frac{u_y\sqrt{1 - v^2/c^2}}{1 - vu_x/c^2}, \qquad (13.16)$$

$$u'_z = c\frac{U'_3}{U'_0} = c\frac{U_3\sqrt{1 - v^2/c^2}}{U_0 - v/cU_1} = \frac{u_z\sqrt{1 - v^2/c^2}}{1 - vu_x/c^2}.$$

Таким образом, мы получили так называемые ф о р м у л ы   с л о ж е н и я   с к о р о с т е й. Нетрудно доказать, что эти формулы обратимы, т. е.

$$u_x = \frac{u'_x + v}{1 + vu'_x/c^2},$$

$$u_y = \frac{u'_y\sqrt{1 - v^2/c^2}}{1 + vu'_x/c^2},$$

$$u_z = \frac{u'_z\sqrt{1 - v^2/c^2}}{1 + vu'_x/c^2}. \qquad (13.17)$$

Из этих формул, в частности, видно, что путем перехода от одной инерциальной системы отсчета к другой невозможно получить скорость тела, превышающую скорость света c, если в исходной системе скорость этого тела была меньше c.

Действительно, согласно формулам (13.17) квадрат трехмерной скорости $\mathbf{u}$ преобразуется как

$$\mathbf{u}'^2 = u'^2_x + u'^2_y + u'^2_z = \frac{(u_x - v)^2 + (u_y^2 + u_z^2)(1 - v^2/c^2)}{(1 - vu_x/c^2)^2}. \qquad (13.18)$$

Вычитая из правой и левой частей этого уравнения c^2, после простых преобразований получаем

$$u'^2_x + u'^2_y + u'^2_z - c^2 = (u^2_x + u^2_y + u^2_z - c^2)\frac{(1 - v^2/c^2)}{(1 - vu_x/c^2)^2}. \qquad (13.19)$$

Преобразования Лоренца, очевидно, имеют физический смысл только при $v < c$, так как в противном случае (т. е. при $v > c$) штрихованные координаты становятся мнимыми. Следовательно, физический смысл имеют лишь системы отсчета, движущиеся относительно исходной со скоростями, не превышающими скорость света *. Поэтому множитель, стоящий в правой части (13.19), всегда положителен, т. е.

$$\frac{(1 - v^2/c^2)}{(1 - vu_x/c^2)^2} > 0, \tag{13.20}$$

откуда согласно (13.19) получаем, что выражение

$$\mathbf{u}^2 - c^2 \tag{13.21}$$

не изменяет знака при любых переходах из одной инерциальной системы отсчета в другую. Следовательно, для любых систем отсчета

$$\begin{aligned} &\mathbf{u}'^2 < c^2, \text{ если } \mathbf{u}^2 < c^2, \\ &\mathbf{u}'^2 = 0, \text{ если } \mathbf{u}^2 = 0, \\ &\mathbf{u}'^2 > c^2, \text{ если } \mathbf{u}^2 > c^2. \end{aligned} \tag{13.22}$$

Иначе говоря, досветовые скорости ($u < c$) любых физических процессов во всех инерциальных системах отсчета остаются досветовыми, сверхсветовые ($u > c$) всегда остаются сверхсветовыми, а процессы, распространяющиеся со скоростью света в любых системах отсчета, имеют скорость, равную скорости света, в соответствии с постулатом о постоянстве скорости света в пустоте.

---

* Подчеркнем, что речь идет здесь о системах отсчета, а не о физических процессах, для которых не исключены и скорости, превышающие c.

IV

ПАРАДОКСЫ РЕЛЯТИВИСТСКОЙ ДИНАМИКИ

§ 14. Механика материальной точки

Релятивистски ковариантными уравнениями механики материальной точки являются уравнения Минковского, которые представляют собой естественное обобщение уравнений Ньютона, заключающееся в замене трехмерных скоростей и ускорений на четырехмерные. Такая замена не противоречит надежно проверенной механике Ньютона, так как при тех скоростях, при которых последняя применима, первые три компоненты четырехмерной скорости и ускорения не отличаются от компонент трехмерной скорости и ускорения.

Уравнения Минковского имеют вид

$$M \frac{d^2 x_k}{d\tau^2} = F_k, \tag{14.1}$$

или

$$M \frac{dU_k}{d\tau} = F_k. \tag{14.2}$$

Здесь M — собственная масса, являющаяся четырехмерным инвариантом и представляющая естественное обобщение обычной ньютоновской массы; F_k — компонента четырехмерной силы, представляющей обобщение трехмерной силы $\mathfrak{f}$.

Вводя четырехмерный вектор импульса

$$P_k = MU_k, \tag{14.3}$$

уравнения Минковского можно также записать в виде

$$dP_k/d\tau = F_k. \tag{14.4}$$

Для того чтобы раскрыть физический смысл нулевой компоненты уравнений Минковского и нулевой компоненты четырехмерного импульса, а также для того чтобы установить связь с обычными трехмерными физическими измерениями, уравнения Минковского записываются также в нековариантной трехмерной форме.

Учитывая, что в силу перпендикулярности четырехмерной скорости и четырехмерного ускорения, т. е. в силу условия (13.10) согласно (14.2) получаем

$$F_k U_k = 0, \tag{14.5}$$

или в обозначениях (13.14) имеем

$$F_0 U_0 - \mathbf{FU} = 0. \tag{14.6}$$

Вводя далее обозначения

$$m = \frac{M}{\sqrt{1 - u^2/c^2}}; \tag{14.7}$$

$$\mathbf{f} = \sqrt{1 - u^2/c^2}\,\mathbf{F}, \tag{14.8}$$

получаем согласно (14.6) вместо (14.2) следующие уравнения:

$$\frac{d}{dt}(m\mathbf{u}) = \mathbf{f}, \quad \frac{d}{dt}(mc^2) = (\mathbf{uf}), \tag{14.9}$$

первое из которых при $u \ll c$ переходит в обычное трехмерное векторное уравнение Ньютона

$$\frac{d}{dt}(M\mathbf{u}) = M\frac{d\mathbf{u}}{dt} = \mathbf{f}, \tag{14.10}$$

а второе превращается в уравнение энергии

$$\frac{d}{dt}\left(\frac{Mu^2}{2}\right) = (\mathbf{uf}), \tag{14.11}$$

в силу того, что при $u \ll c$

$$mc^2 = \frac{Mc^2}{\sqrt{1 - u^2/c^2}} = Mc^2 + Mu^2/2 + \cdots \tag{14.12}$$

Таким образом, выясняется, что первые три компоненты четырехмерного импульса являются обобщением трехмерного импульса, а нулевая компонента имеет смысл обобщенной энергии, деленной на скорость света. Следовательно,

$$\mathbf{P} = M\mathbf{U} = m\mathbf{u} = \frac{M\mathbf{u}}{\sqrt{1 - u^2/c^2}} = \frac{\mathbf{u}E}{c^2}, \qquad (14.13)$$
$$P_0 = MU_0 = mc = \frac{Mc}{\sqrt{1 - u^2/c^2}} = E/c.$$

В релятивистской четырехмерной механике величину $\mathbf{P}$ называют импульсом, а величину E — энергией точки. Величину m, определяемую по (14.7), называют инертной массой, поскольку она входит в уравнения движения (14.9), так же, как входит инертная масса в уравнения движения Ньютона. Чтобы отметить зависимость m от скорости, ее называют также массой движения.

Согласно (14.13) энергия выражается в виде

$$E = mc^2 \qquad (14.14)$$

или

$$E = \frac{Mc^2}{\sqrt{1 - u^2/c^2}}. \qquad (14.15)$$

Так получается известная формула Эйнштейна, которая иногда истолковывается как выражение эквивалентности массы и энергии или как «закон инертности энергии».

С точки зрения четырехмерной геометрии пространства-времени истинный физический смысл могут иметь лишь четырехмерно-ковариантные величины. В механике точки такими величинами являются четырехмерный скаляр — собственная масса M и четырехмерные вектора скорости U_k, ускорения $dU_k/d\tau$, импульса P_k. Энергия E в четырехмерной теории не имеет самостоятельного физического смысла, как это имеет место в трехмерной ньютоновской теории. В четырехмерной теории энергия E является умноженной на c нулевой компонентой четырехмерного импульса P_k. Следовательно, рассматриваемая только вместе с тремя первыми компонентами импульса $\mathbf{P}$, энергия $E = cP_0$ имеет ковариантный физический смысл. С этой четырехмерной точки зрения собственная масса является лишь величиной, пропорциональной «абсолютной длине» четырехмерного импульса, так как согласно (14.3) и (13.7)

$$P_k P_k = M^2 U_k U_k = M^2 c^2 \tag{14.16}$$

или

$$(E/c)^2 - \mathbf{P}^2 = M^2 c^2. \tag{14.17}$$

Уравнение (14.17) можно рассматривать как определение ковариантной собственной массы тела посредством четырехмерного вектора импульса.

Итак, с четырехмерно-геометрической точки зрения импульс $\mathbf{P}$ и энергия E деления на скорость света c, т. е. $P_0 = E/c$, составляют компоненты одной физической величины — четырехмерного вектора импульса. Инвариантная длина этого четырехмерного вектора равна собственной массе частицы M, умноженной на скорость света c.

§ 15. Смысл утверждения об эквивалентности массы и энергии

Соотношению Эйнштейна (14.14) обычно придается глубокий физический смысл. Оно трактуется как выражающее эквивалентность массы и энергии. Такая интерпретация формулы Эйнштейна представляется неизбежной, если величину m, определяемую формулой (14.7), рассматривать как релятивистское обобщение понятия инертной массы. В этом случае формула (14.14) выражает пропорциональность релятивистской энергии релятивистской инертной массе. Поскольку же соотношение (14.14) является универсальным, постольку во всех законах E можно заменить на m, и наоборот, что и выражает их эквивалентность с точностью до постоянного множителя c^2, который соответствующим выбором единиц измерения может быть сделан равным единице.

Если соотношению (14.14) придавать указанный выше смысл, то необходимо рассматривать его как релятивистски ковариантное. Однако с четырехмерной точки зрения величина E является нулевой компонентой четырехмерного вектора

$$E_k = c P_k. \tag{15.1}$$

Поэтому и величина m должна рассматриваться не как инвариант, а как нулевая компонента четырехмерного вектора

$$m_k = P_k/c. \tag{15.2}$$

Следовательно, и соотношение (14.14), если его рассматривать не как определение энергии через величины, входящие в уравнение (14.9), а как некоторое новое физическое утверждение, приобретает смысл только в том случае, если его записать в виде

$$E_k = m_k c^2, \tag{15.3}$$

приписывая при этом самостоятельный физический смысл векторам E_k и m_k.

Очевидно, и выражение (14.7) для инертной массы m имеет ковариантный смысл только в том случае, если его рассматривать как нулевую компоненту выражения четырехмерной величины m_k через M и U_k, т. е. как нулевую компоненту вектора

$$m_k = MU_k/c. \tag{15.4}$$

Тем более не имеет четырехмерного ковариантного физического смысла такое понятие как *кинетическая энергия*, определяемая выражением

$$(m - M)\, c^2 = Mc^2 \left(\frac{1}{\sqrt{1 - u^2/c^2}} - 1\right), \tag{15.5}$$

которое при $u \ll c$ совпадает с классическим выражением $Mu^2/2$. Первый член (15.5) является нулевой компонентой четырехмерного вектора, а второй член — четырехмерным скаляром. Ясно, что такой «гибрид» из компоненты вектора и скаляра не может считаться ковариантно вводимой физической величиной, если мы имеем дело со скоростями тел, сравнимыми со скоростью света. Только при скоростях $u \ll c$ кинетическая энергия приобретает смысл трехмерного скаляра.

Величины, определяемые выражениями (14.7), (14.14), (14.15), (15.5), и уравнения (14.9) имеют определенный физический смысл только при переходе от четырехмерных представлений к трехмерным, связанным с фиксированной системой отсчета, т. е. при игнорировании принципа относительности и четырехмерности пространства-времени. Последнее же может быть оправдано либо для случая $u \ll c$, либо как некоторый суррогат строгой теории, предназначенный для примирения четырехмерной теории с наглядными трехмерными представлениями привычной классической физики. Действительно, если игнорировать релятивистскую ковариантность и строить теорию для одной единственной системы отсчета, то релятивистские эффекты можно изобразить как учет того обстоятельства, что входящая в обычные уравнения механики (14.9) «инертная масса» m зависит от скорости согласно «закону» (14.7). Тогда определение энергии E как нулевой компоненты четырехмерного импульса, умноженной на c, можно представить как «закон инертности энергии» и т. д.

Многие недоразумения и парадоксы, возникающие при истолковании формул релятивистской механики, происходят вследствие того, что так называемые законы, имеющие оправдание только в трехмерной нековариантной формулировке, пытаются понять и истолковать с релятивистских четырехмерных позиций.

В четырехмерной теории нет понятия инертной массы как некоторого зависящего от скорости скаляра. В ней есть лишь понятие собственной массы M, неразрывно связанное с импульсом и энергией. Поэтому «закон зависимости инертной массы от скорости» может быть включен в четырехмерную теорию лишь при обобщении понятия инертной массы вида (15.4). Но такое обобщение представляется искусственным, ибо вектор массы m_k (15.4) ничем, кроме постоянного множителя $1/c$, не отличается от вектора четырехмерного импульса P_k.

Так же обстоит дело и с законом Эйнштейна $E = mc^2$. С трехмерной точки зрения это действительно закон, так как в нем связываются две качественно различной природы величины, одна из которых является характеристикой движения, другая — мерой инертности материи. С четырехмерной же точки зрения это соотношение имеет смысл лишь в форме (15.3). Но в этом случае оно низводится до уровня определения вектора массы m_k через вектор энергии E_k, причем оба эти вектора физически определяются лишь через вектор импульса P_k, который только и имеет прямой физический смысл.

Соотношению Эйнштейна (14.14) может быть придан, однако, и иной, ковариантный, но отличный от (15.3), смысл. В таком смысле это соотношение и применяется в ядерной энергетике. Этот вопрос мы разберем, однако, в следующем параграфе, посвященном закону сохранения энергии и импульса для систем частиц.

§ 16. Закон сохранения энергии, импульса и общей собственной массы

В нерелятивистской механике закон сохранения энергии получается как следствие уравнений динамики системы материальных точек

$$\frac{d}{dt}(m\mathbf{u}_k) = \mathbf{f}_k. \tag{16.1}$$

Из них же вытекает сохранение импульса при условии

$$\sum_k \mathbf{f}_k = 0. \tag{16.2}$$

В релятивистской механике, вследствие невозможности пренебречь запаздыванием взаимодействий, условие (16.2), выражающее закон равенства действия противодействию, уже не выполняется. Однако законы сохранения для систем, содержащих как точечные частицы, так и поле, обеспечивающее взаимодействие этих частиц, могут быть получены как следствие общего закона сохранения энергии-импульса в дифференциальной форме

$$\frac{\partial T_{k\alpha}}{\partial x_\alpha} = 0, \tag{16.3}$$

где $T_{k\alpha}$ – четырехмерный тензор энергии-импульса. Этот вопрос может быть подробно рассмотрен лишь в связи с полевой теорией элементарных частиц. Поэтому здесь мы остановимся лишь на анализе самих законов сохранения импульса и энергии, приведя их без вывода.

Пусть $(e_k/c,\ \mathbf{p}_k)$ — четырехмерный импульс частицы номера k, а $(e^{(f)}/c,\ \mathbf{p}^{(f)})$ – четырехмерный импульс всех полей. Тогда для изолированной системы полей и частиц имеет место закон сохранени общего импульса $(E/c,\ \mathbf{P})$. Следовательно, для системы состоящей из N частиц, величины

$$\begin{aligned} \mathbf{P} &= \sum_{k=1}^{N} \mathbf{p}_k + \mathbf{p}^{(f)}, \\ E &= \sum_{k=1}^{N} e_k + e^{(f)} \end{aligned} \tag{16.4}$$

остаются неизменными, откуда согласно (14.17) сохраняется также общая собственная масса системы, определяемая как

$$M = \frac{1}{c}\sqrt{(E/c)^2 - (\mathbf{P})^2}. \qquad (16.5)$$

Таким образом, в релятивистской теории трехмерные законы сохранения импульса, энергии и массы объединяются в единый четырехмерный закон сохранения четырехмерного импульса или закон сохранения энергии-импульса.

В системе центра масс, где

$$\boldsymbol{P} = 0, \qquad (16.6)$$

закон сохранения полной собственной массы совпадает с законом сохранения энергии (с точностью до постоянного множителя c^2). В этой системе

$$M = \frac{1}{c^2}E_0 = \sum_{k=1}^{N}\frac{M_k}{\sqrt{1-\beta_k^2}} + e^{(f)}/c^2. \qquad (16.7)$$

При скоростях частиц, малых по сравнению со скоростью света, полевая энергия $e^{(f)}$ приобретает смысл энергии попарного взаимодействия частиц, т. е. смысл потенциальной энергии системы U. Таким образом, уравнение (16.7) при $\beta_k \ll 1$ превращается в соотношение

$$\sum_{k=1}^{N} M_k + \frac{1}{c^2}\left\{\sum_{k=1}^{N}\frac{M_k u_k^2}{2} + U\right\} = M = \mathrm{const}. \qquad (16.8)$$

Это соотношение можно записать также в виде

$$\Delta M = \frac{1}{c^2}\left\{\sum_{k=1}^{N}\frac{M_k u_k^2}{2} + U\right\}, \qquad (16.9)$$

где величина

$$\Delta M = M - \sum_{k=1}^{N} M_k \qquad (16.10)$$

называется *дефектом массы*. Таким образом, формула (16.9) означает, что *классическая механическая энергия тела равна дефекту массы, умноженному на квадрат скорости света*.

При отсутствии превращений частиц одних в другие число частиц каждого сорта остается неизменным и в силу сохранения M остается постоянным и дефект массы ΔM и равенство (16.9) приобретает смысл механического закона сохранения энергии, т. е.

$$\sum_{k=1}^{N} \frac{M_k u_k^2}{2} + U = \mathscr{E} = c^2 \Delta M. \qquad (16.11)$$

При этом утверждение, что $\Delta M = \text{const}$ совпадает согласно (16.10) с классическим механическим законом сохранения массы

$$\sum_{k=1}^{N} M_k = \text{const}. \qquad (16.12)$$

Таким образом, в нерелятивистском случае, при отсутствии превращений частиц одних в другие закон сохранения собственной массы распадается на классический закон сохранения энергии и закон сохранения массы.

В случае же наличия превращения частиц одних в другие, распадений систем на части или воссоединений величина $\sum_{k=1}^{N} M_k$, а следовательно, и ΔM изменяется, в силу чего изменяется и механическая энергия системы в соответствии с (16.11). Поэтому классического характера формула (16.11) приобретает смысл соотношения, указывающего, насколько может изменяться механическая энергия системы при качественных превращениях материи из одной формы в другую. Следовательно, классическая теория дополняется некоторым новым положением, означающим возможность возникновения или уничтожения механической энергии при качественных превращениях материи.

Соотношения (16.9), (16.11) могут быть записаны и для случая больших скоростей частиц в релятивистски ковариантном виде.

Пусть $E_0 = Mc^2$ — энергия системы частиц при $\mathbf{P} = 0$, или собственная энергия системы в заданном состоянии. Рассмотрим другое возможное состояние той же системы, когда ее собственная энергия минимальна при заданном составе, определяемом набором M_k, т. е.

собственными массами N частей системы. Если в число этих собственных масс включить также собственную массу полевой энергии $e^{(f)}$, определяемую как

$$M_{N+1} = M^{(f)} = e_0^{(f)} / c^2, \tag{16.13}$$

то собственная энергия в этом состоянии согласно (16.7) равна

$$E_0^{(\min)} = \sum_{k=1}^{N+1} M_k c^2, \tag{16.14}$$

а собственная масса

$$M^{(\min)} = \sum_{k=1}^{N+1} M_k. \tag{16.15}$$

Величины M, $M^{(\min)}$, E_0, $E_0^{(\min)}$ являются четырехмерными скалярами, поэтому согласно (16.7) можно написать следующее ковариантное соотношение

$$\mathscr{E} = E_0 - E_0^{(\min)} = c^2 (M - M^{(\min)}). \tag{16.16}$$

Это релятивистски ковариантное соотношение имеет более глубокий физический смысл, чем формула (16.7), выражающая пропорциональность собственной энергии и собственной массы системы. Оно выражает то максимальное количество тепла, которое может быть выделено системой, имеющей начальную собственную массу M при ее превращении в систему, состоящую из N частиц с собственными массами M_k и имеющей минимальную полевую энергию взаимодействия, равную $c^2 M^{(f)}$. Таким образом, величина $\mathscr{E}$ является энергией в термодинамическом смысле. Поэтому величину $\mathscr{E}$ можно назвать активной энергией системы заданного состава.

Согласно (16.10), (16.11) и (16.15) формулу (16.16) можно записать также в виде

$$\mathscr{E} = c^2 \Delta M - U, \tag{16.17}$$

где использовано обозначение

$$U = e_0^{(f)} = M^{(f)} c^2.$$

Если в конечном состоянии частицы пространственно разобщены, так что полевой энергией взаимодействия U можно пренебречь, то формула (16.17) приобретает упрощенный вид

$$\mathscr{E} = c^2 \Delta M. \tag{16.18}$$

Эта формула определяет ту максимальную энергию, которая может быть использована при распаде исходной системы собственной массы M на совокупность N частиц с собственными массами M_1, M_2, ..., M_N.

Если система, состоящая из N' частиц с собственными массами M_1', M_2', ..., $M_{N'}'$, превращается в систему N частиц с собственными массами M_1, M_2, ..., M_N, то приращение активной энергии такой системы согласно (16.16) будет равно

$$\Delta\mathscr{E} = -c^2 \Delta M^{(\min)}. \tag{16.19}$$

В этой формуле, очевидно,

$$\Delta M^{(\min)} = \sum_{k=1}^{N'+1} M_k' - \sum_{k=1}^{N+1} M_k. \tag{16.20}$$

Если же в начальном и конечном состоянии мы имеем дело с совокупностями свободных частиц, т. е. можем пренебречь их энергией связи, то полевыми массами $M_{N'+1}'$ и M_{N+1} можно пренебречь и вместо (16.20) писать

$$\Delta M^{(\min)} = \sum_{k=1}^{N'} M_k' - \sum_{k=1}^{N} M_k. \tag{16.21}$$

Формула (16.19) определяет активную энергию, образующуюся при ядерных превращениях или превращениях элементарных частиц, т. е. при превращениях материи из одной формы в другую.

§ 17. Превращается ли масса в энергию?

Формула (16.16) имеет иное физическое содержание, чем формула (14.14), однако она также называется соотношением Эйнштейна. В смысле (16.17) соотношение или закон Эйнштейна применяется в ядерной энергетике. На основании этого соотношения утверждают, что масса тела есть мера содержащейся в нем энергии. Часто, ссылаясь на эту формулу или на формулу (16.19), говорят о превращении массы в энергию.

Утверждение о превращении массы в энергию вызывает особенно энергичные возражения со стороны философов-материалистов и, наоборот, используется философами-идеалистами для опровержения материализма. Дело в том, что под энергией привыкли понимать количественную меру движения при его превращении из одной формы в другую. Под массой же понимают количественную меру материи.

Такие воззрения сложились, во всяком случае, в конце прошлого века и они использовались Энгельсом в его «Диалектике природы». С этой точки зрения превращение массы в энергию означает уничтожение материи и порождение движения, т. е. превращение материи в движение. Это утверждение противоречит диалектическому материализму, исходящему из неуничтожимости материи и ее движения, и не встречает возражений в идеалистической философии. Оно может быть даже использовано для опровержения основного тезиса материализма о существовании материи как объективной реальности, существующей вне нашего сознания. Для того чтобы разрешить эти противоречия, рассмотрим, возможно ли с точки зрения физики в каком-либо смысле говорить о превращении массы в энергию.

Если, исходя из четырехмерных представлений о пространстве и времени, рассматривать энергию как величину, пропорциональную нулевой компоненте четырехмерного импульса, а под собственной массой понимать лишь абсолютную величину четырехмерного импульса (его «длину»), то утверждение о превращении массы в энергию теряет какой бы то ни было смысл, ибо энергия и масса — это лишь разные проекции одной и той же физической величины и во всех превращениях обе эти величины сохраняются. При любых физических превращениях в изолированной системе полная энергия и полная собственная масса системы остаются неизменными. Следовательно, в последовательной релятивистской теории никаких превращений массы в энергию не происходит, и формулы (14.14), (14.17), (16.5), (16.7) выражают лишь органическую связь, существующую между понятиями массы, энергии и импульса.

Законы же сохранения энергии-импульса и собственной массы (16.4), (16.5) представляют в четырехмерной теории единый закон, выражающий как сохранение материи при ее превращениях из одной формы в другую, так и неуничтожимость движения.

Таким образом, никакой возможности уничтожения материи или движения из релятивистских формул, связывающих энергию, импульс и собственную массу, не следует. Из них лишь следует необходимость объединения всеобщего закона сохранения и превращения материи и закона сохранения движения в единый закон природы, из которого следует как неуничтожимость материи, так и неуничтожимость движения. Это новое положение, очевидно, никак не противоречит диалектическому материализму, но дополняет и развивает его.

С другой стороны, если отвлечься от точных релятивистских понятий энергии и собственной массы и под энергией понимать лишь ту часть величины $E = cP_0$, которая способна превращаться в тепло, то формулы (16.9), (16.16) можно истолковать как выражающие превращение собственной массы как некоторой новой формы энергии в энергию, понимаемую в классическом смысле. Действительно, согласно (16.9), (16.16) эта «классическая энергия» не сохраняется при превращениях, идущих с изменением суммы собственных масс частей, составляющих систему. Она увеличивается согласно (16.9) на величину, пропорциональную уменьшению суммы собственных масс системы. Таким образом, с этой классической точки зрения при качественных превращениях материи, идущих с изменением суммы собственных масс составляющих частей, классическая энергия не сохраняется. Но так же как и в классической механике не сохраняется кинетическая энергия, а сохраняется сумма кинетической и потенциальной энергии, в случае качественных превращений материи можно считать, что сохраняется классическая энергия плюс некоторая скрытая форма энергии, равная разности сумм собственных масс частей после превращения и до превращения, умноженной на c^2. Следовательно, если $c^2 \Delta M^{(\min)}$ рассматривать как скрытую энергию, то при качественном превращении материи происходит не превращение массы в энергию, а п р е в р а щ е н и е скрытой формы энергии в активную классическую энергию. При этом толковании, однако, формулу (16.19) правильнее писать в виде

$$\Delta \mathscr{E}_{\text{кл}} = \Delta \mathscr{E} + c^2 \Delta M^{(\min)} = 0. \qquad (17.1)$$

Совершенно аналогично можно спасти и классический закон сохранения массы, если выделяющейся классической энергии приписать некоторую собственную массу $\Delta\mathscr{E}/c^2$ и, записав (16.19) в виде

$$\Delta M_{\text{кл}} = \Delta M^{(\min)} + \Delta\mathscr{E}/c^2 = 0, \tag{17.2}$$

говорить о сохранении суммы собственных масс плюс собственная масса энергии. Здесь опять же масса превращается в массу другого рода, но не в энергию.

Итак, утверждение, что масса превращается в энергию, лишено физического смысла. Использование этого утверждения для извлечения философских выводов также лишено основания.

Даже истолкование выводов из соотношения (16.19) как несохранения классической энергии при качественных превращениях материи не может служить основанием для утверждения об уничтожимости движения, поскольку классическая энергия является только частью полной энергии, которая сохраняется.

§ 18. Неаддитивность собственных масс

В отличие от энергии и импульса собственная масса не является аддитивной величиной, т. е. собственная масса системы (M) отлична от суммы собственных масс составляющих ее частиц (M_k) Последнее непосредственно видно из общего выражения для собственной массы системы частиц, которая согласно (16.4), (16.5) равна

$$M = \frac{1}{c}\sqrt{\Big(\sum_k e_k / c\Big)^2 - \Big(\sum_k \mathbf{p}_k\Big)^2}\,. \tag{18.1}$$

Собственные же массы отдельных частиц равны *

$$M_k = \frac{1}{c}\sqrt{(e_k / c)^2 - (\mathbf{p}_k)^2}\,. \tag{18.2}$$

Замечая, что согласно (14.13) импульсы можно представить как

$$\mathbf{p} = \frac{e}{c}\boldsymbol{\beta}, \text{ где } \boldsymbol{\beta} = \frac{\mathbf{u}}{c}, \tag{18.3}$$

---

* В случае наличия полевой массы и полевого импульса их также надо включить в суммы, стоящие под корнем выражения (18.1).

выражение (18.1) запишем в виде

$$M=\frac{1}{c}\sqrt{\left(\sum_k \frac{e_k}{c}\right)^2-\left(\sum_k \frac{e_k}{c}\,\boldsymbol{\beta}_k\right)^2}. \tag{18.4}$$

Из этого выражения видно, что $M=\sum\limits_k M_k$ лишь в том случае, если все скорости частиц одинаковы, т. е.

$$\boldsymbol{\beta}_k=\boldsymbol{\beta}. \tag{18.5}$$

Если же скорости частиц различны и все $e_k>0$ *, то $M>\sum\limits_k M_k$, поскольку в системе центра масс выражение (18.4) имеет вид

$$M=\frac{1}{c^2}\sum_k e_k'=\sum_k\frac{M_k}{\sqrt{1-\beta_k'^2}}>\sum_k M_k, \tag{18.6}$$

где $\boldsymbol{\beta}_k'=\mathbf{u}_k'/c$; $\mathbf{u}_k'$ — скорости частиц в системе центра масс.

Особый интерес представляет случай, когда система состоит из частиц, движущихся со скоростью света. Для таких частиц $\beta_k^2=1$ и согласно (18.2) и (18.3) собственные массы равны нулю. Однако формула (18.4) показывает, что общая собственная масса такой системы, состоящей из частиц с нулевой собственной массой, может быть равна нулю только при выполнении условия (18.5), т. е. только в том случае, если все частицы движутся в одном и том же направлении. Если же направления скоростей различных частиц не совпадают одно с другим, то

$$\left(\sum_k \frac{e_k}{c}\boldsymbol{\beta}_k\right)^2<\left(\sum_k \beta\frac{e_k}{c}\right)^2=\left(\sum_k\frac{e_k}{c}\right)^2, \tag{18.7}$$

поскольку $\beta_k^2=\beta^2=1$. Отсюда согласно (18.4) получаем $M>0$.

Таким образом, любой пространственно ограниченный световой пакет, состоящий из фотонов, имеющих скорости, равные скорости света, но распределенные в некотором пространственном угле, будет иметь собственную массу, отличную от нуля.

---

* Вопрос о возможности частиц, для которых $e_k<0$ и $M_k<0$, специально рассмотрен в гл. VII.

Итак, если об отдельных фотонах мы говорим, что они имеют собственную массу, равную нулю, то о свете вообще мы этого уже сказать не можем. *Любой реальный световой поток имеет отличную от нуля собственную массу.* Только бесконечная плоская световая волна, т. е. поток строго коллинеарных фотонов имеет равную нулю общую сообственную массу. Но такой случай светового потока практически никогда не осуществим, так как любой реальный световой поток пространственно ограничен, т. е. не является бесконечной плоской волной.

V

ВОЗМОЖНЫ ЛИ СКОРОСТИ, БО́ЛЬШИЕ СКОРОСТИ СВЕТА?

§ 19. Возможные скорости физических процессов

В физике приходится иметь дело как со скоростями, меньшими или равными скорости света, так и со скоростями, превышающими скорость света. Макроскопические тела и все известные элементарные частицы, имеющие действительную, положительную собственную массу, во всех системах отсчета движутся со скоростью, меньшей скорости света. Фотоны, имеющие нулевую собственную массу, движутся со скоростью света. Известны также физические процессы, распространяющиеся со скоростью, большей скорости света. Например, фазовая скорость распространения волн де Бройля или фазовая скорость электромагнитных волн в разреженной электронной плазме превышают скорость света. Таким образом, физические процессы могут характеризоваться как скоростями, меньшими c, так и скоростями, превышающими c.

В конце § 13 было доказано, что посредством перехода от одной инерциальной системы отсчета к другой никакие досветовые скорости не могут быть сделаны сверхсветовыми и, наоборот, сверхсветовые скорости не могут быть превращены в досветовые. Процессы же, распространяющиеся со скоростью света во всех инерциальных системах, имеют одну и ту же скорость c. При выводе этого утверждения, конечно, предполагалось, что все инерциальные системы отсчета имеют досветовые скорости, т. е. $v < c$. Это ограничение связано с самой структурой преобразований Лоренца, которые теряют смысл при $v > c$, так как штрихованные координаты при этом становятся мнимыми, последнее же не поддается какому-либо физическому истолкованию.

Итак, принимая, что реальное физическое пространство и время могут изображаться инерциальными системами отсчета, лишь движущимися друг относительно друга со скоростями $v < c$, мы приходим к заключению о качественном различии процессов, распространяющихся с досветовыми, световыми и сверхсветовыми скоростями.

С четырехмерной точки зрения сверхсветовые скорости описываются четырехмерным вектором, компоненты которого должны быть мнимыми величинами, поскольку при $u > c$ согласно (13.2)

$$(ds)^2 = (dx_0)^2(1 - u^2/c^2) < 0, \tag{19.1}$$

т. е. ds — мнимая величина, откуда компоненты U_k, определяемые из (13.5), также могут быть лишь мнимыми. Следовательно, вместо (13.14) для четырехмерных компонент сверхсветовой скорости надо написать

$$(U_0, \mathbf{U}) = \left(\frac{ic}{\sqrt{u^2/c^2 - 1}}, \frac{i\mathbf{u}}{\sqrt{u^2/c^2 - 1}}\right), \tag{19.2}$$

причем $\mathbf{u}$ и $\sqrt{u^2/c^2 - 1}$ — действительные величины. Из этого выражения также видно, что даже с формально кинематической точки зрения сверхсветовые скорости качественно отличны от досветовых и не могут быть превращены в них посредством преобразований Лоренца.

Как мы уже отметили, сверхсветовыми могут быть фазовые скорости распространения некоторых видов волн. В качестве примера рассмотрим волны спинорного поля ψ_α. Для монохроматической плоской волны*.

$$\psi = \psi_0 e^{i[\omega t - \mathbf{kr}]} \tag{19.3}$$

фазовой скоростью является величина

$$\mathbf{u} = \frac{\omega}{k}\mathbf{n}, \tag{19.4}$$

---

* Спинорный индекс α в дальнейшем опускается и подразумевается, что $\psi_1^*\psi_1 + \psi_2^*\psi_2 + \psi_3^*\psi_3 + \psi_0^*\psi_0 = \psi^*\psi$.

где

$$\mathbf{n} = \mathbf{k}/k \tag{19.5}$$

есть нормаль фронта волны или направление ее распространения. Используя (19.4), монохроматическую волну поля ψ можно записать в виде

$$\psi = \psi_0 e^{i\omega(t - \mathbf{nr}/u)}. \tag{19.6}$$

Из этого выражения видно, что $\mathbf{u}$ является скоростью перемещения поверхности заданной постоянной фазы монохроматической плоской волны.

Монохроматическая плоская волна любого поля является, однако, практически никогда не осуществимым идеальным частным решением уравнений поля. В действительности приходится иметь дело с ограниченными в пространстве и во времени волновыми пакетами, которые можно изобразить как суперпозицию плоских волн с различными амплитудами и волновыми векторами $\mathbf{k}$, т. е. как

$$\psi(\mathbf{r}, t) = \int\limits_{-\infty}^{+\infty} g(\mathbf{k})\, e^{i[\omega(\mathbf{k})t - \mathbf{kr}]}\, dk, \tag{19.7}$$

где $d\mathbf{k} = dk_x dk_y dk_z$, а функция $\omega(\mathbf{k})$ определяет *закон дисперсии* рассматриваемых волн.

Такой пространственно ограниченный волновой пакет перемещается в пространстве как целое с некоторой скоростью, которую можно определить как скорость движения «центра тяжести» пакета. Под координатами «центра тяжести» пакета можно понимать компоненты вектора

$$\mathbf{R} = \frac{\int\limits_{-\infty}^{+\infty} \mathbf{r}\,|\psi|^2\, d\mathbf{r}}{\int\limits_{-\infty}^{+\infty} |\psi|^2\, d\mathbf{r}}. \tag{19.8}$$

Знаменатель этого выражения в силу (19.7) равен

$$\iiint\limits_{-\infty}^{+\infty} g(\mathbf{k})\, g^*(\mathbf{k}')\, e^{i[\omega(\mathbf{k}) - \omega(\mathbf{k}')]\, t - i(\mathbf{k} - \mathbf{k}')\mathbf{r}}\, d\mathbf{k}\, d\mathbf{k}'\, d\mathbf{r} =$$

$$= (2\pi)^3 \int\limits_{-\infty}^{+\infty} g(\mathbf{k})\, g^*(\mathbf{k})\, d\mathbf{k}, \tag{19.9}$$

так как

$$\int\limits_{-\infty}^{+\infty} e^{-i(\mathbf{k}-\mathbf{k}')\mathbf{r}}\, d\mathbf{r} = (2\pi)^3 \delta(\mathbf{k}-\mathbf{k}'), \tag{19.10}$$

где

$$\delta(\mathbf{k}-\mathbf{k}') = \delta(k_x - k'_x)\,\delta(k_y - k'_y)\,\delta(k_z - k'_z).$$

Числитель же (19.8) равен

$$\iiint\limits_{-\infty}^{+\infty} g(\mathbf{k})\, g^*(\mathbf{k}')\, \mathbf{r}\, e^{i[\omega(\mathbf{k})-\omega(\mathbf{k}')]t-i(\mathbf{k}-\mathbf{k}')\mathbf{r}}\, d\mathbf{k}\, d\mathbf{k}'\, d\mathbf{r} =$$

$$= \iiint\limits_{-\infty}^{+\infty} g(\mathbf{k})\, g^*(\mathbf{k}')\, e^{i[\omega(\mathbf{k})-\omega(\mathbf{k}')]t}\, i\nabla_k e^{-i(\mathbf{k}-\mathbf{k}')\mathbf{r}}\, d\mathbf{k}\, d\mathbf{k}'\, d\mathbf{r}, \tag{19.11}$$

где оператор ∇_k есть вектор с компонентами

$$\nabla_k = \left(\frac{\partial}{\partial k_x}, \frac{\partial}{\partial k_y}, \frac{\partial}{\partial k_z}\right). \tag{19.12}$$

Если $g(\mathbf{k})$ — достаточно быстро убывающая при $\mathbf{k} \to \infty$ функция, то, интегрируя (19.11) по частям и применяя (19.10), получаем для числителя (19.8) выражение

$$-(2\pi)^3 \int\limits_{-\infty}^{+\infty} [g^*(\mathbf{k})\, i\nabla_k g(\mathbf{k}) - g^*(\mathbf{k})\, g(\mathbf{k})\, t\nabla_k \omega(\mathbf{k})]\, d\mathbf{k}. \tag{19.13}$$

Знаменатель (19.9) не зависит от времени, так же как и первый член числителя (19.13). Следовательно, скорость движения «центра тяжести» равна

$$\frac{d\mathbf{R}}{dt} = \mathbf{V} = \frac{\int\limits_{-\infty}^{+\infty} \nabla_k \omega(\mathbf{k})\, |g(\mathbf{k})|^2\, d\mathbf{k}}{\int\limits_{-\infty}^{+\infty} |g(\mathbf{k})|^2\, d\mathbf{k}}. \tag{19.14}$$

Вектор

$$\nabla_k \omega(\mathbf{k}) = \left(\frac{\partial\omega}{\partial k_x}, \frac{\partial\omega}{\partial k_y}, \frac{\partial\omega}{\partial k_z}\right) = \mathbf{v} \tag{19.15}$$

называется групповой скоростью волны. Таким образом, скорость «центра тяжести» пакета согласно (19.14) может рассматриваться как средняя групповая скорость пакета.

Если пакет квазимонохроматический, т. е. $|g(\mathbf{k})|^2$ приближается к δ-функции вида $\delta(\mathbf{k}-\mathbf{k}_0)$, то средняя групповая скорость практически совпадает с истинной групповой скоростью при $\mathbf{k}=\mathbf{k}_0$. Следовательно, «центр тяжести» квазимонохроматической группы волн движется с групповой скоростью $\nabla_k\omega(\mathbf{k}_0)$.

Легко убедиться в том, что волновой пакет (19.7) не только перемещается как целое, но и изменяет свою ширину.

За меру ширины пакета естественно принять величину $\Delta(\mathbf{r})$, определяемую аналогично среднему квадратичному уклонению выражением

$$[\Delta(\mathbf{r})]^2=\frac{\int\limits_{-\infty}^{+\infty}(\mathbf{r}-\mathbf{R})^2|\psi|^2d\mathbf{r}}{\int\limits_{-\infty}^{+\infty}|\psi|^2d\mathbf{r}},\tag{19.16}$$

или

$$\Delta(\mathbf{r})=\sqrt{\frac{\int\limits_{-\infty}^{+\infty}r^2|\psi|^2d\mathbf{r}}{\int\limits_{-\infty}^{+\infty}|\psi|^2d\mathbf{r}}-R^2}.\tag{19.17}$$

Числитель первого члена подкоренного выражения (19.17) согласно (19.7) можно представить в виде

$$\int\limits_{-\infty}^{+\infty}r^2|\psi|^2d\mathbf{r}=\iiint\limits_{-\infty}^{+\infty}g(\mathbf{k})\,g^*(\mathbf{k}')\,r^2e^{i[\omega(\mathbf{k})-\omega(\mathbf{k}')]t-i(\mathbf{k}-\mathbf{k}')\mathbf{r}}\,d\mathbf{k}\,d\mathbf{k}'\,d\mathbf{r}=$$

$$=-\iiint\limits_{-\infty}^{+\infty}g(\mathbf{k})\,g^*(\mathbf{k}')\,e^{i[\omega(\mathbf{k})-\omega(\mathbf{k}')]t}\,\nabla_k^2e^{-i(\mathbf{k}-\mathbf{k}')\mathbf{r}}\,d\mathbf{k}\,d\mathbf{k}'\,d\mathbf{r}.$$

Интегрируя это выражение дважды по частям и применяя (19.10), получаем

$$\int\limits_{-\infty}^{+\infty}r^2|\psi|^2d\mathbf{r}=-(2\pi)^3\int\limits_{-\infty}^{+\infty}g^*(\mathbf{k})\{\nabla_k^2g(\mathbf{k}')-$$

$$-it\,[2\nabla_k\omega(\mathbf{k})\,\nabla_kg(\mathbf{k})-g(\mathbf{k})\nabla_k^2\omega(\mathbf{k})]-$$

$$-t^2g(\mathbf{k})\,[\nabla_k\omega(\mathbf{k})]^2\}\,d\mathbf{k}.\tag{19.18}$$

Подставляя (19.18) и (19.19) в (19.17), а также учитывая, что согласно (19.14)

$$\mathbf{R} = \mathbf{V}t + \mathbf{R}_0, \tag{19.19}$$

получаем

$$\Delta(\mathbf{r}) = A\sqrt{(t - t_0)^2 + D}, \tag{19.20}$$

где

$$A^2 = \frac{\int\limits_{-\infty}^{+\infty} [\nabla_k \omega(\mathbf{k})]^2 \, |g(k)|^2 \, d\mathbf{k} - \left[\int\limits_{-\infty}^{+\infty} \nabla_k \omega(\mathbf{k}) \, |g(\mathbf{k})|^2 \, d\mathbf{k}\right]^2}{\int\limits_{-\infty}^{+\infty} |g(\mathbf{k})|^2 \, d\mathbf{k}}. \tag{19.21}$$

Константы t_0 и D также могут быть выражены через $\mathbf{V}$, $\mathbf{R}_0$ и интегральные члены, входящие в (19.18).

Таким образом, скорость расширения волнового пакета при $t \to \infty$ становится равной A. Эта скорость стремится к нулю для квазимонохроматических пакетов, как это видно из (19.21).

Мы рассмотрели здесь случай комплексного спинорного поля ψ. Можно, однако, показать, что для других полей получаются аналогичные выводы и, таким образом, любые квазимонохроматические волновые пакеты как целое движутся с групповой скоростью *.

Следовательно, не фазовая, а групповая скорость физически характеризует движение волновых пакетов. С этой скоростью, очевидно, перемещается и энергия поля, так как последняя сосредоточена внутри пакета, так сказать, вокруг его центра тяжести.

Оказывается, что для всех известных волновых процессов, распространяющихся со сверхсветовой фазовой скоростью, групповая скорость является досветовой. Так, например, для волн де Бройля

$$\omega = c\sqrt{k^2 + c^2M^2/\hbar^2}, \tag{19.22}$$

откуда согласно (19.4) и (19.15) произведение фазовой скорости на групповую равно

$$uv = c^2. \tag{19.23}$$

Отсюда следует, что при $u > c$, $v < c$.

---

* Эти вопросы рассмотрены в статье Р. Серебряного [6].

Таким образом, подтверждается распространенное мнение, что реальные физические скорости, с которыми переносится энергия, являются досветовыми или, в крайнем случае, равными скорости света. Сверхсветовыми являются лишь фазовые скорости, т. е. скорости переноса фазы волны, с которыми не связаны процессы переноса материи или энергии.

Ниже мы, однако, увидим, что последнее заключение может быть критически пересмотрено на основании более глубокого физического анализа.

§ 20. Предельность скорости света как следствие принципа причинности

В теории относительности доказывается общая теорема о том, что ни один с и г н а л не может распространяться со скоростью, большей скорости света. Под понятием «сигнал» неявно понимается некоторое макроскопическое воздействие, могущее вызвать макроскопически наблюдаемые эффекты. Считается также, то сигнал переносит воздействие из одной точки в другую, т. е. испускается по нашему желанию в точке *1* и поглощается, возбуждая некоторые макроскопические процессы, в точке *2*.

Таким образом, сигнал представляется как некоторая макроскопическая порция энергии, переносимая каким-то физическим агентом из одной точки в другую. Таким переносящим агентом может быть имеющая собственную массу частица, отдельный фотон или группа волн. Фаза волны, очевидно, не может быть сигналом, так как плоская волна должна быть как-то искажена, чтобы переносить какое-то заданное дискретное воздействие. Любое же искажение монохроматической плоской волны составит уже группу волн, которая будет распространяться не с фазовой, а с групповой скоростью.

Докажем упомянутую выше теорему о невозможности сверхсветовых сигналов.

Рассмотрим два события: испускание сигнала в точке x_1 в момент времени t_1 и поглощение сигнала в точке x_2 в последующий момент времени t_2. Скоростью сигнала, очевидно, является величина

$$u = (x_2 - x_1) / (t_2 - t_1). \tag{20.1}$$

При этом предполагается, что

$$t_2 > t_1, \tag{20.2}$$

так как поглощение сигнала, по определению,— более позднее событие, чем его испускание.

Рассмотрим эти же два события с точки зрения другой инерциальной системы отсчета Σ', движущейся относительно исходной со скоростью v. Согласно преобразованиям Лоренца [см. формулу (11.6)]:

$$t_2' - t_1' = \frac{(t_2 - t_1) - \frac{v}{c^2}(x_2 - x_1)}{\sqrt{1 - v^2/c^2}}, \tag{20.3}$$

откуда согласно определению (20.1)

$$t_2' - t_1' = (t_2 - t_1)\frac{1 - vu/c^2}{\sqrt{1 - v^2/c^2}}. \tag{20.4}$$

Если $u < c$, то в силу того, что всегда $v < c$, получаем $vu/c^2 < 1$ и множитель при $(t_2 - t_1)$ является положительной величиной. Следовательно, $(t_2' - t_1') > 0$, если $(t_2 - t_1) > 0$, т. е.

$$t_2' > t_1' \text{ при } u < c. \tag{20.5}$$

Таким образом, временнáя последовательность событий, связанных досветовым сигналом, не может быть изменена посредством перехода к другой системе отсчета.

Иное положение имеет место в случае сверхсветовых сигналов. Если $u > c$, то согласно (20.4) возможно указать такую систему отсчета Σ', движущуюся со скоростью $v < c$, в которой множитель при $(t_2 - t_1)$ становится отрицательным. Действительно, при $u > c$, т. е. $c/u < 1$, существуют такие $v < c$, при которых удовлетворяется неравенство

$$c/u < v/c < 1, \tag{20.6}$$

из которого следует, что $vu/c^2 > 1$.

Таким образом, при скорости системы отсчета, определяемой неравенством (20.6), множитель $(t_2 - t_1)$ в (20.4) становится отрицательным и, следовательно, при $(t_2 - t_1) > 0$ получаем $(t_2' - t_1') < 0$, т. е. временнáя последовательность событий нарушается, так как при $t_1 < t_2$ получаем

$$t_2' < t_1'. \tag{20.7}$$

Следовательно, допущение сверхсветовых сигналов эквивалентно допущению возможности изменения временно́й последовательности испускания и поглощения путем выбора системы отсчета. Но допущение такой возможности противоречит принципу причинности в его принятой в физике формулировке, ибо путем выбора системы отсчета причина — испускание сигнала может быть осуществлена после следствия — поглощения сигнала.

Особенно нелепой представляется такая возможность обращения временно́й последовательности причины и следствия в следующем макроскопическом примере.

Пусть источником сигнала является ружье, сигналом — пуля, а приемником сигнала — мишень (рис. 12).

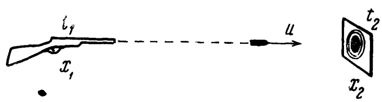

Рис. 12

Если скорость пули меньше скорости света, то с точки зрения любой системы отсчета выстрел ружья, т. е. причина, предшествует попаданию пули в мишень, т. е. следствию. Если же скорость пули превышает скорость света, то всегда можно найти такую систему отсчета, в которой попадание пули в мишень окажется событием более ранним, чем выстрел ружья *. Нелепость, т. е. неосуществимость подобного обратного процесса, приводится как совершенно очевидный аргумент в пользу невозможности нарушения принципа причинности, а следовательно, невозможности сверхсветовых сигналов.

Исходя из этой теоремы, считают, что в теории относительности невозможны какие-либо тела, движущиеся со скоростью, большей скорости света, и какие-либо

* Точнее, это обращение последовательности событий во времени надо представлять как обращение во времени всего процесса, т. е. сперва тепловая энергия мишени так концентрируется, что пуля выбрасывается мишенью и летит в ствол ружья. При этом собираются все продукты распада взрывчатки и полностью восстанавливается первоначальный заряд. Ружье становится заряженным, и на этом процесс прекращается.

процессы, переносящие энергию и распространяющиеся со скоростью, большей скорости света. Отсюда же доказывается невозможность абсолютно твердых тел, т. е. упругих тел, в которых упругие волны могут распространяться с бесконечно большой скоростью. Причем эти выводы надо рассматривать не как следствие взаимосвязи пространства и времени, выражаемой группой Лоренца, а как следствие п р и н ц и п а п р и ч и н н о с т и, примененного в теории относительности. И совершенно неправильным является утверждение, что теория относительности, так сказать, сама по себе, без привлечения других законов природы якобы запрещает сверхсветовые сигналы и переносящие энергию возмущения, распространяющиеся со скоростью, большей скорости света.

Итак, нам пришлось воспользоваться некоторым дополнительным допущением, называемым «принцип причинности», для того чтобы сделать чисто физический вывод о невозможности сверхсветовых сигналов и иных переносящих физическое воздействие возмущений. Но физические следствия могут быть извлечены только из физических же законов или постулатов. Таким образом, принцип причинности надо рассматривать как некоторое физическое утверждение или как некоторый физический закон.

В общефилософском плане принцип причинности связывается с представлением о причинной обусловленности одних явлений другими. От всеобщей взаимосвязи причинная обусловленность отличается направленностью действия от причин к следствиям. Эта обусловленность представляется как всеобщая, т. е. считается, что не существует каких-либо явлений, причинно не обусловленных другими явлениями. В данном параграфе мы, однако, понимали под принципом причинности некоторое более узкое утверждение, касающееся двух событий, происходящих в двух пространственно разобщенных точках x_1 и x_2 в моменты времени t_1 и t_2. Эти события мы считали *причинно обусловленными*, причем событие в точке *1* рассматривалось как п р и ч и н а, а событие в точке *2* — как с л е д с т в и е. Под принципом же причинности понималось утверждение, что причина обязательно п р е д ш е с т в у е т следствию и что эта временна́я последовательность событий абсолютна, т. е. не может быть изменена выбором системы отсчета или точки зрения наблюдателя.

Иначе говоря, под принципом причинности понималось следующее утверждение: «Из двух причинно обусловленных событий, происходящих в двух пространственно разобщенных точках, одно является причиной, другое — следствием, причем причина всегда предшествует по времени следствию, и эта последовательность не может быть нарушена выбором системы отсчета».

Основным в этом утверждении является то, что временна́я последовательность причинно связанных событий абсолютна и не может быть нарушена или, говоря обыденным языком, что «время течет в одном направлении от прошедшего к будущему». Следовательно, «принцип причинности» и представление о направленности течения физических процессов во времени — по существу одно и то же физическое допущение [7,8].

§ 21. Принцип причинности и второе начало термодинамики

Еще в начале этого века со всеобщим признанием статистической физики как основы термодинамики и вообще физики макроскопических процессов было установлено, что направленность физических процессов во времени обусловлена исключительно статистическими законами, следствием которых является второе начало термодинамики.

В силу абсолютной обратимости во времени микроскопических законов движения направление течения времени микроскопически никак не выделено. Все микроскопические процессы совершаются абсолютно симметрично как при рассмотрении их в положительном (от прошедшего к будущему), так и в отрицательном (от будущего к прошедшему) направлении течения времени. Всякая направленность микроскопических процессов возникает лишь вследствие особых начальных условий, целиком определяемых макроскопической обстановкой опыта. Иначе говоря, всякая необратимость и направленность во времени являются следствием макроскопической необратимости, наблюдаемой в окружающем макромире. Макроскопическая же необратимость является следствием второго начала термодинамики, имеющего чисто статистическую природу.

Поскольку «принцип причинности» в его узкой физической формулировке является выражением направленности процессов во времени, а последняя вытекает из второго начала термодинамики, постольку «принцип причинности» можно рассматривать как следствие или специальное выражение второго начала термодинамики.

Альтернативная точка зрения, отрицание связи «принципа причинности» и второго начала, означала бы попытку ввести некоторый особый физический закон, постулирующий для всех физических процессов определенную и неизменную направленность во времени. Для микромира это означало бы приписывание некоторой несимметрии во времени законам микродвижения. Последнее, однако, никем еще не было успешно введено. Таким образом, интуитивное желание рассматривать «принцип причинности» как некоторый абсолютный (т. е. нигде не нарушаемый) закон природы не имеет пока достаточных оснований с точки зрения известных законов движения.

Трудно даже себе представить, что означало бы открытие истинной необратимости в законах микродвижения. Для любых законов, выражаемых какими-либо дифференциальными уравнениями, необратимость во времени означала бы появление членов, несимметричных по отношению к изменению знака времени. Последнее же неразрывно связано с монотонным изменением во времени некоторых физических величин, т. е. с какими-то законами несохранения. В микромире же, наоборот, все построено на сохранении движения и сохранении материи с большей несомненностью, чем в макромире.

Итак, имеются все основания считать «принцип причинности», т. е. утверждение о неизменяемости временно́й последовательности причинно обусловленных пространственно разобщенных событий, лишь правилом, имеющим макроскопическую природу и выделяющим определенное направление течения времени. Это правило есть просто следствие закона возрастания энтропии.

Но с этой точки зрения возможны и нарушения «принципа причинности» подобно тому, как возможны нарушения второго начала термодинамики во флуктуационных процессах. Следовательно, запреты, накладываемые принципом причинности в теории относительности, не могут быть абсолютными и касаются только процессов макроскопического характера. Таким образом, необходимо пересмотреть те выводы, которые делаются на основании «принципа причинности» в теории относительности, выводы о невозможности сигналов и иных возмущений, распространяющихся со сверхсветовой скоростью.

Для того чтобы не внести в рассуждения каких-либо априорных элементов, проистекающих из интуитивного представления о направленности времени, мы будем исходить прямо из второго начала термодинамики и исследуем те запреты, которые оно накладывает на процессы передачи физических воздействий.

§ 22. Сигнал как физический процесс, переносящий негаэнтропию

С точки зрения теории информации получение сигнала означает получение информации. Следовательно, посылка сигналов из точки *1* в точку *2* означает перенос информации от первой точки ко второй.

Известно, что информация пропорциональна негаэнтропии, т. е. избытку равновесной энтропии над неравновесной, поэтому под сигналом надо понимать лишь такое локализованное возмущение, которое, перемещаясь от излучателя к приемнику, переносит негаэнтропию. Локализованное возмущение может переносить энергию * и тем самым изменять равновесную энтропию приемника, но не быть сигналом, если оно не переносит негаэнтропию, являющуюся мерой неравновесности системы.

Последнее понятно и с чисто термодинамической точки зрения, без привлечения кибернетических соображений. Действительно, для того чтобы поглощенный сигнал мог бы быть зарегистрирован, он должен произвести какое-то макроскопическое действие в системе приемника, т. е. возбудить в ней какой-то спонтанный макроскопический процесс. Последнее же возможно лишь при сообщении системе негаэнтропии, необходимой, например, для перевода системы из метастабильного состояния в стабильное через разделяющий эти состояния энтропийный минимум.

* Под энергией понимается ее релятивистское выражение (14.13).

Пусть ΔS_1 — приращение энтропии излучателя, а ΔS_2 — приращение энтропии приемника сигнала. В силу второго начала термодинамики, очевидно

$$\Delta S_1 + \Delta S_2 \geqslant 0. \tag{22.1}$$

Процесс, переносящий энергию ΔE, может быть сигналом лишь при условии

$$\Delta S_2 < \Delta S_2^0 \text{ или } \Delta (S_2 - S_2^0) < 0, \tag{22.2}$$

где ΔS^0 — равновесное приращение энтропии системы при увеличении ее энергии на ΔE. Очевидно, $\Delta S_1^0 = -\Delta S_2^0$, и поэтому, в силу (22.1)

$$\Delta S_1 > \Delta S_1^0 \text{ или } \Delta (S_1 - S_1^0) > 0, \tag{22.3}$$

т. е. при излучении сигнала негаэнтропия излучателя $S_1^0 - S_1$ может лишь уменьшаться.

Условия (22.2) и (22.3), очевидно, не могут быть нарушены при переходе к другой системе отсчета, если сигнал распространяется со скоростью, меньшей скорости света. Однако в случае сверхсветового сигнала возможно выбрать такую систему отсчета [см. § 20, формула (20.6)], в которой поглощение сигнала приемником окажется более ранним событием, чем его излучение излучателем. Иначе говоря, приемник превратится в «излучатель», а излучатель — в «приемник». Однако при этом излучение будет происходить с увеличением негаэнтропии «излучателя», а поглощение — с уменьшением негаэнтропии «приемника», т. е. «сигнал» будет переносить негаэнтропию в обратном направлении.

Согласно предыдущему, такой процесс противоречит условиям (22.1) — (22.3). При помощи такого процесса может быть нарушено второе начало термодинамики, ибо спонтанное излучение с увеличением негаэнтропии излучателя могло бы быть использовано для построения вечного двигателя второго рода.

Итак, мы видим, что *сверхсветовые сигналы действительно запрещаются вторым началом термодинамики*. Однако этот запрет не является абсолютным, ибо возможны нарушения второго начала во флуктуациях. Следовательно, локализованные возмущения, переносящие энтропию со сверхсветовой скоростью, допускаются так же, как любые процессы флуктуационного порядка.

С этой точки зрения разрешаются воздействия, перемещающиеся со сверхсветовой скоростью, однако они не могут быть «сигналами», переносящими негаэнтропию. Такие сверхсветовые процессы не могут быть возбуждены по нашему произволу из одной точки и поглощены в другой удаленной точке, ибо систематическое воспроизведение подобных процессов означало бы систематическое (не флуктуационное) нарушение второго начала.

Вторым началом термодинамики запрещаются лишь макроскопического типа процессы, распространяющиеся со сверхсветовой скоростью, при помощи которых возможно передавать воздействия, способные вызывать макроскопические необратимые процессы. К такого рода процессам относится, например, процесс распространения упругого возмущения в твердом теле или, иначе, процесс распространения звука. Следовательно, скорость звука в твердом теле не может превышать скорость света. Это означает, что *не существует абсолютно твердых макроскопических тел*, ибо в абсолютно твердом теле скорость звука должна была бы стремиться к бесконечности, так как абсолютно твердое тело, по определению, не может деформироваться и должно перемещаться только как целое. При помощи абсолютно твердого тела можно было бы передавать макроскопические воздействия мгновенно из одной точки в другую, пространственно удаленную от первой. Последнее, однако, запрещается вторым началом термодинамики и поэтому неосуществимо в макромире, т. е. абсолютно твердые макроскопические тела неосуществимы.

§ 23. Возможны ли сверхсветовые возмущения внутри элементарных частиц?

Сделанный в предыдущем параграфе вывод о невозможности абсолютно твердых макроскопических тел часто переносится на все без исключения физические объекты, в частности на элементарные частицы. Утверждается, что элементарные частицы не могут иметь пространственных размеров, так как элементарность означает неделимость, а неделимым может быть лишь объект, подобный абсолютно твердому телу, внутри которого возмущения передаются со сверхсветовой скоростью Следовательно, как утверждают некоторые авторы, элементарные частицы должны быть точечными, так как в противном случае внутри элементарных частиц протекали бы процессы со сверхсветовой скоростью, что якобы запрещается доказанной выше теоремой.

Нетрудно видеть, что этот вывод делается лишь на основании гипотезы, что возмущения, распространяющиеся внутри элементарных частиц, имеют характер макроскопических сигналов, т. е. что к процессам, протекающим внутри элементарных частиц, применим «принцип причинности», в его узкой формулировке, приведенной в конце § 20, или что внутри элементарных частиц второе начало термодинамики имеет абсолютный характер.

Более вероятной гипотезой является, однако, допущение внутри элементарных частиц весьма тесной связи пространственно разобщенных событий. Эта связь может быть настолько тесной, что пространственно разобщенные события можно рассматривать лишь как причинно взаимообусловленные, но не разделяемые на причины, находящиеся в одной точке, и следствия в другой. В этом случае изменение временнóй последовательности событий не означает нарушения причинной взаимообусловленности и принципа причинности в более широком понимании. Нарушается лишь некоторое правило, называемое «принципом причинности», являющееся следствием закона возрастания энтропии. Однако нет никаких оснований требовать, чтобы закон возрастания энтропии выполнялся для всех процессов внутри элементарных частиц. Наоборот, в силу сугубой макроскопичности второго начала термодинамики есть все основания считать, что процессы внутри элементарных частиц протекают в основном как флуктуационные, т. е. как идущие с нарушением второго начала.

Итак, более широкий подход к содержанию теории относительности и принципу причинности делает не обязательным вывод о точечности элементарных частиц. При таком подходе элементарные частицы можно представлять как имеющие пространственную структуру, причем внутри частиц физические возмущения распространяются со сверхсветовыми скоростями.

Предположение о нарушении принципа причинности на интервалах порядка 10^{-13} *см* делалось и ранее (Блохинцев, Гайзенберг и др.). Однако против такого предположения выдвигалось серьезное возражение, состоящее в том, что нарушение причинности на интервале порядка 10^{-13} *см* посредством выбора системы отсчета может быть превращено в нарушение причинности на сколь угодно большом интервале, т. е. допущение нарушения причинности в микромире приводит к допущению его нарушения и в макромире.

Рассмотрим это доказательство подробнее. Пусть два взаимообусловленных события x_1, t_1 и x_2, t_2 разделены пространственноподобным интервалом, т. е.

$$s^2_{12} = c^2 (t_2 - t_1)^2 - (x_2 - x_1)^2 < 0 \tag{23.1}$$

или, что то же самое,

$$u = \frac{(x_2 - x_1)}{(t_2 - t_1)} > c. \tag{23.2}$$

Согласно преобразованиям Лоренца, в другой системе отсчета

$$x'_2 - x'_1 = \frac{(x_2 - x_1) - v (t_2 - t_1)}{\sqrt{1 - v^2 / c^2}} = (x_2 - x_1) \frac{1 - v / u}{\sqrt{1 - v^2 / c^2}} . \tag{23.3}$$

Согласно этой формуле при $u \gg c$, очевидно:

$$x'_2 - x'_1 = \frac{x_2 - x_1}{\sqrt{1 - v^2 / c^2}} . \tag{23.4}$$

Следовательно, если между точками x_1 и x_2 возникает процесс, нарушающий «принцип причинности», то в системе Σ' этот процесс возникает между точками x'_1 и x'_2, причем расстояние $l' = x'_2 - x'_1$ при $v \to c$ может быть сделано сколь угодно превышающим расстояние $l = x_2 - x_1$.

Заметим, однако, что вероятность нарушения причинности при этом уменьшается в такое же число раз, в какое увеличивается расстояние между событиями. Действительно, если рассматриваемые «беспричинные» события возникают в системе Σ со средней частотой ν, то в системе Σ' в силу эффекта замедления движущихся часов эти же события будут происходить с частотой

$$\nu' = \nu \sqrt{1 - v^2 / c^2}. \tag{23.5}$$

Следовательно,

$$l'\nu' = l\nu, \tag{23.6}$$

т. е. частота появления «беспричинных» событий уменьшается во столько же раз, во сколько увеличивается расстояние. А это означает, что и вероятность встретить такое событие уменьшается в такое же число раз.

Таким образом, хотя нарушение причинности и может быть сделано макроскопическим эффектом, однако вероятность такого нарушения становится исчезающе малой. Это хорошо согласуется с термодинамическим истолкованием «принципа причинности», так как со статистической точки зрения возможны сколь угодно большие нарушения второго начала термодинамики, однако вероятность таких флуктуаций становится тем меньшей, чем в бóльших масштабах имеет место отклонение от второго начала.

Итак, при термодинамическом истолковании «принципа причинности» передача энергии от точки к точке со сверхсветовой скоростью не запрещается абсолютно, но допускается как процесс флуктуационного типа.

Частицы, движущиеся со сверхсветовой скоростью, также становятся физически допустимыми объектами, не могущими, однако, излучаться по произволу и поглощаться, возбуждая необратимые макроскопические процессы.

Согласно (14.3) и (14.17) для любых частиц

$$\mathbf{P} = \frac{E}{c^2}\mathbf{u}, \quad c^2M^2 = E^2 / c^2 - \mathbf{P}^2; \tag{23.7}$$

поэтому для частиц, движущихся со сверхсветовой скоростью $u > c$,

$$M^2 < 0, \tag{23.8}$$

т. е. собственная масса мнимая.

Следовательно, мы пришли к заключению, что физически допустимо существование частиц *с мнимой собственной массой*, движущихся со скоростью, большей скорости света.

VI

ОТРИЦАТЕЛЬНЫЕ И МНИМЫЕ СОБСТВЕННЫЕ МАССЫ

§ 24. Определение собственной массы

В конце предыдущего параграфа было показано, что аппарат теории относительности допускает частицы, движущиеся со скоростью, большей скорости света, и имеющие мнимую собственную массу. Такие частицы запрещаются лишь принципом причинности. Но этот запрет снимается, если принцип причинности рассматривать не как некоторый абсолютный физический закон, а как следствие второго начала термодинамики [9, 10].

Теория относительности допускает также частицы отрицательной массы. Это было показано еще Дираком [11], в связи с релятивистской квантовой теорией электрона. Такие частицы обычно считаются физически не допустимыми, исходя из дополнительных соображений о ненарушимости общих положений термодинамики и принципа причинности, т. е. из тех же соображений, на основании которых запрещаются и сверхсветовые частицы мнимой массы.

Однако в гл. V было показано, что дополнительные соображения, позаимствованные из принципа причинности или из термодинамики, не могут рассматриваться как приводящие к абсолютным запретам, поэтому гипотеза существования частиц отрицательной и мнимой масс может считаться теоретически оправданной.

Чтобы выяснить перспективы, возникающие в связи с этой гипотезой, рассмотрим вытекающие из нее следствия. Исследуем сперва те общие свойства частиц отрицательной и мнимой массы, которые получаются из общих положений релятивистской динамики.

В соответствии с (14.17) собственная масса M определяется как инвариантная длина четырехмерного импульса P_k, т. е. из соотношения

$$c^2M^2 = P_\alpha P_\alpha = \frac{E^2}{c^2} - \mathbf{P}^2. \qquad (24{\cdot}1)$$

Заметим, что в дальнейшем для сокращения вместо термина «собственная масса» мы часто будем писать просто «масса».

Для системы частиц, очевидно,

$$E = cP_0 = \sum_k e_k, \; \mathbf{P} = \sum_k \mathbf{p}_k, \qquad (24.2)$$

где e_k и $\mathbf{p}_k$ — энергия и импульс отдельной частицы с собственной массой, определяемой из выражения $c^2m_k^2 = e^2/c^2 - \mathbf{p}_k^2$. В обычном случае, когда $M^2 > 0$, знак массы M считается совпадающим со знаком энергии E.

Согласно (24.1) собственная масса действительна, если вектор P_k времениподобен (т. е. $M^2 > 0$), и мнима, если этот вектор пространственноподобен (т. е. $M^2 < 0$). Случай нулевой собственной массы ($M = 0$) можно рассматривать как частный случай действительной собственной массы.

Таким образом, если компоненты вектора P_k считать произвольными действительными числами, то формула (24.1) допускает три качественно различные физические системы:

1) системы с положительной собственной массой, т. е. $M^2 \geqslant 0$, $E > 0$;

2) системы с отрицательной собственной массой, т. е. $M^2 \geqslant 0$, $E < 0$;

3) системы с мнимой собственной массой, т. е. $M^2 < 0$.

Следовательно, аппарат релятивистской динамики допускает три типа качественно различных систем, из которых только системы первого класса считаются физически реальными. Системы же второго и третьего классов запрещаются принципом причинности или общими положениями термодинамики. К системам второго класса, очевидно, надо отнести дираковские антиэлектроны отрицательной массы, а к системам третьего класса—виртуальные частицы квантовой теории поля. И те и другие считаются реально не существующими. Однако в рамках релятивистской кинематики и динамики никаких оснований для исключения таких частиц не имеется. Принцип же причинности, используемый для запретов этих частиц, имеет существенно иную макроскопическую природу и не вытекает из релятивистской кинематики и динамики.

§ 25. Общие свойства систем частиц положительной, отрицательной и мнимой массы

Для исследования общих свойств систем частиц всех трех классов удобно пользоваться диаграммой $(P, E/c)$. приведенной на рис. 13, на котором изображены

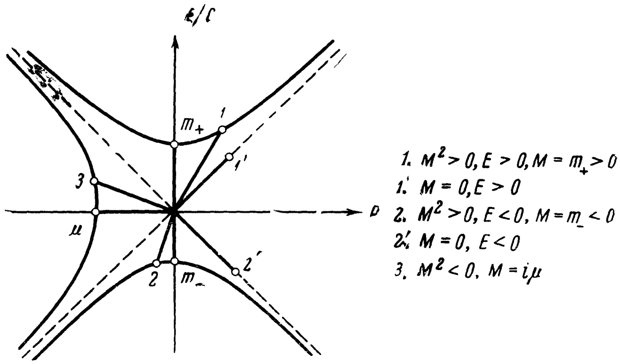

Рис. 13

частицы всех трех классов, включая два подкласса частиц нулевой собственной массы с положительной (*1*) и отрицательной (*2*) энергией. Абсолютные значения масс m_+, m_- и μ находятся по точкам пересечения с осями E/c или P гипербол, проведенных через точки, изображающие соответствующие частицы.

Для более полной характеристики систем введем также понятие средней скорости системы

$$\mathbf{V} = \frac{\mathbf{P}}{E} c^2. \tag{25.1}$$

Очевидно, $|\mathbf{V}| \leqslant c$ для систем первого и второго классов и $|\mathbf{V}| > c$ для систем третьего класса. Если система состоит из одной частицы, то средняя скорость $\mathbf{V}$ совпадает с истинной скоростью частицы $\mathbf{v}$.

Отдельные возможные частицы классифицируются, очевидно, точно так же, как и системы, однако класс совокупности частиц может отличаться от классов отдельных частиц, составляющих эту совокупность.

В силу (24.2) общий вектор $(\mathbf{P}, E/c)$ является простой геометрической суммой векторов $(\mathbf{p}_k, e_k/c)$, а инвариантная длина вектора (M) определяется посредством «проектирования» конца вектора по соответствующей гиперболе на оси E/c или P. Следовательно, достаточно воспользоваться диаграммой (см. рис. 13) для того, чтобы доказать следующие теоремы о собственной массе системы частиц:

I. если все $m_k > 0$, то $M > 0$;

II. если все $m_k = 0$, но $e_k > 0$, то $M \geqslant 0$ $(E > 0)$;

III. если все $m_k < 0$, то $M < 0$;

IV. если все $m_k = 0$, но $e_k < 0$, то $M \leqslant 0$ $(E < 0)$;

V. если все $m_k^2 < 0$, то M может быть любой ($M^2 \geqslant 0$, $M^2 < 0$).

Иначе говоря, совокупность частиц первого класса может образовать только систему первого класса, совокупность частиц второго класса — систему второго класса и совокупность частиц третьего класса — систему любого класса.

Тем же путем легко доказать следующие теоремы для систем, содержащих частицы разных классов:

VI. совокупность частиц, состоящая из более чем одной частицы первого класса и более чем одной частицы второго класса, может образовать систему любого класса;

VII. совокупность двух частиц положительной и отрицательной масс (m_+ и m_-) может образовать систему с общей массой $M > 0$ или $M^2 < 0$ (т. е. первого и третьего классов), если $m_+ + m_- > 0$; с общей массой $M < 0$

или $M^2 < 0$ (т. е. второго и третьего классов), если $m_+ + m_- < 0$, и с общей массой $M = 0$ или $M^2 < 0$, если $m_+ + m_- = 0$;

VIII. совокупность любого числа частиц первого и третьего классов может образовать систему первого или третьего классов и аналогично частицы второго и третьего классов могут образовать лишь систему второго или третьего класса.

Приведенные выше теоремы иллюстрируют тесную связь, существующую между частицами положительной, отрицательной и мнимой собственной массы.

Все известные нам экспериментально обнаруженные частицы принадлежат к первому классу, включая подкласс 1′ (фотоны), т. е. имеют $M^2 \geqslant 0$ при $E > 0$. Системы таких частиц согласно I и II также принадлежат к этому же классу. Однако достаточно допустить существование хотя бы двух частиц второго класса с отрицательной собственной массой и тогда, согласно VI мы будем обязаны считать существующими все три класса собственных масс. Следовательно, вводя отрицательные массы, мы тем самым вводим и мнимые собственные массы.

Изолированные частицы действительной массы, согласно (25.1) могут иметь лишь скорости, меньшие скорости света, частицы нулевой массы — лишь скорость, равную скорости света, и частицы мнимой массы — лишь скорость, бóльшую скорости света. Как, однако, понимать, что согласно (25.1) и VI совокупности частиц положительной и отрицательной массы могут иметь общую мнимую собственную массу и, следовательно, среднюю скорость, превышающую скорость света. Ведь каждая из частиц, составляющих систему, движется со скоростью, меньшей скорости света?

Для выяснения этого вопроса рассмотрим систему, состоящую из двух частиц положительной и отрицательной масс, причем положим $m_+ = |m_-| = m$. Пусть эти частицы движутся вдоль линии x с одинаковой по абсолютной величине скоростью v, но в противоположные стороны. Для такой системы, очевидно, $E = 0$, $P = 2mv$ и, следовательно, $M^2 = -4m^2v^2 < 0$, $V = \pm\infty$. Допустим, что обе частицы *одновременно* поглотились некоторыми телами A и B, расположенными в точках $x = +a$ и $x = -a$. В силу того, что поглощение частицы положительной массы означает увеличение энергии тела A, а поглощение частицы отрицательной массы — уменьшение энергии тела B, этот процесс эквивалентен испусканию некоторой энергии телом B и поглощению такой же энергии телом A. Поскольку оба эти процесса происходят одновременно, постольку весь процесс в целом может рассматриваться как передача энергии от тела B к телу A с бесконечно большой скоростью. Путем выбора системы отсчета эта скорость может быть сделана конечной, однако в любой системе она будет превышать скорость света в соответствии с теоремой, доказанной нами в конце § 13.

Итак, средняя скорость V может пониматься как скорость передачи энергии, определяемой релятивистски. Ясно, однако, что рассмотренный выше процесс не может быть использован для передачи сигнала (т. е. информации) из точки B в точку A, так как события A и B не могут находиться в причинно-следственной связи. Следовательно, сигнал не может быть передан со скоростью, большей скорости света, хотя и возможна передача энергии с такой скоростью.

§ 26. Нарушаются ли законы сохранения энергии и импульса для сверхсветовых частиц?

В соответствии с преобразованиями Лоренца и вытекающими из них формулами преобразования четырехмерной скорости (13.8) компоненты четырехмерного импульса преобразуются как

$$P_1' = \frac{P_1 - \beta P_0}{\sqrt{1-\beta^2}},\quad P_2' = P_2,\quad P_3' = P_3,\quad P_0' = \frac{P_0 - \beta P_1}{\sqrt{1-\beta^2}}, \tag{26.1}$$

где $\beta = v/c$. Для частицы, движущейся вдоль оси x,

$$P' = P\,\frac{1 - v/u}{\sqrt{1 - v^2/c^2}},\quad E' = E\,\frac{1 - vu/c^2}{\sqrt{1 - v^2/c^2}}, \tag{26.2}$$

где

$$u = c^2 \frac{P}{E} \tag{26.3}$$

— скорость частицы.

Согласно этим формулам для частиц первого и второго классов энергия не изменяет знака, а импульс может изменить знак при переходе в другую систему отсчета, поскольку для таких частиц $u < c$. Для частиц же третьего класса $u > c$, и поэтому импульс во всех системах отсчета имеет один и тот же знак, а энергия изменяет знак, когда относительная скорость v достигнет значения

$$v = c^2/u. \tag{26.4}$$

При этой скорости происходит также изменение знака трехмерной скорости, отсчитываемой в системе Σ', как это видно из формул преобразования скоростей (13.16), согласно которым

$$u' = \frac{u - v}{1 - vu/c^2}. \tag{26.5}$$

Таким образом, для сверхсветовых частиц направление трехмерной скорости и направление трехмерного

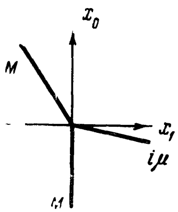

Рис. 14

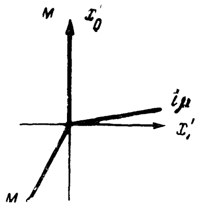

Рис. 15

импульса, определяемого формулами преобразования (26.2), совпадают не во всех системах отсчета. Но это означает, что формулы (26.1), (26.2) приводят к парадоксу, ибо объективный факт совпадения направления трехмерной скорости и трехмерного импульса или несовпадения этих направлений зависит от выбора системы отсчета, т. е. от способа изображения пространства и времени, избранного физиком. Последнее становится еще более парадоксальным, если мы заметим, что формулы преобразования (26.2) приводят к несохранению энергии и импульса в процессах испускания и поглощения сверхсветовых частиц в некоторых системах отсчета

Рассмотрим простейший процесс поглощения частицы мнимой массы $i\mu$ частицей действительной массы M (рис. 14). Известно, что свободная частица неизменной действительной массы M не может поглотить или испустить другую частицу действительной же массы. Подобный процесс, однако, разрешен формулами релятивистской динамики для поглощения или испускания частиц мнимой массы. Действительно, в силу законов сохранения общая масса системы M сохраняется. Следовательно, для сталкивающихся частиц массы M_1 и M_2, превращающихся в частицу массы M, имеем

$$c^2M^2 = \left(\frac{E_1}{c} + \frac{E_2}{c}\right)^2 - (\mathbf{P}_1 + \mathbf{P}_2)^2 =$$

$$= c^2M_1^2 + c^2M_2^2 + 2\left(\frac{E_1}{c}\frac{E_2}{c} - \mathbf{P}_1\mathbf{P}_2\right). \tag{26.6}$$

Если до столкновения частица M_1 покоилась, то $P_1 = 0$, $E_1 = M_1c^2$, откуда согласно (26.6)

$$E_2 = -\frac{c^2}{2M_1}[M_2^2 + M_1^2 - M^2]. \tag{26.7}$$

В рассматриваемом процессе первая частица действительной массы не изменяет свою собственную массу в результате поглощения второй частицы мнимой массы, т. е. $M_1 = M$, откуда

$$E_2 = -\frac{c^2}{2M}M_2^2. \tag{26.8}$$

С другой стороны, для этого процесса

$$E_1 + E_2 = E > E_1 = Mc^2, \tag{26.9}$$

где E — энергия первой частицы после поглощения второй. Следовательно,

$$E_2 > 0. \tag{26.10}$$

Но последнее неравенство может быть удовлетворено согласно (26.8) только при $M_2^2 < 0$, т. е. для поглощаемой частицы мнимой массы. В случае же частицы действительной массы это неравенство может быть удовлетворено только в случае

$$M^2 > M_1^2 + M_2^2 > M_1^2, \tag{26.11}$$

т. е. при $M_1 \neq M$.

Рассмотрим этот же процесс в другой системе координат, в которой первая частица (действительной массы) неподвижна после процесса взаимодействия со второй частицей (мнимой массы) (рис. 15). В этой системе отсчета частица мнимой массы $i\mu$ уже не поглощается частицей действительной массы M, а испускается ею. Поэтому вместо неравенства (26.9) надо написать

$$E_1 = Mc^2 + E_2' > Mc^2, \tag{26.12}$$

где E_2' — энергия испущенной частицы мнимой массы. Следовательно,

$$E_2' > 0. \tag{26.13}$$

Итак, в силу закона сохранения энергии энергия частицы мнимой массы должна быть положительной в обеих системах отсчета. Однако в соответствии с формулами (26.2) и (26.5)

$$E_2' = -E_2, \tag{26.14}$$

поскольку

$$u_2' = -u_2, \tag{26.15}$$

как это очевидно из рис. 14 и 15. Но (26.14) явно противоречит неравенствам (26.10) и (26.13), полученным исходя из закона сохранения энергии. Следовательно, преобразования (26.2), примененные к сверхсветовой частице, приводят к выражению, противоречащему закону сохранения энергии в случае, когда трехмерная скорость в результате этих преобразований изменяет знак. Нетрудно также показать, что в этом случае имеет место и нарушение закона сохранения импульса. Это видно из сравнения рис. 14 и 15. А на рис. 14 импульс частицы мнимой массы направлен влево. На рис. 15 импульс этой частицы направлен вправо. Однако соответственно формулам (26.2) импульс не меняет знака, откуда следует, что эти формулы в данном случае не совместимы с законом сохранения импульса.

Это противоречие между формулами преобразования компонент четырехмерного импульса и законами сохранения может быть снято, если правую часть формул (26.2) умножить на минус единицу в тех случаях, когда

$$vu > c^2. \tag{26.16}$$

Иначе говоря, с законами сохранения энергии и импульса совместимы не преобразования (26.2), а формулы

$$P' = u'\frac{E'}{c^2}, \quad E' = E\frac{|1 - vu/c^2|}{\sqrt{1 - v^2/c^2}}, \quad u' = \frac{u - v}{1 - vu/c^2}. \tag{26.17}$$

Эти формулы, очевидно, совпадают с формулами (26.2) при $vu < c^2$. Следовательно, формулы (26.17) тождественны (26.2) для любых частиц, имеющих действительную массу, так как для них $u < c$. Формулы (26.17) так же можно записать в более общем виде аналогично формулам (26.1):

$$P_1' = \gamma \frac{P_1 - \beta P_0}{\sqrt{1-\beta^2}}, \quad P_2' = \gamma P_2, \quad P_3' = \gamma P_3,$$

$$P_0' = \gamma \frac{P_0 - \beta P_1}{\sqrt{1-\beta^2}}, \quad \gamma = \frac{1 - \beta \frac{P_1}{P_0}}{\left|1 - \beta \frac{P_1}{P_0}\right|}. \tag{26.18}$$

Очевидно, для сверхсветовых частиц четырехмерный импульс необходимо определить как вектор, преобразующийся по формулам (26.17) и (26.18), так как только в этом случае не будут нарушаться законы сохранения энергии и импульса при операции перехода к новой системе отсчета, имеющей скорость, удовлетворяющую условию (26.16).

В соответствии с формулами (26.17) и (26.18) энергия частиц третьего класса не изменяет знака ни при каких лоренцовых преобразованиях. Следовательно, частицы третьего класса могут быть разделены на два подкласса: на частицы, имеющие положительную энергию, и частицы, имеющие отрицательную энергию. Преобразованиями Лоренца невозможно превратить частицы одного из этих двух подклассов в частицы другого подкласса.

§ 27. Частицы отрицательной массы в гравитационном поле

Движение частиц отрицательной массы, очевидно, определяется уравнениями Минковского (14.1) — (14.4). Следовательно, ускорение частиц отрицательной массы направлено противоположно действующей силе F_k, так как согласно (14.1).

$$\frac{d^2 x_k}{d\tau^2} = \frac{dU_k}{d\tau} = \frac{F_k}{M}. \tag{27.1}$$

Таким образом, во внешнем электрическом поле одинаково заряженные частицы отрицательной и положительной масс будут ускоряться в противоположных направлениях. Однако во внешнем гравитационном поле минус-частицы * должны ускоряться в том же направлении, в котором ускоряются плюс-частицы, так как в противном случае минус-частицы нарушали бы принцип эквивалентности, согласно которому все без исключения объекты в одном и том же гравитационном поле приобретают одно и то же ускорение. Следовательно, исходя из принципа эквивалентности, мы должны считать, что во внешнем гравитационном поле на минус-частицу массы $-m$ действует сила $-\mathbf{F}$, равная по абсолютному значению, но противоположная по направлению силе $\mathbf{F}$, действующей на плюс-частицу массы m. Иначе говоря, мы должны считать, что минус-частицы, так же как и плюс-частицы, имеют «гравитационный заряд» **, определяемый одним и тем же выражением

$$e = \sqrt{-\varkappa}\, M, \qquad (27.2)$$

где $\varkappa$ — гравитационная постоянная.

Итак, обычные большие гравитирующие массы (Земля, Солнце, звезды, галактики и т. п.) притягивают как плюс-частицы, так и минус-частицы, сообщая им одно и то же ускорение.

Иначе будут вести себя плюс- и минус-частицы в гравитационном поле, создаваемом большими отрицательными массами. Они будут отталкиваться от этих масс, поскольку гравитационное поле отрицательных масс направлено противоположно полю положительных масс, так как оно может рассматриваться как создаваемое «гравитационным зарядом» противоположного знака.

Очевидно, гравитационное взаимодействие системы частиц отрицательной массы приведет к эффекту, отличному от эффекта гравитационного взаимодействия частиц

---

* В дальнейшем для сокращения будем называть плюс-частицами обычные частицы положительной массы, а минус-частицами — частицы отрицательной массы.

** В нерелятивистской области гравитационное поле определяется законом тяготения Ньютона и может рассматриваться по аналогии с электрическим полем, определяемым законом Кулона, как создаваемое «гравитационным зарядом», определяемым выражением (27.2).

положительной массы. Действительно, гравитационные силы стягивают одну к другой плюс-частицы. Но минус-частицы в результате гравитационного взаимодействия будут отталкиваться одна от другой. Таким образом, в пространстве, заполненном плюс-частицами, должны образовываться сгустки вещества, и оно будет распределяться с весьма неравномерной плотностью, как это и имеет место в реальной вселенной, заполненной положительными массами. Если же пространство заполнено минус-частицами, то расталкивающие силы приведут к распределению с равномерной плотностью.

Допуская существование минус-частиц, возможно рассмотреть модель вселенной, имеющей в среднем нулевую собственную массу. В реальной вселенной средняя плотность массы плюс-частиц чрезвычайно мала, составляя приблизительно 10^{-30} *г/см*^{3}, т. е. примерно один протон в кубическом метре. При этом межгалактическое пространство заполнено средой с несколько меньшей плотностью, а бо́льшие значения, достигающие 10^5 и даже 10^9 *г/см*^{3}, плотность вещества имеет лишь внутри звезд, занимающих относительно ничтожный объем космического пространства. Если считать, что кроме плюс-частиц существуют минус-частицы и полагать, что их средняя плотность приблизительно совпадает со средней плотностью плюс-частиц, то, учитывая взаимное расталкивание минус-частиц, мы должны положить, что и реальная плотность минус-частиц в любой точке вселенной имеет приблизительно такой же порядок. Но эта плотность очень мала и этим может объясняться то обстоятельство, что минус-частицы мы не обнаруживаем в земных условиях и не замечаем производимых ими воздействий в окружающей среде, состоящей из плюс-частиц. Однако в галактических масштабах наличие фона минус-частиц может привести к заметным эффектам. Так, например, распределенные с приблизительно постоянной плотностью ρ минус-частицы должны практически полностью компенсировать гравитационное поле отдельной галактики массы M на расстоянии порядка

$$R = (3M/4\pi\rho)^{1/3}, \qquad (27.3)$$

так как масса минус-частиц, заключенная в сфере радиуса R, по абсолютной величине равна массе галактики, находящейся в центре этой сферы.

Более детально эту гравитационную экранировку можно учесть, считая, что минус-частицы находятся в равновесии при отрицательной температуре $T < 0$ и распределены в поле большой плюс-массы по формуле Больцмана

$$n^- = n_0^- e^{-m^- \varphi/kT}, \tag{27.4}$$

где n^- — концентрация минус-частиц; n_0^- — средняя концентрация минус-частиц в метагалактике; m^- — масса минус-частиц; φ — потенциал гравитационного поля, определяемый в ньютоновском приближении уравнением

$$\nabla^2 \varphi = 4\pi\varkappa\,(m^+ n^+ + m^- n^-). \tag{27.5}$$

Для вселенной с нулевой средней массой, очевидно,

$$m^+ n_0^+ + m^- n_0^- = 0. \tag{27.6}$$

Замечая далее, что плотность свободных плюс-масс, создающих гравитационное поле, равна

$$\rho = m^+ (n^+ - n_0^+), \tag{27.7}$$

и полагая

$$m^- \varphi/kT \ll 1, \tag{27.8}$$

получаем для φ следующее уравнение:

$$\nabla^2 \varphi - \frac{1}{\Lambda^2}\varphi = 4\pi\varkappa\rho, \tag{27.9}$$

где

$$\Lambda = \sqrt{\frac{k\,|T|}{4\pi\varkappa\,|m^-|^2\, n_0^-}}\,. \tag{27.10}$$

Таким образом, для ньютоновского гравитационного потенциала вместо уравнения Пуассона получаем уравнение Неймана с космологической постоянной Λ, зависящей от температуры минус-частиц и их средней плотности.

На расстояниях, превышающих Λ, происходит экранировка гравитационного поля любых объектов, состоящих из плюс-частиц. Отдельные галактики, имеющие размеры, меньшие Λ, на расстояниях, превышающих Λ, не будут притягиваться одна к другой. Если же галактика имеет размеры, превышающие Λ, то переферические ее части не будут удерживаться притяжением центральных частей, что должно привести к выбросам материи из такой галактики.

Полагая $n_0^- = 10^{-6}$ *см*$^{-3}$, $T = 1°$ K, получим для минус-частиц с массой нуклона по формуле (27.10): $\Lambda = 10^{21}$ *см.*

Полезно также заметить, что в релятивистской космологии модель вселенной с нулевой средней плотностью массы представляет особый интерес.

§ 28. Отрицательные массы и принцип причинности

Из всех известных возражений против допущения частиц отрицательной массы наиболее существенным является возражение с точки зрения принципа причинности или, что то же самое, с точки зрения второго начала термодинамики (см. [9, 10]). Все другие возражения либо сводятся к нему, либо в какой-то мере с ним связаны. Проанализируем это главное возражение.

Допустим, что частицы отрицательной массы могут испускаться или поглощаться системами обычных частиц, так же как, например, фотоны или π^0-мезоны. Пусть минус-частица, испущенная системой A в момент времени t_1, поглотилась системой B в момент t_2, причем, очевидно, $t_2 > t_1$. Однако *испускание* минус-частицы означает увеличение энергии и импульса системы A, точно такое же, какое было бы вызвано *поглощением* плюс-частицы той же (по абсолютному значению) массы и скорости. И, аналогично, *поглощение* минус-частицы системой B эквивалентно *испусканию* этой системой плюс-частицы.

Таким образом, с макроскопической точки зрения испускание минус-частицы системой A и поглощение ее системой B эквивалентно процессу испускания плюс частицы системой B и ее поглощения системой A, причем поглощение в последнем процессе происходит *раньше* испускания, т. е. с *нарушением принципа причинности.*

С термодинамической точки зрения последнее означает, что процесс идет с нарушением второго начала термодинамики. Действительно, излучение плюс-частиц — это процесс, идущий с уменьшением негаэнтропии излучателя. При этом испущенные частицы переносят негаэнтропию (т. е. положительную информацию), а негаэнтропия поглощающей системы увеличивается в момент поглощения плюс-частицы (т. е. система переходит в возбужденное состояние, в результате чего в ней могут возникнуть необратимые процессы). Излучение же минус-частиц, очевидно, является процессом, идущим с увеличением негаэнтропии излучателя, а их поглощение ведет к уменьшению негаэнтропии поглотителя, при этом минус-частицы должны переносить отрицательную негаэнтропию (т. е. отрицательную информацию). Но излучение предшествует поглощению, т. е. увеличение негаэнтропии излучающего тела происходит само собой (тело как бы само собой возбуждается), т. е. происходит процесс, противоречащий второму началу термодинамики. Таким образом, физически менее определенное утверждение о нарушении принципа причинности заменяется утверждением о нарушении минус-частицами количественно строго определенного физического закона — второго начала термодинамики.

Обычно на основании рассуждений, аналогичных приведенным выше, делается категорическое заключение о невозможности существования частиц отрицательной массы *. Особенно «очевидным» представляется этот запрет, когда рассуждение опирается на принцип причинности, а не на второе начало термодинамики. Если, однако, признать, что принцип причинности является лишь интуитивно очевидным выражением второго начала, то «очевидность» запрета частиц отрицательной массы теряет физический фундамент.

Действительно, ведь второе начало — статистический закон, господствующий в окружающих нас макроскопических явлениях, причем отступление от него в отдельных флуктуациях допускается статистической механикой. Таким образом, допустимы явления, нарушающие второе начало в отдельных флуктуациях. Следовательно, второе начало не может запретить минус-частицы, если они, по каким-либо причинам, могут проявляться лишь в процессах флуктуационного типа.

Таким образом, ссылки на принцип причинности не могут рассматриваться как аргумент для абсолютного

* Таким же образом доказывается и невозможность существования частиц мнимой массы.

запрета минус-частиц. Необходимо, однако, выяснить, не нарушают ли минус-частицы закона термодинамики не только во флуктуациях, а более существенно, и если такое нарушение имеет место, то может ли оно быть оправдано физически?

§ 29. Отрицательные массы и законы термодинамики

Допуская частицы отрицательной массы, мы полагаем, что физические системы могут иметь как сколь угодно большие положительные энергии, так и сколь угодно малые, ничем не ограниченные снизу отрицательные энергии. Это свойство систем, содержащих минус-частицы, находится, однако, в противоречии с одной из исходных аксиом термодинамики — постулатом существования состояния термодинамического равновесия. Согласно этому постулату любая термодинамическая система должна иметь состояние термодинамического равновесия, т. е. такое состояние, в котором все макроскопические параметры неизменны и система не может выйти из этого состояния спонтанно, т. е. без дополнительных внешних воздействий. Однако такое состояние равновесия возможно не у всех физических систем. В термодинамике и статистической физике обычно рассматриваются лишь системы, имеющие ограниченный снизу энергетический спектр, т. е. минимальную энергию, которая достигается при нулевой температуре. Такие системы имеют состояние термодинамического равновесия. Системы же, не обладающие минимальной энергией, в соответствии с термодинамикой, стремясь к равновесию, должны переходить на все более и более низкие энергетические уровни, т. е. все время находиться в неравновесном состоянии. Согласно статистической физике системы с неограниченным энергетическим спектром, подобные системам, содержащим минус-частицы, также не обладают состоянием термодинамического равновесия, поскольку они не имеют сходящейся суммы состояний

$$Z = \sum_k e^{-E_k/\Theta}, \tag{29.1}$$

где $\Theta = kT$, модуль канонического распределения;

T — абсолютная температура; k — константа Больцмана; E_k — энергетический уровень системы.

Таким образом, допуская существование минус-частиц, мы тем самым соглашаемся с возможностью нарушения законов термодинамики не только во флуктуациях, но и в макроскопических масштабах. Можно, конечно, исходя из концепции абсолютной ненарушимости законов термодинамики, рассматривать обнаруженную выше несовместимость минус-частиц с постулатом существования термодинамического равновесия как аргумент, запрещающий минус-частицы. Однако более широкой является атомистическая концепция, лежащая в основе статистической физики, согласно которой законы термодинамики не имеют абсолютного характера, а являются лишь следствием статистической теории динамических систем Согласно этой концепции допустимы нарушения законов термодинамики не только во флуктуациях, но и систематически в макроскопических масштабах для особого типа динамических микросистем. Необходимо лишь обеспечить ненарушимость законов термодинамики в заметных масштабах для обычных макроскопических систем, встречающихся в повседневном, земном опыте.

Квазиравновесие систем, состоящих из плюс-частиц, при наличии в природе минус-частиц может быть, очевидно, обеспечено, если допустить, что минус-частицы взаимодействуют с плюс-частицами только гравитационно и что другие виды взаимодействия (сильное, электромагнитное, слабое) между плюс- и минус-частицами полностью отсутствуют. В этом случае любая система плюс-частиц может практически считаться адиабатически изолированной от минус-частиц и поэтому может рассматриваться как находящаяся в состоянии термодинамического равновесия. Равновесие может нарушаться лишь вследствие обмена гравитонами системы плюс-частиц с системой минус-частиц. Однако такие процессы крайне маловероятны и не могут привести к заметным эффектам.

Если, в соответствии со сделанным выше допущением, между минус-частицами и плюс-частицами осуществляется только гравитационное взаимодействие, то минус-частицы вообще смогут быть обнаружены лишь в явлениях космических масштабов по гравитационным полям, создаваемым большими скоплениями этих частиц.

Важно рассмотреть также другое, более естественное предположение, допускающее все известные виды взаимодействий (сильное, электромагнитное, слабое, гравитационное) между минус-частицами и плюс-частицами. Принимая это допущение и считая, что минус-частицы распределены во вселенной с очень малой плотностью (как это должно иметь место в соответствии с § 27), мы не придем к заметным нарушениям постулата термодинамического равновесия в земных условиях, если будем рассматривать только процессы взаимодействия плюс-частиц с минус-частицами. При концентрации 10^{-6} частиц на 1 $см^3$ минус-частицы, взаимодействуя с обычным веществом, будут вызывать в земных условиях не более заметные макроскопические эффекты, чем космические лучи, и, таким образом, нарушение постулата равновесия может быть лишь труднообнаружимым эффектом.

Гипотеза существования всех видов взаимодействия между плюс- и минус-частицами встречается, однако, с определенными трудностями, если учесть возможность процессов множественного рождения совокупностей плюс- и минус-частиц. Обычные рассуждения квантовой теории поля не запрещают процессов рождения двух или более плюс-частиц и одновременно двух или более минус-частиц прямо из вакуума, т. е. так сказать *из ничего*. Сам по себе этот процесс не является более странным, чем сами минус-частицы. Однако в соответствии с обычными рассуждениями квантовой теории поля, в рамках теории возмущений, вероятность такого процесса множественного рождения из вакуума возрастает с ростом числа рождаемых частиц *. Поэтому вероятность рождения объектов, сравнимых по числу частиц, с такими космическими объектами, как звезды или галактики, может стать настолько значительной, что подобные процессы окажутся очень частым явлением.

Однако это возражение против распространения всех видов взаимодействия на минус-частицы не является столь радикальным, как оно кажется на первый взгляд. Действительно, ведь это возражение можно рассматривать как указание на еще одну расходимость современной квантовой теории поля, т. е. как на еще одно свидетель-

* На это возражение против гипотезы существования всех видов взаимодействия между плюс- и минус-частицами обратил мое внимание Т. Д. Ли в 1964 г.

ство неудовлетворительности существующей теории элементарных частиц. Неудивительно, что внутреннее противоречие квантовой теории поля столь резко выявляется при введении именно минус-частиц, т. е. объектов, чуждых существующей схеме квантовой теории, построенной для положительных энергий и масс.

Обнаруженная расходимость легко устраняется, если, например, допустить, что энтропия мира минус-частиц не возрастает, а убывает, т. е. макроскопическое время минус-частиц течет в обратном направлении по отношению ко времени, регистрируемому обычными макроскопическими часами. В этом случае вероятность одновременного порождения минус-частиц в одной точке должна подсчитываться как вероятность их одновременного собирания в одну точку, т. е. как очень маловероятное событие. С учетом этого обстоятельства вероятность множественного образования совокупности N плюс-частиц и N минус-частиц не может отличаться от вероятности уничтожения такого же комплекса частиц в одном акте. Следовательно, процессы множественного образования плюс- и минус-частиц из вакуума не могут привести к каким-либо катастрофическим нарушениям равновесия в макроскопических масштабах. Ясно, однако, что гипотеза противоположного направления течения времени мира плюс-частиц и мира минус-частиц должна привести к радикальной перестройке квантовой теории поля.

Итак, не будем считать исключенной гипотезу о возможности всех известных видов взаимодействий между минус- и плюс-частицами. Принимая эту гипотезу, мы, очевидно, допускаем возможность макроскопических нарушений классической термодинамики и признаем необходимость построения новой термодинамики, в рамках которой классическая термодинамика должна быть лишь частным случаем, справедливым для узкого класса систем, имеющих ограниченный энергетический спектр.

Наибольшие затруднения, пожалуй, связаны с допущением электромагнитных взаимодействий минус-частиц с плюс-частицами. Минус-частица, имеющая электрический заряд, обладала бы весьма странными свойствами. Проходя через обычное вещество, состоящее из плюс-частиц, минус-частица в результате кулоновских взаимодействий с электронами и ядрами должна непрерывно терять энергию, т. е. ускоряться. При этом среда, через которую проходит минус-частица, будет нагреваться. Таким образом, имея достаточное число минус-частиц, можно было бы осуществить энергетический источник, который функционирует подобно термодинамической машине второго рода (запрещаемой классической термодинамикой), непрерывно производя работу за счет уменьшения энергии (охлаждения) теплового резервуара, состоящего из минус-частиц.

Очевидно, такого рода машина второго рода может быть осуществлена и с минус-частицами, взаимодействующими с обычным веществом посредством сильных или слабых взаимодействий. Однако в последнем случае процесс энерговыделения будет протекать весьма медленно вследствие малости поперечных сечений столкновений минус-частиц с плюс-частицами. В случае же электромагнитных взаимодействий передача энергии минус-частиц плюс-частицами может протекать весьма быстро, так как заряженная частица производит многочисленные ионизации пронизываемых ею атомов. Поскольку подобные процессы до сих пор не обнаруживались в повседневном опыте, постольку естественно предположить, что электрически заряженные минус-частицы либо вообще не существуют, либо их очень мало в окружающем пространстве и вероятность их порождения чрезвычайно мала. Что же касается минус-частиц, взаимодействующих с плюс-частицами посредством сильных и слабых взаимодействий, то вызываемое ими нарушение запрета термодинамической машины второго рода может оказаться заметным лишь в явлениях космических масштабов и практически необнаружимым в макроскопических земных экспериментах.

Как ни удивительны рассмотренные выше следствия гипотезы существования минус-частиц, взаимодействующих с обычным веществом, однако физики не вправе отрицать их возможность только на основании априорной убежденности в невозможности макроскопических нарушений классической термодинамики.

§ 30. Регистрация частиц отрицательной массы

Все известные приборы — регистраторы частиц — действуют примерно по следующей схеме: частица, попадающая извне в прибор, полностью или частично теряет в нем свою энергию; вместе с этой долей энергии частица сообщает прибору негаэнтропию, выводящую систему прибора из некоторого метастабильного состояния равновесия (или динамического квазиравновесия), в котором она при этом находилась; в результате этого возникает необратимый процесс лавинного типа, приводящий к макроскопически обнаружимым эффектам. По этой схеме действуют камера Вильсона, ионизационная камера, счетчик Гейгера, черенковский счетчик, фотоэмульсии и т. д. Очевидно, частицы отрицательной массы не могут привести в действие приборы, работающие по этой схеме, если эти частицы попадают извне в прибор и поглощаются или замедляются в нем. Действительно, при поглощении прибором минус-частица не сообщает ему энергию, а отнимает. Следовательно, она не может произвести того действия, которое производит плюс-частица при ее поглощении, т. е. она не может сообщить прибору негаэнтропию, необходимую для приведения его в действие.

Приборы обычного типа могут, очевидно, регистрировать не поглощение, а излучение минус-частиц из прибора, так как излученная минус-частица сообщит прибору такую же энергию, какую сообщила бы ему поглощенная плюс-частица. Будет регистрироваться также процесс, при котором скорость минус-частицы, попавшей в прибор, будет нарастать. Например, электрически заряженная минус-частица, проходя через камеру Вильсона, могла бы вызвать ионизацию центров конденсации, однако при каждом процессе ионизации она отдавала бы энергию, т. е. увеличивала скорость. След, оставленный такой минус-частицей, не отличался бы от следа плюс-частицы, двигавшейся в противоположном направлении. Следовательно, заряженная минус-частица, попавшая в камеру Вильсона снизу, т. е. вышедшая из Земли, оставляла бы точно такой же след, как плюс-частица, пришедшая сверху, т. е. из космоса. Поэтому заряженные минус-частицы, порожденные внутри Земли и затем ускоренные в процессе столкновений с плюс-частицами вещества Земли, по выходе из нее регистрировались бы камерой Вильсона точно так же, как заряженные космические плюс-частицы, приходящие из космоса и замедляемые веществом Земли.

Однако, как мы выяснили в предыдущем параграфе, подобные электрически заряженные минус-частицы либо вообще не могут существовать, либо вероятность их порождения по каким-то неизвестным пока причинам исчезающе мала, ибо, в противном случае, вызываемые ими нарушения термодинамики привели бы к заметным эффектам.

Очевидно, возможные процессы самоускорения незаряженных минус-частиц не смогут быть зарегистрированы приборами, подобными камере Вильсона. Лишь спонтанное испускание минус-частицы активной частью прибора типа счетчика может привести его в действие, т. е. быть зарегистрировано. Для улавливания же и регистрации уже существующих минус-частиц такого типа приборы не приспособлены.

Прибор, предназначенный для регистрации минус-частиц, должен приводиться в действие не сообщением ему энергии частицы, а в результате уменьшения энергии активной части прибора, т. е. в результате процесса, эквивалентного излучению плюс-частицы. Следовательно, такой прибор должен содержать систему, первоначально находящуюся на некотором промежуточном уровне (не самом низком), причем перейти на более низкий уровень эта система может лишь в результате поглощения минус-частицы, т. е. сообщения ей отрицательной энергии. После же перехода на более низкий уровень в системе должен начаться спонтанный необратимый макроскопический процесс, приводящий к акту «регистрации», т. е. система должна перейти в какое-то третье состояние, более устойчивое, чем оба предыдущие.

Подобные процессы могут иметь место в системе, находящейся в состоянии с отрицательной температурой. Действительно, в такой системе большей стабильностью обладают более высокие энергетические уровни, поскольку «возбуждение» такой системы означает переход на более низкий энергетический уровень, а не на более высокий, как в случае обычных систем с положительной температурой.

Итак, в принципе, представляется возможным создать приборы, регистрирующие поглощение минус-частиц.

§ 31. Частицы мнимой массы и второе начало термодинамики

Принято считать, что частицы мнимой собственной массы, движущиеся со скоростью, большей скорости света, не могут рассматриваться как реальные объекты, поскольку излучение и поглощение таких частиц переносило бы воздействие излучателя на поглотитель со скоростью, большей скорости света. Сигнализация же со скоростью, большей скорости света, противоречила бы принципу причинности, поскольку согласно теории относительности всегда возможно выбрать такую систему отсчета, в которой момент излучения (причина) окажется более поздним, чем момент поглощения сигнала (следствие), ибо эти события связаны пространственноподобным интервалом. Однако, как это было нами выяснено в § 22, подобное рассуждение приводит к абсолютному запрету частиц мнимой массы лишь в том случае, если принцип причинности рассматривать как абсолютный физический закон. Но с физической точки зрения принцип причинности является некоторым выражением или следствием второго начала термодинамики. Следовательно, применяя те же рассуждения, которые мы применяли в отношении частиц отрицательной массы, мы можем утверждать, что частицы мнимой массы запрещаются лишь как объекты, используемые для сигнализации, однако они могут появляться в процессах флуктуационного типа, не нарушая при этом второго начала термодинамики для макроскопических процессов систематического характера.

Легко видеть, что второе начало термодинамики не нарушается в рассмотренном выше процессе излучения и поглощения частицы мнимой массы, если вероятность процесса испускания равна вероятности процесса поглощения. В этом случае невозможно отличить излучатель сигналов от их приемника и вообще невозможна сигнализация, так как передача воздействия носит не систематический и не направленный характер.

Равенство вероятностей процесса поглощения и испускания, очевидно, имеет место, если пространство изотропно заполнено частицами мнимой массы, не имеющими каких-либо зарядов и поглощающихся и излучающихся обычными частицами положительной массы без изменения собственной массы последних *. В этом случае, как это видно, например, из рис. 14 и 15, процессы испускания и поглощения совершенно симметричны во времени и вообще тождественны, поскольку превращаются один в другой посредством перехода от одной системы отсчета к другой. Следовательно, как испускание, так и поглощение частиц мнимой массы в данном случае не сопровождаются изменением общей энтропии системы, в силу чего второе начало не нарушается.

Такого рода частицы мнимой массы не переносят негаэнтропию и не могут быть использованы для сигнализации, так как всякий сигнал должен переносить информацию, т. е негаэнтропию.

Несколько иная ситуация будет иметь место, если частица, излучающая частицу мнимой массы, изменяет при этом свою собственную массу, заряд, спин и т. д., т. е. превращается в другую частицу. В этом случае симметрия во времени нарушается и утверждение о равенстве вероятностей излучения и поглощения лишается основания.

Симметрия, однако, восстанавливается. если мы будем рассматривать не изолированные процессы излучения или поглощения частиц мнимой массы, а процессы, в которых частица мнимой массы излучается частицей A и поглощается частицей B так, что частица A превращается в частицу сорта B и частица B — в частицу сорта A. Но последний процесс физически совпадает с известным процессом обмена импульсом, зарядом и т. д. между двумя элементарными частицами посредством *виртуальной* частицы.

Таким образом, появляющиеся в квантовой теории элементарных частиц виртуальные частицы можно рассматривать как физически реальные частицы с мнимой собственной массой, которыми обмениваются обычные элементарные частицы. Введение таких частиц не нарушает второго начала термодинамики, и, следовательно, при помощи них невозможно нарушать макроскопический принцип причинности.

---

* В § 26 нами было показано, что возможен процесс, когда частица положительной массы излучает частицу мнимой массы, изменяя лишь свои энергию и импульс, но не массу покоя. Подобный процесс с испусканием частиц действительной массы невозможен.

В конце § 26 было показано, что частицы мнимой массы могут иметь либо положительную, либо отрицательную энергию. Таким образом, допущение частиц мнимой массы не связано с непременным нарушением законов термодинамики в макроскопических масштабах, как это имеет место в случае минус-частиц. Законы термодинамики не будут макроскопически нарушаться, если запретить частицы с отрицательной энергией и допустить лишь частицы мнимой массы с положительной энергией

§ 32. Возможна ли регистрация частиц мнимой массы?

Мы уже видели, что частицы мнимой массы не переносят негаэнтропию и поэтому не могут быть использованы в качестве сигналов. Представляется, таким образом, что они не могут быть вообще зарегистрированы и являются как бы ненаблюдаемыми объектами.

Однако на примере частиц отрицательной массы мы уже видели, что существуют объекты, которые не могут быть зарегистрированы обычными приборами, однако могут быть обнаружены при помощи принципиально новых измерительных устройств. Следует поэтому рассмотреть возможность существования особых измеряющих систем, способных регистрировать частицы мнимой массы.

Поскольку систематическая регистрация поглощения или излучения частиц мнимой массы приводила бы к нарушению второго начала термодинамики, постольку мы должны отказаться от возможности построить приборы, регистрирующие частицу мнимой массы в заданной точке. Это не означает, конечно, что мы вообще отказываемся от возможности обнаружить какие-либо действия частиц мнимой массы в заданной точке, ибо не запрещаются флуктуационные скопления частиц, идущие с нарушением второго начала термодинамики, которые могут привести в действие регистрирующие приборы обычного типа. Это не будет, однако, регистрацией частицы в обычном понимании.

Если исключаются приборы, регистрирующие частицу мнимой массы в заданной точке, то допустимы приборы, регистрирующие испускание частицы в одной точке и ее поглощение в другой точке, как одно единое явление. Так, например, если частица мнимой массы переносит электрический заряд, то процесс ее испускания частицей A и поглощения частицей B может быть зарегистрирован в фотоэмульсии по следу, оставленному частицей A до испускания частицы мнимой массы, и следу частицы B, образуемому после поглощения частицы мнимой массы. Иначе говоря, представляется возможным зарегистрировать процесс обмена зарядом между заряженной и нейтральной частицами при помощи частицы мнимой массы (т. е. процесс, который принято рассматривать как процесс обмена посредством виртуальной частицы).

Следовательно, в принципе, частицы мнимой массы могут быть экспериментально обнаружены, однако лишь при помощи специального типа приборов и особого типа экспериментов, регистрирующих одновременно процесс испускания и процесс поглощения такой частицы.

§ 33. Отрицательные массы и отрицательные температуры

Известно, что системы, имеющие как минимальную, так и максимальную конечные энергии, т. е. ограниченный как снизу, так и сверху энергетический спектр

$$E_{\min} < E_n < E_{\max}, \tag{33.1}$$

могут находиться в равновесном состоянии с отрицательной абсолютной температурой *. Действительно, сумма состояний (29.1) в случае ограниченного энергетического спектра сходится как при $\Theta > 0$, так и при $\Theta < 0$. Следовательно, у системы существует равновесное каноническое распределение

$$W(E_n) = Ae^{-E_n/\Theta} \tag{33.2}$$

как при $\Theta > 0$, так и при $\Theta < 0$, т. е. и при отрицательной температуре $T = \Theta/k$, так как в обоих случаях условие

* Точнее говоря, это состояние является квазиравновесным, так как никакая система не может быть абсолютно изолирована от окружающих тел, которые не могут находиться в равновесии при отрицательной температуре, ибо имеют неограниченный сверху энергетический спектр.

нормировки

$$\sum_k W(E_k) = AZ = 1 \qquad (33.3)$$

удовлетворяется благодаря сходимости суммы состояний.

Квазиравновесные состояния с отрицательной температурой были открыты экспериментально Парселом и Паундом в 1951 г. [12] для спиновых систем, удовлетворяющих условию (33.1).

Для систем, состоящих из плюс-частиц, энергия системы E_n превышает минимальное возможное значение $E_{\min}$ и может принимать сколь угодно большое значение, т. е.

$$E_{\min} \leqslant E_n < +\infty, \qquad (33.4)$$

так как кинетическая энергия частиц может быть как угодно большой.

Для таких систем модуль Θ не может быть отрицательным, так как при $\Theta < 0$ условие нормировки (33.3) не могло бы быть удовлетворено при константе A, отличной от нуля, вследствие расходимости суммы состояний Z. Следовательно, при выполнении условия (33.4) модуль Θ может быть только положительным, т. е. $T > 0$. Однако для минус-частиц, очевидно,

$$-\infty < E_n \leqslant E_{\max} \qquad (33.5)$$

и, следовательно, распределение (33.2) имеет смысл лишь при $\Theta < 0$, т. е. при отрицательных абсолютных температурах. Следовательно, системы, состоящие из частиц отрицательной массы, могут находиться в состоянии термодинамического равновесия только при отрицательных абсолютных температурах.

Легко видеть, что распределение (33.2) не может быть равновесным для систем, содержащих частицы положительной массы, взаимодействующие с частицами отрицательной массы. Для таких систем энергия не ограничена ни сверху, ни снизу, и поэтому сумма вероятностей (33.2) по всем состояниям будет расходящейся при любом, отличном от нуля значении A. Следовательно, такие системы вообще не могут находиться в состоянии термодинамического равновесия с определенной температурой. Для них возможны лишь квазиравновесные состояния, когда подсистема, содержащая частицы положительной массы, находится во внутреннем равновесии, характеризуемом положительной температурой, а подсистема, состоящая из частиц отрицательной массы, имеет отрицательную температуру. В силу взаимодействия таких систем между собой температура каждой из них будет изменяться и общая система вообще не будет равновесной. Если, однако, это взаимодействие очень слабо, то нарушение равновесия будет происходить крайне медленно и для каждой из подсистем можно пользоваться представлениями равновесной термодинамики.

Итак, системы, состоящие из частиц отрицательной массы, с термодинамической точки зрения могут рассматриваться как обладающие отрицательной температурой, если взаимодействие этих систем с обычными системами положительной массы достаточно мало.

§ 34. Частицы отрицательной массы и космические лучи

В § 27 мы рассмотрели модель вселенной, имеющей в среднем нулевую собственную массу, заполненную приблизительно равномерно минус-частицами с плотностью 10^{-30} *г/см*^{3} и плюс-частицами, собранными в виде звезд, планет, межзвездного газа и других космических объектов. Если минус-частицы взаимодействуют с плюс-частицами посредством гравитационных, слабых и сильных взаимодействий *, то их можно рассматривать как систему, слабо взаимодействующую с системой плюс-частиц, т. е. как квазиизолированную. Такая система может находиться в квазиравновесном состоянии при отрицательной температуре. Следовательно, в рассматриваемой модели вселенная состоит из системы плюс-частиц, неравномерно распределенных в пространстве и имеющих положительные температуры, и системы минус-частиц, приблизительно равномерно пространственно распределенных и имеющих отрицательную температуру.

Из термодинамики известно, что имея тепловой резервуар, находящийся в состоянии с отрицательной тем-

* Исключим электромагнитные взаимодействия на основании соображений, изложенных в конце § 29.

пературой, возможно осуществить *термодинамическую машину второго рода*, которая может производить работу лишь за счет непрерывного охлаждения единственного теплового резервуара. Такими машинами второго рода фактически являются лазеры, создающие интенсивный световой поток за счет охлаждения квазиизолированной подсистемы, предварительно приведенной в состояние с отрицательной температурой *.

С тепловым резервуаром, состоящим из минус-частиц, очевидно, также может быть осуществлена термодинамическая машина второго рода. Такая машина в отличие от лазера может производить неограниченное количество работы, поскольку тепловой резервуар, состоящий из минус-частиц, не имеет конечной минимальной энергии и при неограниченном понижении температуры может отдавать неограниченные количества энергии. Следовательно, во вселенной, состоящей из плюс- и минус-частиц, осуществим вечный двигатель второго рода, т. е. принципиально новый, практически неисчерпаемый энергетический источник.

Таким образом, появляется принципиально новая возможность объяснения ряда загадочных астрофизических явлений, таких, например, как излучение радиогалактик и квазаров. Эти объекты излучают столь большие количества энергии, что представляется невозможным объяснить их какими-либо ядерными или термоядерными реакциями. Появляется также новая возможность объяснения генерации космических лучей в межзвездном и межгалактическом пространстве, а также в атмосфере Солнца. Рассмотрим последнюю возможность более подробно.

Пусть $f(E, \mathbf{r}, t)$ — функция распределения релятивистских космических частиц по энергиям и пространству; ρ_0^+ — пространственная плотность нерелятивистских плюс-частиц, т. е. плотность обычного вещества (межзвездного газа, вещества планет, звезд и т. п.); ρ^- — пространственная плотность минус-частиц. Сталкиваясь

---

* Следует заметить, что лазер, будучи термодинамической машиной второго рода, не является вечным двигателем второго рода. Подсистема с отрицательной температурой, питающая лазер, имеет лишь ограниченный запас энергии и при охлаждении до сколь угодно низких температур может произвести лишь ограниченное количество работы.

с минус-частицами, релятивистская частица космических лучей будет увеличивать в среднем свою энергию за счет увеличения отрицательной кинетической энергии минус-частиц *. Предположим, что в результате столкновений с минус-частицами космическая частица в среднем ежесекундно приобретает некоторую долю α своей энергии E. Иначе говоря, будем считать, что средняя энергия частицы увеличивается в соответствии с уравнением

$$d\bar{E}/dt = \alpha\bar{E}. \tag{34.1}$$

Коэффициент α, очевидно, пропорционален концентрации минус-частиц, скорости космических частиц и эффективному поперечному сечению упругого столкновения плюс-частиц с минус-частицами σ_{+-}, т. е. равен

$$\alpha = n^- c\, \sigma_{+-}, \tag{34.2}$$

где $n^- = \rho^-/m^-$ — концентрация; m^- —масса минус-частиц.

В результате этого процесса увеличения средней энергии космических частиц функция распределения будет изменяться со временем в соответствии с уравнением

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial E}(\alpha E f) = 0. \tag{34.3}$$

Действительно, согласно последнему уравнению для функции f, зависящей только от E и t, имеем

$$\frac{d\bar{E}}{dt} = \frac{1}{n^+}\frac{d}{dt}\int_0^\infty Ef\,dE = \frac{1}{n^+}\int_0^\infty E\frac{\partial f}{\partial t}\,dE =$$

$$= -\frac{1}{n^+}\int_0^\infty E\frac{\partial}{\partial E}(\alpha E f)dE = -\frac{\alpha E^2}{n^+}f\Big|_0^\infty + \frac{\alpha}{n^+}\int_0^\infty Ef\,dE,$$

где

$$n^+ = \int_0^\infty f\,dE$$

* При столкновении плюс-частицы с плюс-частицами покоящейся среды средняя энергия частицы убывает, если средняя кинетическая энергия частиц среды меньше кинетической энергии частицы, и возрастает, если кинетическая энергия частицы меньше средней кинетической энергии частиц среды. При столкновении же плюс-частицы с покоящейся средой, состоящей из минус-частиц, средняя кинетическая энергия частицы возрастает во всех случаях.

есть среднее число космических частиц в единице объема, т. е. средняя их концентрация. Отсюда, полагая f конечной при $E = 0$ и исчезающей достаточно быстро при $E \to \infty$, получаем уравнение (34.1).

Пространственно-энергетическая плотность f будет изменяться также в результате столкновений космических частиц с обычным веществом. Если T — среднее время пробега космической частицы, то функция f ежесекундно должна уменьшаться в $1/T$ раз. Очевидно,

$$T = 1 / \overset{+}{n}_0 c \sigma_{+0}, \tag{34.4}$$

где σ_{+0} — полное эффективное поперечное сечение столкновения космической частицы с частицами нерелятивистского вещества; $\overset{+}{n}_0 = \overset{+}{\rho}_0 / m^+$; m^+ — средняя масса плюс-частиц вещества.

Пространственно-энергетическая плотность f может также изменяться в результате обычной пространственной диффузии космических частиц в межзвездных магнитных полях.

Учитывая эти два эффекта, а также наличие первоисточников космических лучей, уравнение (34.3) необходимо дополнить тремя членами, в результате чего пространственно-энергетическое уравнение диффузии космических частиц приобретает вид

$$\frac{\partial f}{\partial t} - D\nabla^2 f + \frac{\partial}{\partial E}(\alpha E f) + \frac{f}{T} = Q, \tag{34.5}$$

где D — коэффициент пространственной диффузии; Q — плотность первоисточников космических частиц.

В соответствии с [13—16] это уравнение диффузии является укороченным, не учитывающим диффузионного разброса по энергиям, т. е. не содержащим члена со второй производной по E. Однако оно достаточно для описания механизма ускорения частиц, подобного механизму Ферми [17].

Простейшим решением этого уравнения является стационарное, пространственно однородное решение, т. е. решение уравнения

$$\frac{\partial}{\partial E}(\alpha E f) + \frac{f}{T} = 0 \tag{34.6}$$

для энергий, превышающих максимальную энергию первоисточника. Такое решение имеет вид

$$f = AE^{-\gamma}, \tag{34.7}$$

где A — константа интегрирования;

$$\gamma = 1 + 1 / \alpha T = 1 + n_0^+ \sigma_{+0} / n^- \sigma_{+-}. \qquad (34.8)$$

Естественно считать минус-частицы обладающими достаточно большой средней отрицательной температурой, так как они постоянно приобретают отрицательную кинетическую энергию от частиц космическиой радиации и не могут скапливаться в виде планет, звезд и других плотных космических объектов, сохраняющих не слишком высокую температуру. Поэтому n^- можно считать приблизительно постоянной величиной для всей метагалактики *. Таким образом, коэффициент γ должен иметь минимальное значение в межгалактическом пространстве и возрастать внутри туманностей и вблизи звезд.

Гипотеза метагалактического происхождения космических лучей, которая легко объясняет факт их анизотропности, соответствует предположению

$$\frac{n_0^+ \sigma_{+0}}{n^- \sigma_{+-}} \simeq 2, \qquad (34.9)$$

так как для первичной компоненты космических лучей $\gamma \simeq 3$.

Подобное соотношение между плотностями и поперечными сечениями представляется естественным, если считать средние метагалактические плотности плюс- и минус-частиц приблизительно одинаковыми и эффективные сечения σ_{+0} и σ_{+-} одного и того же порядка.

Итак, изотропность и энергетический спектр первичной компоненты легко объясняются, если считать, что космические лучи образуются в межгалактическом пространстве, и сделать простейшие предположения о плотности минус-частиц и о поперечных сечениях взаимодействия минус-частиц с плюс-частицами [18].

Вблизи звезд концентрация частиц вещества n_0^+ повышается, и если считать, что концентрация минус-частиц n^- остается практически постоянной, то согласно (34.8)

---

* Плотность минус-частиц может увеличиваться внутри звезд, так как при отрицательных температурах, согласно распределению Больцмана,

$$n^- \sim \exp(-m\varphi / kT) = \exp(-|m|\varphi / k|T|),$$

где φ — ньютоновский потенциал тяготения. Однако при больших $|T|$ это увеличение будет незначительным.

показатель γ увеличивается. Таким образом, вблизи Солнца также должны генерироваться космические лучи, однако их спектр должен быть более крутым, чем для космических лучей, генерируемых всей метагалактикой. Интенсивность энерговыделения, т. е. энергия, приобретаемая космическими частицами в единичном объеме, в единицу времени, очевидно, пропорциональна концентрации n^+ космических частиц и приращению средней энергии в единицу времени, т. е. согласно (34.1)

$$dW / dt = \alpha \overline{E} n^+, \tag{34.10}$$

где W — плотность энергии космических частиц.

Таким образом, при приблизительно постоянном для всех пространственных областей коэффициенте α интенсивность энерговыделения возрастает пропорционально плотности космических частиц. Следовательно, если в какой-либо области пространства, например вблизи сверхновой звезды, образовалось скопление релятивистских частиц, то и энерговыделение в этой области будет повышенным.

Для объяснения радиоизлучения радиогалактик достаточно предположить, что в них в силу каких-то процессов скопились релятивистские частицы с концентрацией, на несколько порядков превышающей концентрацию газа в межгалактическом пространстве, так что в заданный момент мощность общего радиоизлучения этих частиц в магнитных полях радиогалактики равна наблюдаемой величине. Такое энерговыделение может поддерживаться постоянным за счет столкновений релятивистских частиц с минус-частицами, как это видно из формулы (34.10). Действительно, при $n^- \sim 10^{-6}$ $см^{-3}$, $\sigma_{+-} \sim 10^{-26}$ $см^2$, $c = 3 \cdot 10^{10}$ $см/сек$ получаем $\alpha \sim 3 \cdot 10^{-22}$ $сек^{-1}$. Полагая далее $\overline{E} \sim 10^{10}$ $эв \simeq 1{,}6 \cdot 10^{-2}$ $эрг$, $n^+ \sim 10^{-3}$ $см^{-3}$, получаем для радиогалактики объема 10^{68} $см^3$ общую выделяемую мощность $\sim 5 \cdot 10^{41}$ $эрг/сек$, что соответствует наблюдаемой величине. При повышении концентрации релятивистских частиц в разумных пределах выделяемая мощность может быть соответственно значительно увеличена. Таким образом, рассматриваемая гипотеза позволяет естественным образом объяснить и колоссальное энерговыделение в квазарах, не прибегая к сомнительной гипотезе гравитационного коллапса.

ЛИТЕРАТУРА

1. A. Einstein. Ann. Phys., **17**, 891 (1905).
2. H. Poincarè. Rend. Circolo Mat. Palermo, **21**, 129 (1906).
3. H. Minkowski. Phys. Z. S., **10**, 104 (1909).
4. Л. И. Мандельштам. Полное собрание трудов, т. V. Изд-во АН СССР, 1950.
5. H. A. Lorentz. Proc. Acad. Sci. Amsterdam, **6**, 809 (1904).
6. Р. В. Серебряный. ЖЭТФ, **20**, 1130 (1950).
7. Я. П. Терлецкий. Динамические и статистические законы физики. Изд. МГУ, 1950.
8. Я. П. Терлецкий. Статистическая физика. «Высшая школа», 1966.
9. Я. П. Терлецкий. ДАН СССР, **133**, 329 (1960).
10. Y. P. Terletsky. J. Phys. et Radium, **21**, 681 (1960); **23**, 910 (1962).
11. P. A. M. Dirac. Proc. Roy. Soc., **117**, 610; **118**, 341 (1928).
12. E. M. Purcell and R. V. Pound. Phys. Rev., 81, 279 (1951).
13. Я. П. Терлецкий и А. А. Логунов. ЖЭТФ, **21**, 567 (1951).
14. А. А. Логунов и Я. П. Терлецкий. Изв. АН СССР, серия физ., **17**, 119 (1953).
15. P. Morrison, S. Olbert, B. Rossi. Phys. Rev., **94**, 440 (1954).
16. Я. П. Терлецкий. ДАН СССР, **101**, 59 (1955).
17. E. Fermi. Phys. Rev., **75**, 1169 (1949).
18. Yakov P. Terletsky. Ann. Inst. Henri Poincaré, **1**, 431 (1964); Quasi—Stellar Sources and Gravitational Collapse. Chicago, 1965, стр. 466.

ОГЛАВЛЕНИЕ

Яков Петрович Терлецкий

Парадоксы теории относительности

Утверждено к печати Московским обществом испытателей природы

Редактор издательства *К. П. Гуров* Художник *В. А. Назаров*

Технический редактор *Л. И. Матюхина*

Сдано в набор 22/XII 1965. Подписано к печати 30/III 1966 г. Формат $84\times108^1/_{32}$

Печ. л. 3,75. Усл. печ. л. 6,30. Уч.-изд. л. 5,3. Тираж 16.500 экз. Т-03604

Изд. № 506/66. Тип. зак. 25

Цена 33 к.

Издательство «Наука». Москва, К-62, Подсосенский пер., д. 21

2-я типография издательства «Наука». Москва, Г-99, Шубинский пер., 10