

ПРОСТРАНСТВО-ВРЕМЯ,
ГЕОМЕТРИЯ,
КОСМОЛОГИЯ

У.Бёрке

Spacetime, Geometry, Cosmology

WILLIAM L. BURKE

University of California, Santa Cruz

University Science Books
Mill Valley, California, 1980

ПРОСТРАНСТВО-ВРЕМЯ, ГЕОМЕТРИЯ, КОСМОЛОГИЯ

У.Бёрке

Перевод
с английского
кандидатов физ.-мат. наук
Ю. В. Граца, В. И. Хлебникова и
З. А. Штейнграда
под редакцией
д-ра физ.-мат. наук
Д. В. ГАЛЬЦОВА



Москва
«Мир»
1985

ББК 22.313
Б48
УДК 530.1:531.51 + 523.12

Бёрке У.

Б48 Пространство-время, геометрия, космология. Пер. с англ. —
М.: Мир, 1985. — 416 с., ил.

Написанное ученым из США учебное пособие по специальной и общей теории относительности, которое может служить элементарным и вместе с тем современным введением в теорию гравитации и космологию. Большое внимание уделяется геометрическим аспектам теории, широко используется современная дифференциальная геометрия.

Рассчитана на студентов физических и математических факультетов, преподавателей физики, а также читателей-неспециалистов, интересующихся современными проблемами общей теории относительности и космологии.

Б $\frac{1704020000-204}{041(01)-85}$ 59—85, ч. 1

ББК 22.313
530.1

Редакция литературы по физике

Предисловие редактора перевода

Книга профессора Калифорнийского университета У. Бёрке написана на основе курса лекций для студентов, намеревающихся специализироваться в области космологии. В последнее время интерес к космологическим проблемам чрезвычайно возрос, и появление учебного пособия такой направленности закономерно и полезно. Книга содержит доступное младшекурсникам изложение основ специальной и общей теории относительности и элементарное введение в современную теорию строения Вселенной. По уровню сложности она находится примерно посередине между курсами общей и теоретической физики, принятыми в наших университетах. Описательному материалу отводится подчиненная роль. Основную свою задачу автор видит в физическом обосновании и более глубоком осмыслении геометрических структур, составляющих основу наших представлений о пространстве-времени.

В отличие от большинства учебников такого уровня изложение в книге Бёрке, оставаясь элементарным, ведется на современном языке дифференциальной геометрии. Автору удалось найти удачный компромисс между степенью абстрактности применяемых математических средств и конкретным физическим содержанием. В книге много интересных и неожиданных примеров, выявляющих существо таких понятий, как метрика, системы отсчета, собственное время и т. д. Так, смысл метрики искривленного пространства-времени поясняется на основе анализа распространения волновых пакетов на поверхности воды: метрический тензор связывается с законом дисперсии выделенного типа волн. В этой «теории относительности для волн на воде» ход часов задается колебаниями поверхности жидкости, световые сигналы изображаются пакетами волн ряби, имеющих другой закон дисперсии. Эту задачу, кстати, читателю будет полезно вспомнить и при изучении квантовой механики.

В название книги не случайно включено слово «геометрия». Довольно много места в ней уделено линейным пространствам, отображениям, дифференциальной геометрии многообразий. Автор пользуется языком линейных форм, безындексны-

ми обозначениями для векторов и тензоров. Это не дань моде, а путь к более глубокому раскрытию геометрического смысла тензорных величин. Усвоению такого языка помогают многочисленные графические иллюстрации и физические примеры. Для дальнейшего изучения методов дифференциальной геометрии многообразий в общей теории относительности читателю можно рекомендовать книгу Хокинга и Эллиса [1].

Стремление к большей наглядности характеризует и основную часть книги, посвященную космологии. Обсуждая геометрию однородных изотропных пространств, Бёрке апеллирует к «четырёхмерному» пространственному воображению, используя в качестве четвертой координаты цвет. Здесь довольно много описательного материала, однако, предостерегая от опасности сбиться на тон бездоказательных общих рассуждений, автор основное внимание уделяет математическим аспектам теории. Разумеется, эта часть книги может служить лишь начальным введением в космологию. Читателю, желающему продолжить изучение предмета, можно рекомендовать, помимо книг, указанных автором, книги Зельдовича и Новикова [2] и Пиблса [3].

Следует иметь в виду, что многие «старые» загадки космологии, о которых говорится в тексте, — проблема скрытой массы, проблема однородности (горизонта), трактовка сингулярности, физика «очень ранней» Вселенной — получили новое освещение в самые последние годы, что не могло найти отражения в книге Бёрке. Характерно, что это продвижение было связано не столько с накоплением и уточнением данных астрономических наблюдений, сколько с успехами теории элементарных частиц и развитием теории гравитации.

Большие усилия прилагаются сейчас в направлении создания последовательной квантовой теории гравитации. И хотя в окончательном виде такая теория еще не сформулирована, по видимому, уже сейчас многие элементы ее понимаются правильно. Теоретики научились рассчитывать квантовые поправки к уравнениям Эйнштейна [4 — 6]. Было показано, что одним из решений вакуумных уравнений Эйнштейна с учетом квантовых поправок является несингулярное пространство де Ситтера [7]: Все большую популярность завоевывает точка зрения, что деситтеровская Вселенная предшествовала фридмановской и переходила в последнюю в процессе эволюции. Оказалось, что такое представление хорошо согласуется и с новыми идеями, возникшими в теории элементарных частиц.

Космология — несомненно самая «романтическая» из существующих физических теорий. Ее задачи завораживают воображение. Перекидывая мостик между микрофизикой немалых малых расстояний (где характерным масштабом является планковская длина 10^{-33} см) и физикой сверхгигантских мас-

штабов, характеризующих Вселенную в целом, космология призвана служить венцом физической картины мира. Для описания сверхплотного состояния ранней Вселенной необходимо знать законы, управляющие поведением элементарных частиц при энергиях вплоть до 10^{19} ГэВ ($1 \text{ ГэВ} = 10^9 \text{ эВ}$) [8]. Ясно, что достичь столь высоких энергий с помощью ускорителей или иных искусственно созданных устройств никогда не удастся (хотя не исключены косвенные методы проверки теорий, такие, как ведущийся в настоящее время поиск распада протона). К счастью, естественную возможность проверки теорий элементарных частиц при сверхвысоких энергиях дает сама космология. Сформулировав те или иные гипотезы о свойствах частиц при сверхвысоких энергиях, можно построить соответствующий «сценарий» ранней Вселенной и затем проверять согласование извлекаемых следствий с данными астрономических наблюдений. Поразительно то, что, несмотря на единичность и невозпроизводимость космологического «эксперимента», возможности теоретических построений довольно жестко ограничиваются совокупностью наблюдательных данных. Можно поэтому сказать, что задача состоит не в том, чтобы сначала угадать «правильную» теорию элементарных частиц и затем применить ее к ранней Вселенной, а в том, чтобы построить космологически согласованную теорию элементарных частиц. Существенный прогресс в этом направлении был связан с созданием моделей Великого объединения.

Открытие нейтральных токов, а затем обнаружение W- и Z-бозонов блестяще подтвердили объединенную теорию слабых и электромагнитных взаимодействий. В этой теории важную роль играет скалярное поле (поле Хиггса), которое может иметь отличное от нуля среднее значение в состоянии вакуума. За счет этого поля приобретают массы фермионы и промежуточные бозоны, а само единое электрослабое взаимодействие оказывается разделенным на взаимодействия слабые и электромагнитные [9]. Появление скалярного поля аналогично возникновению спонтанной намагниченности ферромагнетика — симметрия основного состояния (вакуума) при этом становится отличной от более высокой симметрии гамильтониана (спонтанное нарушение симметрии). При нулевой температуре состояние нарушенной симметрии с отличным от нуля средним значением скалярного поля энергетически более выгодно. Однако с повышением температуры квантовые флуктуации будут разрушать «ферромагнитную» фазу вакуума, и при температуре порядка 10^{15} К (соответствующей энергии 100 ГэВ) симметрия должна восстанавливаться, а скалярное поле исчезать.

В моделях Великого объединения предполагается, что аналогичная ситуация имеет место и для сильных взаимодействий, причем их слияние с электрослабыми взаимодействиями отве-

чает энергии 10^{15} ГэВ. При температурах порядка 10^{28} К, соответствующих этой энергии, должен происходить фазовый переход: при более высоких температурах устойчива симметричная фаза, при более низких — «ферромагнитная». Такой фазовый переход сопровождается появлением еще одного скалярного поля [10].

Космологические следствия моделей Великого объединения оказались многообещающими. Одним из важных для космологии предсказаний этих моделей является нестабильность протона со временем жизни $10^{29} - 10^{33}$ лет (отметим, что экспериментально поиск распада протона пока не увенчался успехом). Возможность распада протона означает, что барионный заряд не сохраняется строго. Это открывает принципиальную возможность объяснения барионной асимметрии Вселенной, т. е. наблюдаемого отсутствия антивещества [11]. Еще более интересные последствия могут иметь описанные выше фазовые переходы.

Согласно обсуждаемой в разд. 49 книги теории горячей Вселенной, в моменты времени, предшествовавшие 10^{-31} с, температура была выше 10^{28} К, скалярные поля отсутствовали и вакуум находился в симметричной фазе. По мере остывания при расширении должен сначала происходить фазовый переход, в результате которого отделяются сильные взаимодействия ($t \sim 10^{-31}$ с), а затем второй фазовый переход ($t \sim 10^{-19}$ с), при котором разделяются слабые и электромагнитные взаимодействия. Наконец, при еще меньших энергиях (порядка сотен МэВ), вероятно, происходит еще один фазовый переход, сопровождающийся «запиранием» кварков и образованием сильно-взаимодействующих частиц — адронов — из кварк-глюонного «супа». Таким образом, теория Великого объединения приводит к сценарию горячей Вселенной, по которому происходит не просто ее остывание, но и имеют место фазовые переходы, сопровождающиеся более существенными перестройками ее структуры.

Изложенный сценарий также приводит к возможности де-ситтеровской стадии в эволюции ранней Вселенной, характеризующейся экспоненциальным «раздуванием» ее размеров (иначе эту стадию называют «инфляционным» периодом: по экспоненциальному закону происходит рост цен во время инфляции [12]; оба термина, «раздувание» и «инфляция», являются вариантами перевода английского термина *inflation*, введенного в работе [13], где такой сценарий был впервые четко сформулирован). Инфляционная стадия эволюции Вселенной весьма привлекательна, так как позволяет естественно решить проблемы горизонта и плоскостности, т. е. близости наблюдаемой плотности вещества к критической (см. разд 50).

Экспоненциально быстрое расширение Вселенной (соответствующее пространство-время, как уже указывалось, является деситтеровским) могло начаться, если фазовый переход при энергиях порядка 10^{15} ГэВ был переходом первого рода. В этом случае симметричное состояние вакуума («ложный» вакуум) метастабильно, и для образования новой фазы требуется некоторое время. Если это время достаточно велико, Вселенная успевает расшириться, пока фазовый переход еще не завершился — происходит «переохлаждение» ложного вакуума. Существенно, что такое состояние вакуума характеризуется конечной положительной плотностью энергии, в то время как плотность энергии несимметричного вакуума, который должен образоваться в результате фазового перехода, равна нулю. Если переохлаждение достаточно сильное, то плотность энергии обычного вещества становится много меньше плотности энергии ложного вакуума и первой можно вовсе пренебречь. Тензор энергии-импульса вакуума пропорционален метрическому тензору $T_{\mu\nu} = \rho_{\text{вак}} g_{\mu\nu}$, поэтому мы приходим к вакуумным уравнениям Эйнштейна с положительной эффективной космологической постоянной. Это эквивалентно «веществу», создающему отрицательное давление, по абсолютной величине равное плотности энергии $p = -\rho_{\text{вак}}$ (ср. с формулой (41.9) в тексте). Из уравнения (4.88) тогда следует, что величина $\rho_{\text{вак}}$ остается при расширении постоянной, а само расширение происходит по экспоненциальному закону в соответствии с уравнением (4.87) при $\rho = -\rho$ и $\kappa = 0$, т. е. $R \sim \exp(Ht)$, где $H = \left(\frac{8\pi}{3} \rho_{\text{вак}}\right)^{1/2}$

Постоянная H столь велика, что за ничтожное время порядка 10^{-36} с Вселенная успеет расшириться в 10^{30} раз [12]. Это дает простое объяснение одинаковости Вселенной в областях, столь удаленных друг от друга, что находящиеся в них наблюдатели не успели бы обменяться световыми сигналами за все время существования Мира. Нетрудно также заметить, что даже отличное от нуля первоначальное значение параметра кривизны κ в уравнениях (48.1), (48.2) экспоненциально быстро становится несущественным — соответствующий вклад убывает как $\exp(-2Ht)$. Таким образом, для объяснения наблюдаемой в настоящее время плотности, близкой к критической, не нужно предполагать чересчур точной «плоскостности» Вселенной на ранних стадиях расширения, как этого требует стандартная модель. Интересно, что полная энергия Вселенной при «раздувании» возрастает в колоссальной степени, поскольку плотность $\rho_{\text{вак}}$ остается величиной постоянной. Тем самым оказывается, что энергия, заключенная во Вселенной, возникает из вакуума.

Дальнейший анализ показал, что реалистичность подобного сценария критически зависит от деталей фазового перехода, которые в свою очередь определяются выбором конкретного варианта модели Великого объединения. Если, как это было предложено в первоначальном варианте сценария инфляционной Вселенной, термализация и выход на режим фридмановского расширения происходят за счет столкновений зародышей несимметричной фазы, то Вселенная в результате окажется неприемлемо сильно неоднородной. При таком фазовом переходе должны также рождаться магнитные монополи в количестве, явно противоречащем наблюдениям. Однако существуют модели, в которых переход вакуума в несимметричную фазу происходит не скачком, а постепенно, так что значение плотности энергии новой фазы медленно убывает. За это время каждый зародыш новой фазы успеет расшириться до размеров, значительно превышающих размеры видимой части Вселенной, которая, таким образом, может полностью оказаться внутри единственного «пузыря» несимметричной фазы [14]. Такой вариант сценария инфляционной Вселенной решает одновременно и проблему монополей, поскольку плотность рождающихся монополей должна быть одного порядка с плотностью пузырей новой фазы.

Если инфляционный сценарий выдержит проверку временем, то на повестку дня выйдет исследование еще более ранней истории Вселенной. Не исключено, однако, что вопрос будет решен принципиально иным способом. В последнее время высказывается интригующая мысль, что классическое гравитационное поле (т. е. метрика, а с ней и само представление о пространстве и времени) возникло в результате квантового перехода из некоторого «состояния», описание которого должна дать полная квантовая теория гравитации и в котором обычные понятия пространства и времени лишены смысла. Возможно, что именно деситтеровская Вселенная окажется начальным классическим состоянием мира, возникшим из квантового хаоса [15].

Говоря о новых веяниях в космологии восьмидесятых годов, нельзя не коснуться вопроса о возможной роли нейтрино во Вселенной. Перечисляя в разд. 36 виды материи, играющие главную роль в космологии, Бёрке не упоминает нейтрино. Между тем из описываемой в книге теории горячей Вселенной следует, что при высоких температурах нейтрино находились в термодинамическом равновесии с другими частицами и их концентрация по порядку величины равнялась концентрации фотонов. В ходе расширения общая концентрация частиц уменьшалась, и вследствие слабости нейтринных взаимодействий по истечении времени порядка $0,1$ с после начала расширения нейтрино стали практически свободными. В современную эпоху

концентрация реликтовых нейтрино (каждого сорта) примерно в три раза меньше концентрации реликтовых фотонов [11]. Если нейтрино, как раньше полагали, имеют массу, равную нулю, то энергия каждой частицы будет чисто тепловой и отвечает температуре порядка 2 К. Соответствующая плотность энергии (в пересчете на эффективную массу) имеет порядок 10^{-34} г/см³, и ее вклад, как и вклад реликтовых фотонов, составляет лишь 10^{-4} от плотности обычного вещества. Поэтому если нейтрино безмассовы, их роль в динамике Вселенной в целом малосущественна.

Однако в 1980 г. группа ученых из Института экспериментальной и теоретической физики во главе с В. А. Любимовым сообщила о результатах эксперимента, в котором с большой точностью измерялся энергетический спектр электронов, вылетающих при β -распаде ядер тяжелого водорода (трития). Если нейтрино массивны, то спектр электронов должен обрываться раньше, поскольку массивные нейтрино в отличие от безмассовых не могут уносить сколь угодно малую энергию — ее минимальное значение равно энергии покоя. Из результатов измерений следовало, что масса нейтрино лежит в интервале 30 ± 16 эВ. Хотя этот вывод еще нельзя считать окончательным, вытекающие из него следствия для астрофизики и космологии столь радикальны, что не могут быть обойдены молчанием.

Действительно, простой подсчет показывает, что при массе электронных нейтрино 30 эВ ($5 \cdot 10^{-32}$ г) в реликтовых нейтрино должна быть сосредоточена почти вся масса вещества во Вселенной (97%). Эта цифра еще увеличится, если другие типы нейтрино также массивны. Полная плотность вещества во Вселенной при этом оказывается порядка 10^{-29} г/см³, т. е. порядка критической плотности. Заметим, что именно из космологических соображений следуют наиболее жесткие ограничения на суммарную массу всех типов нейтрино — она не может существенно превышать 40 эВ, иначе время существования Вселенной окажется слишком малым [11].

Существуют причины, по которым массивные нейтрино крайне «нужны» в астрофизике и космологии. Наличие массы у нейтрино существенно изменяет картину возникновения структурных неоднородностей в первоначально однородном распределении вещества [16]. Под действием собственного тяготения нейтрино будут достаточно быстро образовывать сгущения, которые своим гравитационным полем должны увлекать и обычное вещество. Если же нейтрино безмассовы и создаваемое ими гравитационное поле пренебрежимо мало, процесс развития гравитационной неустойчивости должен происходить гораздо медленнее и для объяснения наблюдаемой в настоящее

время структуры Вселенной приходится предполагать слишком высокий уровень флуктуаций плотности в раннюю эпоху. Одновременно масса нейтрино разрешает парадокс скрытой массы, о котором упоминается в разд. 37. Согласно наблюдениям, в скоплениях галактик действует гравитационное поле, для создания которого необходима по крайней мере на порядок большая масса, чем суммарная масса самих галактик. Величина недостающей массы хорошо согласуется с массой сгущений, образующихся при развитии гравитационной неустойчивости массивных нейтрино. Конечно, нельзя утверждать, что проблемы гравитационной неустойчивости и скрытой массы не могут быть разрешены иначе, однако простота приведенных выше объяснений позволяет говорить о наличии астрофизических указаний на существование массы у нейтрино, что делает саму проблему экспериментального поиска массы нейтрино еще более острой.

Этот краткий обзор идей, получивших развитие уже в восьмидесятые годы, показывает, какими гигантскими темпами создается сейчас самая грандиозная физическая теория — теория Мира в целом. Надеемся, что предлагаемая вниманию читателей книга окажется полезной не только тем, кто хочет получить общее представление о космологии, но и тем, кто посвятит себя работе в этой захватывающе интересной области физики.

Перевод книги был выполнен кандидатами физико-математических наук Ю. В. Грацем (предисловия, раздел «Как пользоваться книгой», разд. 1 — 6 гл. I и гл. II), З. А. Штейнградом (введение, разд. 7 — 14 гл. I и гл. III) и В. И. Хлебниковым (гл. IV).

Д. Гальцов

ЛИТЕРАТУРА

1. Хокинг С., Эллис Дж. Крупномасштабная структура пространства-времени. — М.: Мир, 1977, с. 431.
2. Зельдович Я. Б., Новиков И. Д. Строение и эволюция Вселенной. — М.: Наука, 1975, с. 736
3. Пиблс Ф. Дж. Э. Структура Вселенной в больших масштабах. — М.: Мир, 1983, с. 408.
4. Гриб А. А., Мамаев С. Г., Мостепаненко В. М. Квантовые эффекты в интенсивных внешних полях. — М.: Атомиздат, 1980, с. 1 — 295.
5. Биррел Н., Девис П. Квантованные поля в искривленном пространстве-времени. — М.: Мир, 1984, с. 356.
6. Квантовая гравитация. Труды 2-го семинара «Квантовая теория гравитации», Москва, 13 — 15 окт. 1981 г. — М.: ИЯИ, 1982, с. 1 — 319.

7. *Starobinsky A. A.* A new type of isotropic cosmological model without singularity. — *Phys. Lett.*, 1980, v. 91 B, No 1, p. 99 — 102.
8. *Вейнберг С.* Первые три минуты. — М.: Энергоиздат, 1981.
9. *Окунь Л. Б.* Лептоны и кварки. — М.: Наука, 1981, с. 304.
10. *Киржниц Д. А., Линде А. Д.* Фазовые превращения в микромире и во Вселенной. — *Природа*, 1979, № 11, с. 20 — 30.
11. *Долгов А. Д., Зельдович Я. Б.* Космология и элементарные частицы. — *УФН*, 1980, т. 130, № 4, с. 559 — 614.
12. *Зельдович Я. Б.* Современная космология. — *Природа*, 1983, № 9, с. 11 — 24.
13. *Guth A. H.* Inflationary universe: a possible solution to the horizon and flatness problems. — *Phys. Rev.*, 1981, v. 23 D, No. 2, p. 347 — 356.
14. *Linde A. D.* A new inflationary universe scenario: a possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problem. — *Phys. Lett.*, 1982, v. 108, No. 6, p. 389 — 393.
15. *Гришук Л. П., Зельдович Я. Б.* Полные космологические теории: см. [6, с. 39 — 47].
16. *Бисноватый-Коган Г. С.* Нейтрино во Вселенной. — *Природа*, 1982, № 10, с. 26 — 31.

Предисловие автора к русскому изданию

Мне приятно, что моя книга переведена на русский язык. Надеюсь; что содержащиеся в ней конкретные примеры и иллюстрации послужат полезным и интересным дополнением к обычному более абстрактному изложению специальной и общей теории относительности. Обращение к моделям может помочь не только студенту, изучающему предмет, но и исследователю, развивающему его. Я благодарен доктору Д. В. Гальцову за усилия, сделавшие возможным появление этого перевода.

Санта-Крус, 1984

У. Бёрке

Посвящается Дж. Сингу

Предисловие автора

Короче говоря, я старался использовать весь свой опыт исследовательской работы и преподавания, чтобы дать читателю возможность понять, как работают другие, и поработать самостоятельно.

Дж. Полиа

Эта книга посвящена геометрии пространства-времени и ее применению к космологии. Она представляет собой вводный курс, знакомящий читателя с геометрическими понятиями и идеями, которые играют важную роль в современной теоретической физике. В тексте приводится много примеров и рассматриваются приложения теории, что должно облегчить ее усвоение. Книга рассчитана на серьезных студентов, которые хотят научиться проводить вычисления в рамках стандартных космологических моделей и понимать такие вычисления. В конце почти каждого раздела приводятся задачи. По уровню изложения теории относительности эта книга занимает промежуточное положение между книгами «Пространство-время» Тейлора и Уилера и «Гравитация» Мизнера, Торна и Уилера. Как учебник по космологии она необычна. В ней очень мало места отводится описательной стороне космологии и сложной физике Большого Взрыва. Вместо этого большое внимание уделяется фундаментальным идеям, лежащим в основе теории относительности и космологии, и геометрическому языку, необходимому для их обсуждения. Именно это нужно студенту, который хочет начать активно работать и должен для этого овладеть соответствующим аппаратом. Обычный описательный подход слишком легко может привести к тому, что студенты, на которых слишком большое впечатление производят красивые, убедительно звучащие фразы, станут заниматься пустыми и глупыми умозрительными рассуждениями.

По сложности данная книга примерно соответствует начальному курсу электромагнетизма, например книге Лоррайна и Корсона. От читателя требуется по крайней мере краткое предварительное знакомство с курсами линейной алгебры и обыкновенных дифференциальных уравнений. Было бы полезно прослушать (или изучить самостоятельно) курс классической механики, однако для чтения книги вполне достаточно тех знаний механики и электродинамики, которые даются вводными курсами физики. Предполагается, что читатель уже имеет некоторое представление (в объеме трех или четырех лекций) о специальной теории относительности. Нужно знать, что такое пара-

докс близнецов, но не обязательно уметь его объяснять. Сложность этой книги связана не с громоздкими аналитическими вычислениями, а с тем, что изложенный в ней материал требует основательного размышления. Это делает книгу полезным дополнением к обычным курсам, которые перегружены подробными выкладками. Я старался сделать книгу такой, чтобы ее можно было использовать как для занятий в аудитории, так и для самостоятельного изучения, возможно, в качестве дополнения к более традиционному курсу космологии.

Книга начинается с изложения специальной теории относительности. При этом я не следую аксиоматическому подходу, с помощью которого специальная теория относительности была построена. Вместо этого используется феноменологический подход, который не выглядит столь таинственно и более пригоден для критического рассмотрения. Я уделяю мало места парадоксам и всякого рода тайнам, а вместо этого стараюсь подчеркнуть действительно новое, что содержится в этой теории. При этом я рассматриваю новые идеи как разумную возможность, отнюдь не считая, что они в корне уничтожают все старые представления. Новые идеи, которыми оперирует специальная теория относительности, требуют нового языка. В книге разработан такой язык, который применяется затем к эйнштейновской теории гравитации — общей теории относительности. В заключение весь разработанный в книге аппарат используется для рассмотрения фридмановских моделей Вселенной, которые являются неотъемлемой частью современной космологии.

У меня было сильное искушение подчеркнуть особенности используемого в книге геометрического языка и дать ей подзаголовок «Книга 1-форм». Геометрический язык исчисления на многообразиях действительно представляет собой «новую математику» для физиков, и его должны изучить все. В качестве иллюстрации я рассматриваю применение аппарата 1-форм не только в космологии, но и при решении более земных задач, в частности для исследования распространения диспергирующих волн. Материал этой книги может служить введением в тензорный анализ. Однако после него следует изучить соответствующий более строгий математический курс. Математический аппарат используется в книге в объеме, минимально необходимом для того, чтобы аккуратно проводить все нужные выкладки. Я надеюсь, что даже люди, страдающие своего рода аллергией к математике, например некоторые нетерпеливые геофизики, с удовольствием прочитают эту книгу.

Прилежный студент, прочитав книгу, будет достаточно хорошо и неформально понимать, каким образом проводятся вычисления на многообразиях, иметь представление о геометрии

искривленного пространства-времени и уметь применять свои знания в рамках стандартных космологических моделей. В этой книге мы почти не рассматриваем физическую сторону космологии: образование галактик, синтез элементов и т. д. К сожалению, это значительно более сложные вопросы, и большая часть того, что мы знаем по этому поводу, может быть получена только с использованием сложного и непонятного машинного языка. Изучение этих вопросов в значительной степени менее привлекательно и доставляет меньше удовольствия, чем изучение описанного здесь геометрического подхода к космологии.

Уильям Бёрке

Я счастлив, что мне довелось познакомиться с теорией относительности в той форме, как ее излагал Дж. Л. Синг. В этом мне помог мой учитель Фрэнк Истебрук. Ученик сам решает, каким образом ему использовать то, чему его научили. Поэтому ни Дж. Синг, ни Ф. Истебрук не несут ответственности за мое пристрастие к 1-формам.

Мысль написать эту книгу мне подали мои студенты, окончательное же решение пришло при чтении биографии Моцарта: я спросил себя, почему я читаю биографии вместо того, чтобы делать что-нибудь самому. Я начал писать заметки по курсу космологии, который только что закончил читать студентам, и к своему удивлению в течение трех недель без особого труда исписал четыре сотни страниц. Эта книга является расширенным и в значительной степени переработанным вариантом той рукописи. Окончательная переработка заняла шесть месяцев и потребовала значительных усилий.

Студенты, слушавшие мой курс космологии, оказали мне огромную помощь в устранении имевшихся в рукописи погрешностей, они же опробовали первые варианты предложенных в книге задач. Эл. Келли, Петер Ренц и Джон Фолкнер прочитали черновой вариант рукописи. Бриан Хэтфилд проверил большую часть задач. Эйдан Келли блестяще осуществил редактирование. Их помощь неоценима. Чрезвычайно мудрые советы, которым я отдаю должное и которые я хочу передать другим, дали мне Джон Летчер («Пишите, не останавливаясь») и Ральф Байерлейн («Быстрее заканчивайте книгу, иначе вы просто устанете от нее»). Кэрл Фэрхэст, Джуди Роуз и Дотти Холлинджер прекрасно справились с работой по перепечатыванию окончательного варианта рукописи. И, конечно, огромной благодарности заслуживают Кип, Пако и в особенности Пэт.

Введение

Эта книга о пространстве-времени. В пределах небольших областей пространство-время приближенно описывается специальной теорией относительности подобно тому, как малый участок поверхности Земли приближенно описывается геометрией на плоскости. Однако явление гравитации в общей теории относительности Эйнштейна характеризуется свойствами пространства-времени в больших масштабах. Гравитация описывается кривизной пространства-времени, близким аналогом которой является кривизна поверхности Земли. По мере перехода ко все большим и большим областям пространства-времени гравитация становится доминирующим физическим фактором. Наконец, модель Вселенной, которую мы собираемся здесь рассмотреть, представляет собой самогравитирующий конгломерат энергии-массы, погруженный в искривленное пространство-время. Цель книги — подробно обсудить значение и смысл очерченного выше круга представлений и вооружить читателя языком и инструментами, необходимыми, чтобы пользоваться этими представлениями для предсказания и объяснения результатов астрономических наблюдений.

Что значит построить модель Вселенной? Мы должны поставить этот вопрос, хотя, казалось бы, он столь фундаментален, что относится лишь к сфере философии. Хорошим аналогом построения модели Вселенной является составление карты города. Город и карта города — вещи разные, и их не следует путать. Сложнее избежать такой путаницы в космологии. Далее, чтобы от карты был прок, она должна верно отражать некоторые существенные особенности города. Кроме того, она должна составляться в соответствии с определенными соглашениями; например, главные дороги должны быть нанесены красным цветом; север должен быть в верхней части и т. д. И еще, карта вовсе не должна определять, что такое дорога. Мы усваиваем такие понятия с помощью наглядной демонстрации.

Пространство-время

Модели

Нам говорят: вот это дорога, а это не дорога, и мы по индукции приобретаем представление о том, что такое дорога. Подобные элементарные представления, которые не определяются, а усваиваются в результате наглядной демонстрации, называются первичными понятиями.

Первичное понятие — это фундаментальный элемент физической теории, который не определяется в рамках теории, а предполагается известным из описания либо из более фундаментальной теории.

Все сказанное справедливо и для моделей физического мира, которые нам предстоит здесь построить. Мы будем составлять «карту» мира, руководствуясь некоторыми конкретными, но, вообще говоря, произвольными соглашениями. Для построения образа нам потребуется не бумага, как в случае дорожной карты, а определенная математическая структура — линейное векторное пространство. Поскольку мы можем толковать о физической реальности, только опираясь на ту или иную модель, самое простое — вообще забыть о различии между объектом и построенной нами моделью этого объекта. Как и в ситуации с картой города, в нашу модель тоже войдут некоторые элементы, которые не подлежат определению в ее рамках. Для пространства-времени первичными будут понятия *события, свободной частицы, часов и световых сигналов*. Наша модель позволяет дать словесное описание каждого из этих понятий подобно тому, как дорожная карта могла бы подсказать, что следует считать подходящей дорогой.

Крупномасштабное усреднение

Важным шагом при построении любой модели является отбрасывание излишних подробностей — операция, которую мы назовем крупномасштабным усреднением. При построении модели многое не принимается в расчет. Из обычной дорожной карты нельзя узнать ни точной ширины, ни качества покрытия дороги, ни его цвета. Но дорожная карта тем и удобнее точной фотографии, что в ней опущены такие несущественные детали. Подобное крупномасштабное усреднение имеет место и при описании наших первичных понятий, предназначенных для модели пространства-времени. В частности, мы не станем описывать часы путем построения подробной модели их внутреннего устройства. Такого рода подробности даже не будут упоминаться. В модели предполагается, что они несущественны, а уж задача эксперимента — выяснить, насколько это соответствует истине.

Построив модель и сформулировав правила обращения с ней, мы обретаем «язык», необходимый для обсуждения всего,

что связано с пространством-временем. Не овладев соответствующим языком, не имеет смысла вообще приступать к обсуждению. Само собой разумеется, что можно рассмотреть различные модели, и каждая из них может иметь свою, отличную от других, степень усреднения. Определенная доля содержания физической теории является неотъемлемой частью самого языка, на котором эта теория формулируется. Это очень тонкий момент, и он имеет как положительные, так и отрицательные стороны. Так, предположения, которые скрыты в самом языке и используются для обсуждения рассматриваемой ситуации, часто проходят незамеченными. Это представляет определенную опасность, ибо такие предположения нередко остаются непроверенными. В то же время язык, в который автоматически входит значительная часть содержания физической теории, обычно является исключительно эффективным рабочим языком. Так что, переходя от критического анализа специальной теории относительности к конкретным расчетам, мы каждый раз будем отдавать предпочтение именно тому языку, в который уже изначально «встроена» большая часть интересующей нас структуры.

Процесс математического моделирования той или иной физической ситуации мы будем называть построением *представления*. Обратный процесс, в результате которого той или иной математической структуре ставится в соответствие описываемая ею физическая ситуация, договоримся называть *реализацией* математической структуры. Обычно существует много различных, но эквивалентных представлений заданной физической ситуации. Здесь снова полезно прибегнуть к аналогии с дорожной картой, которую можно было бы составить, например, так, чтобы север был слева, а восток внизу. Такая карта была бы столь же достоверным представлением, что и обычная, но потребовала бы иных правил обращения, скажем при выборе направления пути на пересечении дорог. Правила перехода между подобными различными эквивалентными представлениями я называю преобразованиями *ковариантности*.

17.

[Внимание! Не все понимают ковариантность именно в таком смысле.]

Ковариантность — это смена правил и математических структур при переходе от одного эквивалентного представления к другому.

Смысл ковариантности заключается в том, что при изменении представления физическая ситуация остается той же самой. Это схематически показано на рис. 0.1.

Какую математическую структуру следует использовать для представления пространства-времени? При составлении до-

Теория многообразий — язык для описания пространства-времени

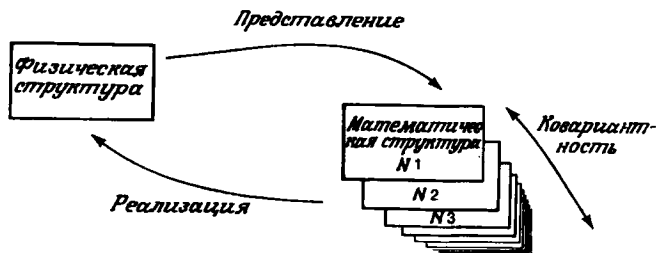


Рис. 0.1

Представление о ковариантности как о соотношениях между несколькими различными математическими структурами, каждая из которых моделирует одну и ту же физическую ситуацию.

рожной карты используется евклидова геометрия на плоскости. Для представления пространства-времени мы тоже воспользуемся геометрической структурой, но она должна быть более гибкой. Ведь нам нужно будет описывать не только статические конфигурации, но также движения и скорости протекания различных процессов. Подходящие для этой цели математические структуры называются многообразиями, и мы в дальнейшем уделим им соответствующее внимание. Но математику, разумеется, не следует путать с физикой. Общая теория относительности не является разделом математики. Поэтому математика в этой книге будет излагаться лишь в самом необходимом для наших задач объеме. Мы постараемся найти золотую середину между слишком большой дозой рафинированной математики и тем экстремально малым объемом, когда студент, заведенный в тупик наглядностью изложения, уже не способен без переучивания вникать в суть рассматриваемого вопроса. Таким образом, хотя многообразие и является подходящей геометрической структурой, мы дадим лишь нестрогое определение этого понятия. Строгое определение слишком сложно и довольно мало способствует как интуитивному восприятию, так и процессу вычислений. Если человек не способен уловить идею при нестрогом изложении, то возможно, что он вообще не готов к восприятию любой абстрактной идеи. Далее, несмотря на то что нам придется сталкиваться с вопросами топологического характера, мы не станем в них вдаваться сколько-нибудь глубоко. Топологические идеи довольно прозрачны и хорошо укладываются в голове уже на уровне интуитивного восприятия. Ведь несложно понять, что кофейная чашечка похожа на бублик, но не на апельсин или круг швейцарского сыра. Нам не понадобится и строгий анализ этих идей, и мы не будем им заниматься, тем более что он связан с огромными трудностями математического характера.

Мы используем вычисления на многообразиях для того, чтобы распространить хорошо знакомые идеи математического анализа на множества, которые отличаются от евклидова пространства. Основная идея вычислений проста. Это метод линейной аппроксимации нелинейных объектов: так, мы рассматриваем прямые линии как линейные аппроксимации кривых, плоскости как линейные аппроксимации поверхностей и т. д. и т. п. Именно эта простая идея лежит в основе всех приводимых здесь вычислений на многообразиях. Как мы увидим в дальнейшем, проще рассматривать саму линейную аппроксимацию, например касательную, чем ее представление — тангенс угла наклона этой касательной. Мы неоднократно будем переходить от обсуждений тех или иных объектов к обсуждению их представлений. Иногда есть смысл различать объекты и их представления, а иногда в этом нет необходимости.

В книге дается неформальное изложение математического аппарата, в котором очень важную роль играют тщательно выполненные диаграммы. Хорошо построенные диаграммы не только полезны для создания интуитивного представления о предмете обсуждения, они часто служат неплохим алгоритмом, позволяющим получить корректное доказательство того или иного утверждения, а подчас и выполнить определенные расчеты. Этот подход к теории относительности, впервые предложенный Дж. Л. Сингом, интенсивно используется в великолепной популярной книге Тейлора и Уилера [40].

Насколько подробно мы собираемся рассматривать те или иные разделы физики? Нам предстоит серьезно разобраться в теории относительности, но, к сожалению, мы не сможем достаточно полно обсудить проблемы астрофизики. Наше обсуждение специальной теории относительности будет настолько основательным и критическим, насколько я смогу его таковым сделать. Однако в общую теорию относительности мы не сможем погрузиться столь же глубоко, поскольку связанные с ней расчеты исключительно сложны и требуют специальных навыков, для приобретения которых понадобился бы годовой курс. Поэтому я хочу, чтобы читатель понял лишь конечный результат общерелятивистских расчетов, а именно, что такое пространство-время. Вы узнаете, как описывать ход часов, свободные частицы и световые сигналы в любом наперед заданном пространстве-времени. Но уравнения, которые описывают само пространство-время и его эволюцию, останутся вне нашего поля зрения. Они, конечно, будут упоминаться. Более того, мы даже проанализируем кое-какие интуитивные идеи, касающиеся этих уравнений. Однако серьезные трудности технического характера все же не позволят нам работать с ними на должном уровне. Мы сможем продемонстрировать применение этих уравнений только в случае однородных и изотропных кос-

Идея наших вычислений

Диаграммы

Насколько мы углубимся в физику?

мологических моделей. К тому моменту нам уже будет по силам построить стандартные космологические модели, исследовать их свойства и обсудить принципы их отбора, но не удастся сделать ни шага за пределы этого ограниченного класса. Однако такое ограничение не столь серьезно, как может показаться, поскольку мы так мало знаем о Вселенной, что не вправе выбирать какую-нибудь более обстоятельную модель¹⁾.

Симметрия

Через все наши рассуждения красной нитью проходит идея симметрии. Вся математика специальной теории относительно-сти — это фактически просто исследование таких математических структур, которые совместимы с лоренцевой симметрией. Для практики мы обсудим также симметрию более распространенных физических явлений, таких, как волны на воде. Построенные нами космологические модели будут обладать самой высокой симметрией, совместимой с наблюдаемой эволюционной природой Вселенной. В то же время существует и противоположная версия, согласно которой симметричные ситуации настолько специфичны, что их возникновение весьма маловероятно. Этот конфликт между порядком и хаосом насчитывает уже не одно столетие. Что касается нашей Вселенной, то, как показывают наблюдения, в больших масштабах она очень симметрична. Кроются ли корни этой симметрии в известных физических законах или она является следствием особых начальных условий, пока не известно.

Эта книга необычна в нескольких отношениях. Я излагаю здесь специальную теорию относительности, опираясь на наблюдения за ходом движущихся часов, а не на предположение о существовании лоренцевой симметрии, как это обычно принято. В результате такие проблемы, как парадокс близнецов, мы сможем разрешать, апеллируя к эксперименту, а не к логическим построениям. Новым элементом является и довольно подробное обсуждение принципа относительности для волн на воде. В результате этого обсуждения возникает еще одна реализация представления о релятивистской симметрии.

Задачи

Важной составной частью книги является большое число содержащихся в ней задач. Они охватывают весь спектр сложности от самых элементарных до очень трудных, а некоторые задачи требуют известной находчивости. Я снабдил каждую за-

¹⁾ С этим замечанием автора нельзя согласиться. Несмотря на кажущуюся скудость наблюдательных данных, получаемой из космоса информации оказывается достаточно не только для того, чтобы осуществить выбор между различными вариантами космологических моделей (предметом дискуссии является ранняя Вселенная), но и для того, чтобы отвергнуть некоторые из моделей элементарных частиц. — *Прим. ред.*

дачу оценкой в баллах по шкале сложности, которую разработал Д. Нут:

- (00) задачи, которые решаются в уме;
- (10) простые задачи, требующие на решение около минуты;
- (20) задачи средней трудности, на которые можно потратить от десяти до двадцати минут;
- (30) задачи умеренной трудности и сложности, на решение которых нужно, скажем, два часа;
- (40) тема курсовой работы.

Как пользоваться книгой

Сначала это покажется сложным, но сначала все сложно.

Мусаши

Как я могу сказать, что я думаю, пока размышляю о том, что говорю.

Э. Форстер

Многие читатели будут использовать эту книгу для самостоятельной работы. Поэтому мне хотелось бы высказать свои соображения по поводу стратегии и тактики изучения изложенного в ней материала, т. е. по поводу того, что при работе в аудитории должен был бы определить преподаватель.

Физика

Эта книга по физике, поэтому естественно предположить, что читатель ставит своей целью овладеть искусством физиков. Это искусство, как мне кажется, имеет три аспекта. Во-первых, вы должны быть способны рассмотреть ситуацию и составить мысленную картину происходящего. Нужно уметь «мысленно повернуть ее». Во-вторых, необходимо уметь построить соответствующую математическую модель, т. е. перевести физические закономерности на язык математических объектов и операций над ними. В-третьих, вы должны быть в достаточной степени подготовлены технически, чтобы умело работать в рамках построенной математической модели. Все три момента важны, хотя легко обучать только третьему. Однако не будем забывать и два других аспекта.

Новый язык

Важно понять, что мы собираемся изучать совершенно новый язык. В новом языке не только присваиваются новые названия знакомым вещам, но и используются новые идеи, которые трудно сформулировать другим способом. Сначала вам будет нелегко, поскольку при использовании нового языка меняется смысл некоторых привычных слов и им придается новое точное значение. Например, мы используем слово «часы» для обозначения понятия, которое несколько отличается от того, что мы обычно понимаем под часами, но все же настолько близкого к нему, что было бы безумием придумать какое-то новое слово. Кроме того, многие понятия, которые кажутся простыми при традиционном подходе, перестают быть таковыми при использовании нового языка. В частности, требуется строго оговорить, что мы понимаем под скоростью света, скоростью хода движущихся часов и т. д. В самой структуре языка неизбежно содержится существенная информация о том, что описывается языком. Язык — это часть теории. Будьте бдительны! Старый язык неизбежно привязывает нас к старой теории.

Полезно уметь должным образом рассматривать гипотетические ситуации. Если я спрошу вас: «Предположим, что в метре от нас находятся четыре шестиногие собаки. Сколько всего у них ног?» Поддержите ли вы игру? Или же скажете: «Этого не может быть»? Пример тривиальный, но не тривиальна проблема. Понятия означают то, что мы условились под ними понимать, даже если мы дали им знакомые имена. В принципе легко согласиться с этим утверждением. Однако будьте осторожны, если рассматриваете понятия, знакомые по повседневной практике. Вспомните, что даже очень умные люди испытывают трудности при рассмотрении величины $\sqrt{-1}$. Аналогичные проблемы возникнут у вас при чтении книги.

Нетривиальную область, требующую обдуманного подхода, образует совокупность вопросов, связанных с представлением физических объектов. Чтобы эффективно работать с объектами физического мира, обычно находят соответствующее математическое представление. Математические объекты в свою очередь также могут обладать представлениями, и в таком случае становится важным, до какой степени можно провести различие между объектом и его представлением.

Представления

Пример

Квадратная матрица представляет собой числовую таблицу, которую можно использовать, чтобы задать линейное преобразование. Соответствующий детерминант может рассматриваться или как определенный набор операций над числами, входящими в таблицу, или как число, показывающее, во сколько раз линейное преобразование изменяет объем. Теорема о том, что детерминант произведения равен произведению детерминантов, в первом случае доказывается в результате весьма запутанных вычислений, а во втором представляет собой очевидный результат.

Аналогичная проблема возникает при использовании тензоров. Физики обычно рассматривают тензоры как числовые таблицы. Однако сейчас уже стало понятно, что удобнее иметь дело с линейными операторами, которые представляются этими числами; именно такой точки зрения я придерживаюсь в книге.

Один из путей абстрагироваться от конкретных особенностей выделенного представления заключается в рассмотрении нескольких представлений. Мы в таких случаях, как правило, будем приводить одновременно графическое и алгебраическое представления. На пространственно-временных диаграммах мы найдем объекты, которые являются представлениями тензоров. Я думаю, что не возникнет недоразумений, если эти графические представления я также буду называть тензорами.

Что нужно знать

Хотя для чтения этой книги достаточно знания анализа и линейной алгебры, затронутые в ней вопросы потребуют от читателя умения абстрактно мыслить. Студенты, на которых рассчитана книга, как раз постигают это искусство. Приводимый в книге материал дает хорошую возможность попрактиковаться в этом. Цель абстрагирования состоит в том, чтобы найти общее в различных частных ситуациях и затем рассмотреть это общее вне зависимости от конкретной ситуации. Такой подход, если он применяется должным образом, углубляет наши представления об объекте исследования. Но будьте осторожны. Чрезмерное абстрагирование — очень распространенная ошибка. Помните о задаче, которую рассматриваете, и не давайте возможности абстрактным объектам начать жить самостоятельной жизнью.

Абстрагирование

Необходима конкретность

Успешный способ изучения нового абстрактного материала состоит в постепенном переходе от простого к сложному. Благодаря поразительному успеху геометрии Евклида как дедуктивной системы людям стало казаться, что наука должна строиться по дедукции от общих правил к конкретным примерам. Это неверно, в особенности если речь идет о процессе обучения. Наш мозг, по-видимому, хорошо справляется с воссозданием целого по элементам мозаичной картины и с обобщением конкретных представлений. Построение же цепочки логических рассуждений, вероятно, весьма неестественный для нас вид деятельности.

Хорошим примером постепенного перехода от специфических особенностей и частных примеров к некоторому общему утверждению, который я всячески пропагандирую, может служить использованный в книге способ введения касательного вектора. Прежде всего дается определение касательных к прямым, а затем — к кривым в векторном пространстве; только после этого я перехожу к рассмотрению касательных к кривым на многообразиях. Нужно стараться аналогичным образом подходить к рассмотрению и более простых проблем. Когда вы следите за ходом доказательства или пытаетесь понять новое определение, придумывайте примеры, нарисуйте несколько диаграмм. Нельзя ожидать, что вы усвоите материал, просто прочитав эту книгу. Необходимо работать активнее. Делайте заметки на полях, проделывайте промежуточные выкладки, рисуйте больше диаграмм, решайте и решайте задачи и даже придумывайте новые задачи. Часто оказывается полезным проделать одно и то же вычисление дважды. Особенно полезно предварительно решить задачу графически. Ошибки, допущенные при графическом решении, обычно значительно более поучительны, чем ошибки в алгебраических выкладках.

Наглядное представление

В этой книге широко используется графический материал. Диаграммы не только дополняют численные представления, но

во многих случаях играют важную роль при обосновании тех или иных утверждений. Хорошо выполненная диаграмма часто подсказывает алгоритм доказательства, в то время как детальное проведение доказательства может оказаться громоздким и утомительным. Обычно мы используем двумерные диаграммы, где по одной оси откладывается пространственная, а по другой — временная координата. Двумерность определяется двумерностью поверхности листа бумаги и соображениями наглядности. Обычно основные особенности могут быть продемонстрированы на двумерных диаграммах, а распространение полученных результатов на случай большего числа измерений может служить полезным и несложным упражнением.

Тензорный анализ имеет репутацию запутанного и сложно-го предмета отчасти благодаря тому, что лишь немногие могут освоить представление тензора в виде таблицы. Вместо этого я предлагаю начать с геометрического представления тензоров. Например, метрический тензор на пространственно-временной диаграмме представляется гиперболой. При произвольной ориентации гиперболы ее уравнение имеет достаточно сложный аналитический вид, хотя сама по себе гипербола совсем не сложное понятие. Как мне кажется, точно так же обстоит дело и с тензорами. Проявив определенный оптимизм, я попытался обойти сложности, связанные с использованием тензоров, путем введения геодезического квадрата для обсуждения кривизны.

Некоторые люди не принимают всерьез диаграммы. Я уверяю вас, что здесь все диаграммы построены очень тщательно и предназначены для того, чтобы их тщательно изучали. Если вы относитесь к числу людей, которые обычно не обращают особого внимания на диаграммы, то рискуете упустить из виду значительную часть материала этой книги.

Проработав предложенный материал, вы должны сами оценить, насколько хорошо вы его понимаете и умеете решать задачи. Отыскивайте свои слабые места и работайте над ними. Будьте осторожны с абстракциями. Слишком мало лучше, чем слишком много. Чрезмерное абстрагирование обыкновенно проявляется в использовании жаргона и некритическом манипулировании словами. Такая плохая привычка вырабатывается, когда пытаются изучить слишком много за слишком короткое время.

Какие еще книги я могу порекомендовать? Чтобы познакомиться с основами астрономии, вы должны прочитать одну из посвященных ей популярных книг, скажем очень хорошую книгу Фреда Хойла «Астрономия и космология». Основы специальной теории относительности хорошо изложены в книге Тейлора и Уилера. Если вы хотите узнать больше о проблемах и задачах, я посоветовал бы обратиться к книгам Вейнберга и По-

[См. второй эпиграф.]

Трудности

Решайте задачи

Другие книги

лиа. Вслед за ними рекомендовал бы книгу Риндлера и Оханиана, где примерно на том же уровне дается традиционное изложение материала. Для дальнейшей работы может служить книга Мизнера, Торнера и Уилера, имеющая характер энциклопедии; в ней приводится и список литературы¹⁾.

¹⁾ Советскому читателю можно рекомендовать следующие книги: Новиков И. Д. Эволюция Вселенной. — М.: Наука, 1979; Шкловский И. С. Звезды, их рождение, жизнь и смерть. — М.: Наука, 1974; Зельдович Я. Б., Новиков И. Д. Строение и эволюция Вселенной. — М.: Наука, 1975; Пиблс П. Структура Вселенной в больших масштабах. — М.: Мир, 1983. — Прим. ред.

I. Специальная теория относительности

Информация о происходящих во Вселенной процессах поступает к нам в основном двумя путями. С неба на нас льются потоки света. Земля изобилует горными породами, естественная радиоактивность которых представляет в наше распоряжение своего рода часы, запущенные в далеком прошлом. Чтобы разобраться в этой информации, нужна теория, описывающая ход часов и распространение световых сигналов. Именно такой теорией, пригодной в случаях, когда гравитационные эффекты несущественны, является специальная теория относительности.

В этой книге специальная теория относительности будет излагаться с трех различных точек зрения. В настоящей главе при построении теории мы будем исходить из нескольких фундаментальных постулатов, комбинируя их с результатами наблюдений за движущимися часами. При таком изложении симметрия теории относительно преобразований Лоренца не предполагается заранее, а вытекает из сделанных предположений. Этот подход аналогичен построению планиметрии в духе Евклида. В следующей главе мы введем математический аппарат, который позволит нам рассмотреть специальную теорию относительности снова, уже на языке аналитической геометрии. Наконец, в гл. III мы вернемся к исходным положениям, переопределив и расширив понятие свободной частицы, для того чтобы можно было перейти к изложению общей теории относительности.

В своем изложении специальной теории относительности мы не придерживаемся исторической последовательности. Это не описание того, как специальная теория относительности была открыта в действительности, и не соображения о том, как она должна была бы быть открыта. Надеюсь, что принятая форма изложения позволит простым и эффективным способом изучить специальную теорию относительности и понять ее тонкую логическую структуру.

1. Структура пространства-времени

Для меня нет ничего более таинственного, чем пространство и время и тем не менее ничто не трогает меня меньше, чем они, так как я никогда не думаю о них.

Чарлз Лэмб

Предмет этой книги — пространство-время и его структура. Движение тел происходит в пространстве-времени, которое представляет собой совокупность всех возможных положений тел во все возможные моменты времени. Пространство-время обладает структурой, которая проявляется в характере законов, управляющих движением объектов. Под структурой пространства-времени мы будем понимать наиболее общие закономерности в поведении свободных частиц, часов и световых сигналов. Знание этой структуры позволяет детально исследовать проблему хода часов и описать эксперименты, доказывающие справедливость теории. Такой теорией, предсказания которой совпадают со всеми проведенными до настоящего времени экспериментами, и является специальная теория относительности. Основные понятия, которыми мы пользуемся при обсуждении специальной теории относительности, определяются в ходе описания структуры пространства-времени.

В настоящей главе книги обсуждается структура пространства-времени в малых областях, в которых гравитацией можно пренебречь. В гл. II развивается математический аппарат, который в гл. III позволит нам обобщить теорию таким образом, чтобы можно было включить в рассмотрение гравитацию, и сделает возможным последовательное обсуждение самогравитирующей Вселенной в гл. IV.

Событие

Основную роль в пространстве-времени играет понятие *события*. Событие определяется местом, где оно произошло, и временем, когда оно произошло. Событие в пространстве-времени то же самое, что точка в евклидовой геометрии. Мы говорим о событии, а не о точке, чтобы подчеркнуть, что понятие события объединяет в себе как пространственное положение, так и время. Понятно, что реальные события характеризуются определенной протяженностью. Однако при нашем усредненном описании этой внутренней структурой мы пренебрегаем.

[О крупномасштабном усреднении см. введение.]

Мы будем изображать физические события и их связи с помощью четырехмерной картины, которую называют *пространственно-временной* диаграммой. Поскольку во многих практически интересных случаях зависимостью от некоторых координат можно пренебречь, наши пространственно-временные диаграммы чаще всего будут двумерными. Такие диаграммы, где по одной оси отложена пространственная координата, а по другой — время, изображают события, которые происходят вдоль определенной линии в пространстве. В случае большего числа измерений мы будем выполнять рисунки в перспективе либо приводить сечения пространственно-временной диаграммы, либо проецировать диаграмму на одну из координатных плоскостей (пренебрегая остальными измерениями). Объекту, существующему в течение продолжительного времени, в пространстве событий соответствует непрерывная кривая, которая называется *мировой линией*.

Пространственно-временная диаграмма

Пример

Рассмотрим пространственно-временную диаграмму, изображающую движение автомашин по улице, вдоль которой расположены светофоры. Здесь мы имеем, например, такие события: на данном светофоре загорается красный сигнал в тот момент, когда к нему подъезжает определенный автомобиль, и т. д. Пространственно-временная диаграмма приведена на рис. 1.1. На нем отмечены некоторые события (изменение сигнала светофоров, начало движения автомобилей). Изображены мировые линии двух автомобилей и двух светофоров. Используются обычные временная и пространственная координаты.

Надеюсь, что приведенный пример позволяет понять, что такое пространственно-временная диаграмма. Это графический аналог обычного расписания движения, например расписания движения самолетов. Важно то, что такая простая идея при надлежащем развитии может стать мощным инструментом исследования, необходимым для нашей последующей работы. Из определения события следует, что совсем обязательно, чтобы в данной точке пространства в данный момент действительно что-либо произошло. Ситуация полностью аналогична той, с которой мы сталкиваемся в евклидовой геометрии, где большая часть точек плоскости не принадлежит никаким геометрическим фигурам.

[Это проекция. На самом деле автомобили не сталкиваются.]

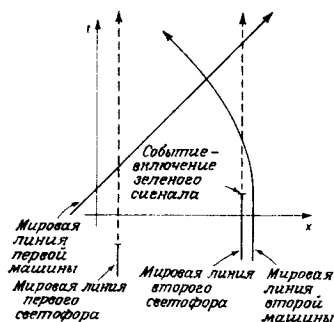


Рис. 1.1.

Двумерная пространственно-временная диаграмма, изображающая движение автомобилей по улице. Зависимость от координаты по оси, проведенной поперек улицы, не учитывается.

Чтобы пояснить, в каком смысле пространственно-временная диаграмма представляет физические события, нужно дать ее операциональное определение, т. е. перечислить последовательность физических измерений и математических операций, необходимых для ее построения. Однако мы отложим обсуждение этих вопросов до следующего раздела, поскольку понятие операций не имеет смысла, пока мы не рассмотрели некоторых проблем, касающихся часов и световых сигналов. Следует также четко представлять, что сама по себе пространственно-временная диаграмма — это математическая модель, основанная на понятии линейного векторного пространства, понятия, с которым мы будем часто сталкиваться в этой книге. Линейное векторное пространство является обобщением привычного трехмерного пространства, с которым мы работаем в механике и электродинамике.

Множество элементов мы будем называть линейным векторным пространством, если выполняются определенные аксиомы. Должно быть задано правило сложения двух элементов векторного пространства, в результате которого получается элемент того же пространства. Должна быть определена операция умножения вектора на действительное число. Требуется, чтобы при этом выполнялись обычные свойства операций сложения и умножения.

Линейное векторное пространство

Аксиомы линейного векторного пространства. Пусть a , b и c — векторы, принадлежащие заданному линейному векторному пространству, и пусть k и m — действительные числа; тогда операции сложения и умножения на число должны удовлетворять следующим требованиям:

$$a + (b + c) = (a + b) + c, \quad (1.1)$$

$$a + b = b + a. \quad (1.2)$$

Должен существовать нулевой вектор, который обозначается символом 0 , такой, что для любого вектора a

$$a + 0 = a. \quad (1.3)$$

Для каждого вектора a должен существовать вектор $(-a)$, такой, что

$$a + (-a) = 0. \quad (1.4)$$

Умножение на число должно удовлетворять следующим требо-

ваниям:

$$(km)a = k(ma), \quad (1.5)$$

$$(k + m)a = ka + ma, \quad (1.6)$$

$$k(a + b) = ka + kb. \quad (1.7)$$

Наконец, умножение на единицу не должно изменять вектор

$$1a = a. \quad (1.8)$$

Пример 1

Рассмотрим множество числовых пар (x, y) , где x и y — произвольные действительные числа. Правила сложения

$$(x, y) + (z, w) \equiv (x + z, y + w) \quad (1.9)$$

и умножения на число

$$k(x, y) \equiv (kx, ky) \quad (1.10)$$

удовлетворяют сформулированным выше аксиомам и задают на множестве пар чисел $\mathbb{R}^2$ структуру линейного векторного пространства. Для рассматриваемого множества это обычное определение операций сложения и умножения на число. Его легко обобщить на случай произвольного числа измерений.

[$\mathbb{R}$ — стандартное обозначение для множества всех действительных чисел. $\mathbb{R}^2$ используется для обозначения декартовой плоскости, т.е. множества всех числовых пар.]

Пример 2

Следующие правила:

$$[x, y] + [z, w] \equiv \quad (1.11)$$

$$\equiv \left[\frac{x(z^2 + w^2) + z(x^2 + y^2)}{(x + z)^2 + (y + w)^2}, \frac{y(z^2 + w^2) + w(x^2 + y^2)}{(x + z)^2 + (y + w)^2} \right],$$

$$k[x, y] \equiv \left[\frac{x}{k}, \frac{y}{k} \right] \quad (1.12)$$

также превращают множество $\mathbb{R}^2$ в линейное векторное пространство. Мы используем квадратные скобки, чтобы подчеркнуть отличие сформулированных выше правил от обычных. Совершенно не очевидно, что эти правила удовлетворяют аксиомам линейного векторного пространства. Действительно, мы

[Понятие линейного векторного пространства является фундаментальным. Чтобы познакомиться с ним ближе, вы должны проверить, что рассмотренные выше пространства действительно удовлетворяют перечисленным аксиомам. Приведите другие примеры. Подробное изложение теории линейных пространств можно найти в книге [21].]

должны исключить из $\mathbb{R}^2$ элемент $(0,0)$ и добавить бесконечно удаленную точку. Теперь нулевым вектором является любая пара чисел $[a, f]$, для которой $a^2 + f^2 = \infty$. Все такие пары должны рассматриваться как один и тот же вектор. Точка $(0, 0)$ должна быть исключена из $\mathbb{R}^2$, чтобы правило сложения (1.11) имело смысл. Несмотря на то что второй пример выглядит несколько надуманным, он представляет собой вполне естественный для рассматриваемого множества способ задания структуры линейного векторного пространства. В разд. 15 мы получим другое более удобное представление.

Эти векторы похожи на хорошо знакомые нам 3-векторы, за исключением того, что все они выходят из начала системы координат. Такие векторы иногда называют закрепленными (bound) векторами, чтобы отличить их от свободных векторов, которые могут начинаться в любой точке.

Наши пространственно-временные диаграммы должны изображаться в четырехмерном линейном векторном пространстве, которое обычно представляет собой множество упорядоченных четверок чисел со стандартной структурой. Чтобы это линейное векторное пространство было для нас интересным, должно существовать правило, которое связывает определенные в нем математические объекты с объектами реального физического мира. (Связь между математическим и физическим миром достаточно сложна. Отложим обсуждение ее деталей до разд. 4, а здесь просто предположим, что такие связи существуют.) Такого рода взаимоотношения между множествами называются *отображениями*. Мы будем использовать слово «отображение» в смысле математического термина. Принятая в математике система обозначений для отображений также очень полезна, и ею следует чаще пользоваться.

Отображения

Отображение: математическая операция, ставящая в соответствие каждому элементу одного множества некоторый элемент другого множества. В математике предложение « f отображает элементы множества A на множество B » записывается в следующем сокращенном виде:

$$f: A \rightarrow B. \quad (1.13)$$

Кроме того, мы будем использовать «ограниченную» стрелку $\mapsto$ либо для обозначения элементов второго множества, либо для

задания отображения в явном виде. Так, предложим «элемент множества B , который получается из элемента a множества A при отображении f , будет обозначаться символом $f(a)$ » мы записываем так:

$$f: A \rightarrow B; a \mapsto f(a). \quad (1.14)$$

Пример

Отображение g , состоящее в удвоении действительных чисел, записывается следующим образом:

$$g: \mathbb{R} \rightarrow \mathbb{R}; x \mapsto 2x. \quad (1.15)$$

Отображение является обобщением понятия функции. При этом функция представляет собой отображение множества действительных чисел в множество действительных чисел.

Вторым по важности после события является понятие *свободной частицы*. В физике под свободной частицей понимают тело, размеры которого малы настолько, что ими можно пренебречь, и на которое не действуют никакие силы. Мировые линии таких частиц образуют особый класс среди всех возможных мировых линий в пространстве-времени. Вспомним сформулированный Галилеем закон, согласно которому свободные частицы движутся с постоянной скоростью. На пространственно-временной диаграмме движению с постоянной скоростью соответствует прямая мировая линия. Мы включили этот закон в нашу теорию, используя в качестве модели пространства-времени линейное векторное пространство. Введем следующий постулат.

Свободные частицы

Постулат свободных частиц: существует отображение событий пространства-времени на пространственно-временную диаграмму, при котором мировые линии свободных частиц изображаются прямыми линиями.

Прямая линия — это математическое понятие, хорошо определенное в линейном векторном пространстве. Под прямой лини-

ей мы понимаем множество векторов v , таких, что

$$v = a + sb, \quad (1.16)$$

где a и b — заданные векторы, а s — произвольное действительное число.

[Первичные понятия были рассмотрены во введении.]

Понятие свободной частицы представляет собой первичное понятие в нашей теории и потому не может быть определено. Можно только продемонстрировать, что мы под этим понимаем. Вот почему изучение физики обязательно предполагает работу в лабораториях.

Может сложиться впечатление, что постулат свободных частиц представляет собой утверждение, которое касается только отображения и ничего не говорит о реальном мире. Это неверно. Обратимся к рис. 1.2. На нем изображены две кривые в векторном пространстве. Пусть эти кривые представляют собой образы двух мировых линий при некотором отображении. Можно ли сказать, что соответствующие частицы не являются свободными частицами? Вовсе нет. Линии могли быть искривлены как раз из-за неудачного выбора отображения. С другой стороны, рассмотрим отображение, в результате которого получен рис. 1.3. Две различные прямые линии могут пересечься только один раз, поэтому никакое отображение не выпрямит одновременно обе эти кривые. Так как не всякие множества мировых линий могут быть отображены в прямые линии одновременно, предположение о существовании прямых мировых линий свободных частиц не бессодержательно.

Отображение, при котором движение свободных частиц изображается прямыми, далеко не единственно. Существует огромное количество таких отображений. Как разобраться в этой ситуации, мы увидим в разд. 4, который посвящен инерциальным системам отсчета.

Следующее важное понятие, которое вводится в пространстве-времени — это понятие часов. Часы представляют собой устройство для измерения отрезков мировой линии. В них может использоваться какое-либо повторяющееся явление, например качание маятника или регулярный стандартный временной интервал. Таким интервалом может быть время жизни нестабильной частицы. В любом случае скорость хода часов должна определяться некоторым внутренним свойством. Действие большей части современных часов основано на атомных и ядерных процессах. Когда мы будем рассматривать аналог принципа относительности для волн на воде, мы столкнемся с другой физической реализацией понятия часов, имеющих

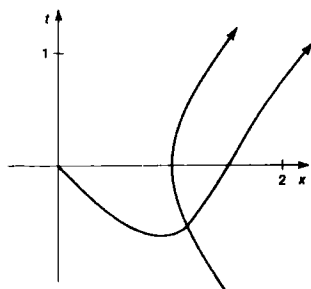


Рис. 1.2

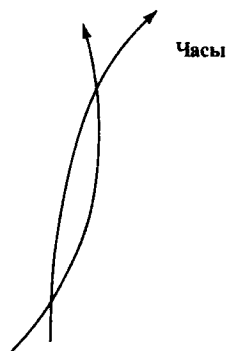


Рис. 1.3

некоторую внутреннюю определенную скорость хода. Чтобы данный алгоритм измерения отрезков мировой линии мог ассоциироваться с часами, он должен удовлетворять следующим двум постулатам.

Постулат универсальности хода часов: относительная скорость хода двух различных часов, переносимых вдоль заданной мировой линии, не зависит от выбора точки и мировой линии в пространстве-времени.

Это еще один постулат, который можно было бы назвать постулатом крупномасштабного усреднения. Он утверждает, что не нужно создавать различных теорий для часов различной конструкции, например, атомных и ядерных часов. Легко понять физическое содержание сформулированного постулата, а развитие технологии дает возможность постановки новых экспериментов, позволяющих проверить его обоснованность. В одном из будущих экспериментов, который сейчас обсуждается, предполагается послать группу различных часов к Солнцу, чтобы проверить, влияет ли присутствие массивного тела на их относительную скорость хода.

Второй постулат, с одной стороны, описывает ход часов, а с другой — накладывает дополнительные ограничения на отображение физических событий в линейное векторное пространство, в результате которого мы получаем пространственно-временную диаграмму.

Постулат равномерности хода часов: существует отображение физического пространства событий в векторное пространство, при котором мировые линии свободных частиц изображаются прямыми, а ход часов равномерен.

Равномерность хода часов означает, что величины интервалов, которые мы измеряем с помощью часов, должны согласовываться с линейной структурой пространства. Отрезок мировой

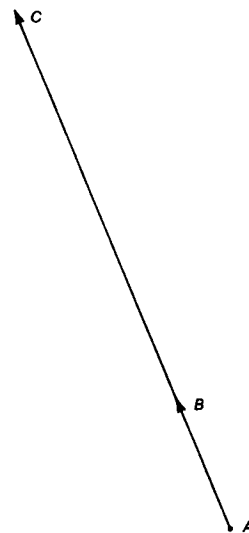


Рис. 1.4

Если вектор AC в k раз больше вектора AB , то часы, которые переместились вдоль отрезка AC мировой линии, отсчитают время, в k раз большее, чем время, которое они отсчитали бы при перемещении вдоль AB . На рисунке k примерно равно 4.

[На этих рисунках умышленно не изображены координатные оси, чтобы подчеркнуть независимость введенных понятий от конкретного выбора системы координат.]

Инерциальная система отсчета

[Числа в круглых скобках указывают относительную сложность задачи; соответствующая шкала приведена на стр. 25.]

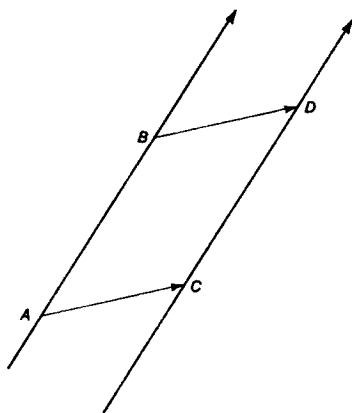


Рис. 1.5.

Если векторы AC и BD равны, то часы, которые переместились вдоль отрезка AB мировой линии, покажут то же самое время, что и часы, переместившиеся вдоль отрезка CD . Это один из аспектов равномерности хода времени.

линии, имеющий удвоенную длину, должен соответствовать в два раза большему временному интервалу. Промежутки времени, измеренные на противоположных сторонах параллелограмма, должны быть равны (рис. 1.4 и 1.5).

Рассмотрение свободных частиц приводит лишь к представлению о линейной структуре векторного пространства. Часы вносят дополнительно понятие параллельности. Обратите внимание, что мы не используем понятие перпендикулярности, да в векторном пространстве и нельзя ввести его естественным образом. Существует много хроноструктур, которые согласуются с приведенными выше постулатами. Выяснить, какая из них соответствует реальным часам, должен эксперимент. Представление реальных процессов в виде пространственно-временной диаграммы, при которой свободным частицам соответствуют прямые мировые линии и ход часов равномерен, называется *инерциальной системой отсчета*.

ЗАДАЧИ

1.1. (15) Перечислите некоторые идеализации, использованные при построении рис. 1.1.

1.2. (15) Если на рис. 1.1 светофоры расположены на расстоянии квартала друг от друга, то достаточно ли времени у водителя второй машины, чтобы среагировать на смену сигнала светофора?

1.3. (24) Используя пространственно-временную диаграмму, выясните, почему нельзя синхронизовать светофоры, если движение по улице двустороннее, а светофоры расположены на различных расстояниях друг от друга.

1.4. (12) Убедитесь, что множество, рассмотренное в примере 1 на стр. 35, действительно является векторным пространством, т. е. что для него выполняются соответствующие аксиомы.

1.5. (18) Покажите, что преобразование координат

$$(t, x) \mapsto (u, v) = \left(\sqrt{x^2 - t^2}, \operatorname{ar th} \left(\frac{t}{x} \right) \right)$$

отображает кривые, изображенные на рис. 1.2, в прямые.

1.6. (35) Проведите еще несколько мировых линий свободных частиц в криволинейном представлении, которым мы пользовались при построении рис. 1.2. Какой вид, в частности, будет иметь картина вблизи начала координат?

1.7. (15) Покажите, что никакое преобразование координат не превратит в прямые семейство окружностей, изображенное на рис. 1.6. Уравнение семейства имеет вид

$$(x - a)^2 + (y - b)^2 = 1,$$

где различным окружностям соответствуют различные значения a и b .

1.8. (30) Найдите преобразование координат, которое отображает семейство парабол, изображенных на рис. 1.7,

$$x = a + bt + t^2$$

в прямые линии.

1.9. (15) Пусть электроны движутся под действием однородного электрического поля. Можно ли рассматривать их как свободные частицы? Что можно сказать о смеси частиц с различным отношением заряда к массе?

1.10. (12) Рассмотрите пространственно-временные диаграммы, изображенные на рис. 19.6, 19.7 и 29.15. Опишите ситуацию в различные моменты времени.

1.11. (34) Придумайте трехмерный аналог странного векторного пространства, рассмотренного в примере 2 на стр. 35.

2. Часы

Если бы кое-кто не совался в чужие дела, — хрипло проворчала Герцогиня, — Земля бы вертелась быстрее.

Льюис Кэррол

Фундаментальная структура пространства-времени, рассмотренная в разд. 1, не определяет однозначно поведение набора идентичных часов, которые находятся в данной точке пространства-времени. При обсуждении этого вопроса воспользуемся инерциальной системой отсчета. Нам потребуется эффективный способ представления, дающий возможность описать поведение любых часов, ход которых согласован с нашими постулатами. Оказывается, что такое представление легко найти. Для начала применим его к хорошо знакомому случаю ньютоновой механики. Это позволит нам попрактиковаться на известном материале, прежде чем переходить к ново-

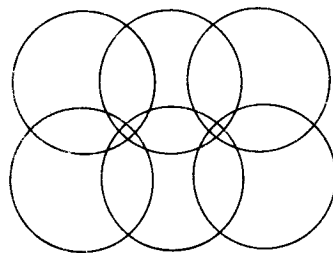


Рис. 1.6

[Оценка сложности этой задачи дана в предположении, что вам известны результаты задачи 1.8, и не учитывает сложности задачи 1.8.]

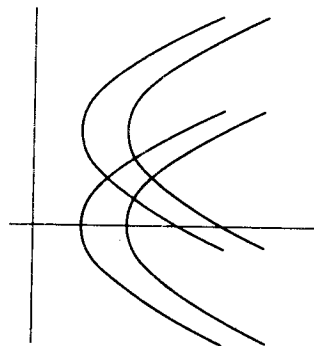


Рис. 1.7

му и достаточно тонкому вопросу о поведении часов в специальной теории относительности. Поскольку инерциальная система отсчета не единственна, необходимо рассмотреть ход различных часов в различных инерциальных системах отсчета. В результате мы придем к важному понятию ковариантности, которое играет фундаментальную роль во всем тензорном анализе.

Представление хроноструктуры

Каким образом мы будем представлять поведение различных часов? Предположим только, что ход часов согласуется с нашими постулатами и что мы каким-то образом выбрали некоторую инерциальную систему отсчета, в которой будем работать. Хотя постулат равномерности хода часов накладывает ограничения на поведение часов, которые перемещаются вдоль любой выделенной мировой линии, а также вдоль всех мировых линий, параллельных выделенной, он ничего не говорит о различных семействах параллельных мировых линий, т. е. о семействах часов, движущихся с разными скоростями. Наши различные теории характеризуются различным поведением движущихся часов.

Каким образом учесть такую дополнительную информацию? Мы можем сделать это непосредственно. Рассмотрим различные часы, которые пролетают через начало отсчета (нулевой вектор) со всевозможными скоростями. На каждой мировой линии отметим точку, которой соответствующие часы достигли за единицу времени. Множество $\mathcal{S}$ всех таких отметок для всевозможных часов и будет нашим представлением поведения часов.

Раз мы знаем множество $\mathcal{S}$ всех событий, которые отстоят от начала отсчета на единицу времени, мы имеем полное описание поведения любых часов. Постулат равномерности хода часов позволяет переносить стандартный временной интервал куда угодно и таким образом измерять любые промежутки времени, как изображено на рис. 2.1. Подчеркнем, что знание $\mathcal{S}$ в одной инерциальной системе вполне достаточно для полного описания поведения часов. Не нужно каждый раз рассматривать систему отсчета, где часы «покоятся». Мы еще даже не определили, что понимается под словом «покоится». Обратите внимание также на то, что наше представление доступно экспериментальному изучению. Физик-экспериментатор в состоянии в пределах ошибки измерения осуществить прямое измерение множества $\mathcal{S}$.

Применим сформулированную выше идею для описания поведения часов в ньютоновой механике. Будем называть такие

[Мы будем использовать прописные рукописные буквы, например $\mathcal{S}$, для обозначения как множеств, так и хроноструктур, которые они представляют. Если необходимо использовать различные обозначения, то множество можно обозначать как $\mathcal{M}$. Обычно в книгах при рассмотрении метрик используется обозначение ds^2 , а не $\mathcal{S}$.]

[Заметьте, что на рис. 2.1 часы, движущиеся влево, ведут себя не так, как часы, движущиеся вправо.]

Часы, измеряющие абсолютное время

часы часами, измеряющими абсолютное время. Это самосогла-сованная математическая модель, которая соответствует пове-дению реальных часов при низких скоростях. Кроме того,

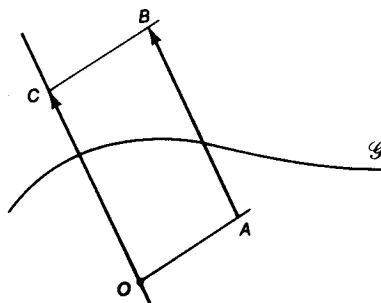


Рис. 2.1
Использование равномерности хода ча-сов для измерения временного интерва-ла AB . Для ситуации, изображенной на рисунке, он равен примерно 1,5 времен-ной единицы.

сравнение с ньютоновой механикой позволяет лучше понять то новое, что вносит в представление о времени специальная тео-рия относительности. Обычные пространственная и временная координаты ньютоновой механики определяют некоторую инерциальную систему отсчета. При этом показания часов не-посредственно даются временной координатой. Обозначим че-рез $\mathcal{A}$ множество событий, которое представляет поведение ча-сов, измеряющих абсолютное время: $\mathcal{A}$ изображается горизон-тальной линией, проведенной на расстоянии, равном единице, от начала отсчета (рис. 2.2). В случае большего числа измере-ний это будет плоскость или гиперплоскость. Соответствующая теория часов самосогласована и даже «приятна» в некото-ром смысле. Мы увидим, что она обладает своей собственной симметрией, которую называют галилеевой симметрией и ко-торая аналогична лоренцевой симметрии специальной теории относительности. В действительности специальная теория от-носительности была придумана потому, что галилеева симмет-рия абсолютного времени оказалась несовместимой с лоренце-вой симметрией электродинамики Максвелла.

Теперь мы сталкиваемся с проблемой, состоящей в том, что наши постулаты не определяют однозначно инерциальную

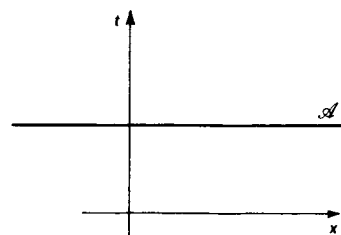


Рис. 2.2

Возможно много представлений

систему отсчета. Глубже этот тонкий вопрос рассматривается в разд. 4, который посвящен специально инерциальным системам отсчета, однако имеет смысл начать его обсуждение уже сейчас. Описание различия между двумя инерциальными системами отсчета, т. е. между двумя отображениями ψ и ψ' в наше векторное пространство V , удобно свести к рассмотрению преобразования пространства V , при котором пространственно-временная диаграмма, получаемая при отображении ψ , переходит в пространственно-временную диаграмму, получаемую при отображении ψ' .

Преобразование: отображение множества на себя.

При таком подходе мы остаемся в рамках нашей математической модели. Отображения ψ и ψ' представляют собой достаточно сложные операции, так как они связаны, с одной стороны, с физическим миром, а с другой стороны, с нашей математической моделью. Если в результате этих отображений мы получаем инерциальные системы отсчета, то на обеих пространственно-временных диаграммах мировые линии свободных частиц будут прямыми. Таким образом, наше преобразование должно переводить прямые линии в прямые. Кроме того, ход часов должен быть равномерным в обеих системах отсчета, следовательно (это не трудно увидеть), параллельные прямые должны переходить в параллельные прямые. Отображения, которые удовлетворяют сформулированным требованиям, называются линейными преобразованиями. Поскольку линейные преобразования важны и будут часто использоваться, приведем их явный вид в случае пространства двух измерений (рис. 2.3). Это следующие преобразования:

Линейные преобразования

[Смысл обозначения с помощью «ограниченной» стрелки $\mapsto$ объяснен на стр. 36.]

[При $k = 2$ отображение (2.1) приобретает вид $(1, 0) \mapsto (2, 0)$, т.е. преобразование переводит точку, которая лежит на оси x на расстоянии, равном единице от начала отсчета, в точку, которая лежит на расстоянии, равном двум.]

[Эти два сдвига отличаются только поворотом на 45° .]

$$\text{растяжения} \quad (x, y) \mapsto (kx, ky); \quad (2.1)$$

$$\text{повороты} \quad (x, y) \mapsto (\cos \theta x + \sin \theta y, \cos \theta y - \sin \theta x); \quad (2.2)$$

$$\text{сдвиги} \quad (x, y) \mapsto \left(\alpha x, \frac{y}{\alpha} \right); \quad (2.3)$$

$$(x, y) \mapsto \frac{1}{\sqrt{1-v^2}} (x - vy, y - vx). \quad (2.4)$$

Любое линейное преобразование является комбинацией пере-

численных выше преобразований. Кроме того, мы можем просто перенести начало координат в другую точку. Линейные преобразования могут быть представлены в виде квадратных матриц.

В различных представлениях множество $\mathcal{S}$ будет иметь разный вид, несмотря на то что оно соответствует одним и тем же часам. На рис. 2.3 показано, каким образом изменяется при линейных преобразованиях множество $\mathcal{A}$, которое представляет поведение часов, измеряющих абсолютное время. Такое изменение представления называется *ковариантностью*. Всюду в этой книге мы будем иметь дело с представлениями, для которых любые линейные преобразования или некоторое подмножество множества линейных преобразований переводит одно допустимое представление в другое допустимое представление. Понятие ковариантности включает в себя такие изменения представления.

Ковариантность

[Возможно, потребуется некоторое время, прежде чем вы поймете идею ковариантности.]

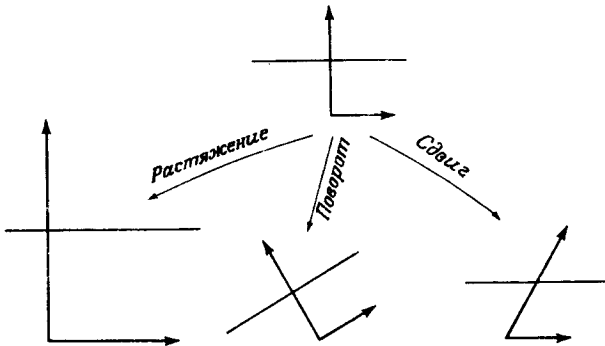


Рис. 2.3

Примеры линейных преобразований в пространстве двух измерений: рас-

тяжение $(x, y) \mapsto (2x, 2y)$, поворот $(x, y) \mapsto \frac{1}{2}(\sqrt{3}x - y, x + \sqrt{3}y)$ и сдвиг $(x, y) \mapsto (x + \sqrt{3}y, y)$.

ЗАДАЧИ

2.1. (10) Найдите явное линейное преобразование, которое переводит перпендикулярные в евклидовом пространстве прямые в неперпендикулярные.

2.2. (9) Какой интервал времени отсчитают часы, которые, двигаясь с постоянной скоростью, перемещаются из события O в событие B на рис. 2.1?

2.3. (20) Покажите, что для часов, измеряющих абсолютное время, временной интервал, который отсчитывается часами при переносе из одного события в другое, не зависит от пути.

3. Евклидова геометрия

Ковариантность представления по отношению к линейным преобразованиям — важное и тонкое понятие. Здесь мы сопоставим обычный подход к построению евклидовой геометрии с подходом, который обладает такой ковариантностью. Сразу заметим, что справедливы и та и другая точки зрения. Выбор между ними должен основываться не на предрассудках, а исключительно на соображениях удобства проведения конкретных вычислений. Поскольку измерение пространственных расстояний очень похоже на измерение временных интервалов, с аналогичной проблемой мы столкнемся, когда вернемся к специальной теории относительности. Здесь важны как отличия, так и сходные черты. Мы увидим, что введенное нами представление хроноструктуры события также описывает евклидову геометрию.

Евклидова геометрия в обычном изложении

Для начала обсудим традиционный подход к построению евклидовой геометрии. Не нарушая общности, можно считать наше пространство двумерным. Линейное пространство проще всего сделать евклидовым, если определить в нем скалярное произведение следующим образом.

[Причина, по которой для обозначения компонент используются верхние индексы, станет понятной в следующей главе, которая посвящена геометрии. Не путайте их с показателями степени.]

Евклидово скалярное произведение:

$$\mathbf{a} \cdot \mathbf{b} \equiv a^x b^x + a^y b^y, \quad (3.1)$$

где $\mathbf{a}$ является 2-вектором с компонентами a^x и a^y ; то же самое относится и к $\mathbf{b}$.

Для вычисления скалярного произведения векторов, которые не выходят из начала системы координат, будем пользоваться параллельным переносом (помните об однородности). Алгебраически это находит отражение в дистрибутивности операции скалярного произведения

[Здесь $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$ — векторы.]

$$(\mathbf{a} + \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c}. \quad (3.2)$$

Последнее свойство непосредственно следует из определения (3.1). Знакомые нам понятия длины и угла можно ввести, воспользовавшись скалярным произведением. Длина вектора $\mathbf{a}$, обозначаемая как $|\mathbf{a}|$, имеет вид

$$|\mathbf{a}|^2 = \mathbf{a} \cdot \mathbf{a}, \quad (3.3)$$

а угол θ между двумя векторами $\mathbf{a}$ и $\mathbf{b}$ дается соотношением

$$\cos \theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}| |\mathbf{b}|}. \quad (3.4)$$

В евклидовой геометрии скалярное произведение имеет вид (3.1) только при специальном выборе системы координат. Оси должны быть ортогональны, и масштаб на обеих осях должен быть одинаков. Такие координаты называются *ортогональными*. Они не единственны. Если повернуть систему координат, то правило (3.1) останется в силе. Простота этого правила обычно делает эффективным использование именно ортогональных координат. Аналогичное положение мы обнаружим в специальной теории относительности, где хроноструктура $\mathcal{S}$ также может быть представлена с помощью некоторого скалярного произведения. Вычисления, проводимые в рамках специальной теории относительности, выполняются наиболее эффективно в представлении, где скалярное произведение имеет простой вид. К сожалению, общую теорию относительности нельзя рассматривать в таких координатах. Действительно, присутствие гравитации проявляется в том, что мы не можем найти такое представление. К счастью, представления, которые ковариантны относительно произвольных линейных преобразований, не столь ограничены и могут быть нужным образом обобщены. Это более общее представление является и «более понятным», поскольку свойства индивидуального представления не путаются со свойствами самого множества $\mathcal{S}$.

Как сформулировать евклидову геометрию в форме, ковариантной относительно линейных преобразований? Поступим так же, как мы делали при рассмотрении часов. Возьмем множество точек $\mathcal{E}$, которые расположены на единичном расстоянии от начала координат. В случае ортогональных координат $\mathcal{E}$ представляет собой окружность. После выполнения произвольного числа растяжений, сдвигов и поворотов мы получим эллипс произвольного размера, эксцентриситета и ориентации. Правило измерения отрезков прямых линий называется метрикой; будем называть множества точек типа $\mathcal{E}$ *метрическими*

Ортогональные координаты

Произвольные координаты

Метрические фигуры

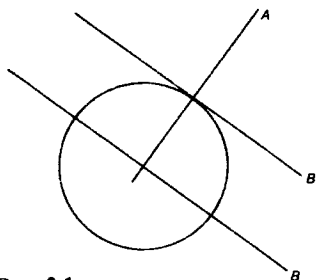


Рис. 3.1

Углы

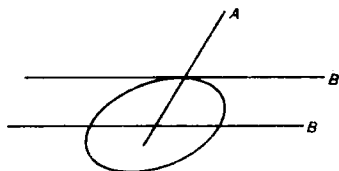


Рис. 3.2

Длины

фигурами. Мы видим, что в общем случае метрическая фигура для евклидовой геометрии представляет собой эллипс. Каким образом мы используем $\mathcal{S}$ для измерения длин и углов? Внимательно проследите, как это делается. Метод достаточно тонкий; кроме того, мы будем им много раз пользоваться. Мы знаем, что такое длина и угол в ортогональных координатах. Если сформулировать определение этих понятий в терминах операций, таких, как параллельный перенос, которые ведут себя должным образом при линейных преобразованиях, то те же операции можно выполнить в любой другой линейной системе координат; результат будет тот же.

Рассмотрим свойство перпендикулярности. На рис. 3.1, который соответствует выбору ортогональной системы координат, прямые A и B перпендикулярны, прямая B' параллельна B и касается единичной окружности в точке ее пересечения с прямой A . Если прямые параллельны или касаются кривых, это свойство сохраняется при линейных преобразованиях, поэтому мы можем использовать описанное выше построение для доказательства перпендикулярности прямых в любом линейном представлении. Типичный случай изображен на рис. 3.2. Если он кажется вам странным, постарайтесь мысленно растянуть рисунок таким образом, чтобы эллипс стал окружностью. Еще лучше посмотреть на чертеж под углом.

Чтобы найти длину отрезка прямой, будем действовать, как изображено на рис. 3.3. Найдем линию, которая параллельна заданному отрезку и проходит через начало координат. Затем с помощью параллельных линий перенесем отрезок единичной длины на измеряемый отрезок и таким образом найдем длину последнего (однородность). Опять все используемые при построении операции сохраняют свой вид при линейных преобразованиях, поэтому описанная процедура может быть осуществлена в общем случае. Скалярное произведение линейно, поэтому правило нахождения длины вектора дает возможность также вычислять скалярное произведение двух различных векторов:

$$2(a \cdot b) = (a + b) \cdot (a + b) - a \cdot a - b \cdot b. \quad (3.5)$$

Мы достигли намеченной цели, а именно представили евклидову геометрию, которая ковариантна только по отношению к поворотам, в форме, ковариантной по отношению к произвольным линейным преобразованиям. Мы также расширили понятие скалярного произведения. Равенство (3.1) спра-

ведливо только в ортогональных системах координат; в других случаях необходимо использовать приведенное выше обобщение, а именно определять длину векторов с помощью метрических фигур, а скалярное произведение векторов — с помощью равенства (3.5). Следует поупражняться с этим представлением. Оно очень похоже на то, которое мы будем использовать в специальной и общей теории относительности. Кроме того, такая практика должна помочь вам ближе познакомиться с понятиями ковариантности и линейного преобразования. Вам необходимо научиться узнавать величины, которые «ковариантны при линейных преобразованиях».

Пример

Рис. 3.4 соответствует произвольной линейной системе координат. Пусть параллелограмм A в действительности является представлением некоторого квадрата в евклидовом пространстве. Покажем графически, что тогда угол B прямой.

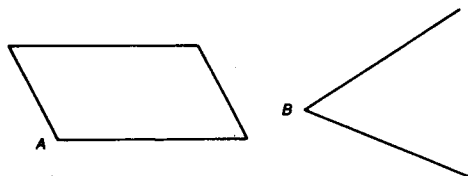


Рис. 3.4

Один способ решения заключается в нахождении соответствующей нашей системе координат метрической фигуры (в данном случае эллипса). Более простой подход состоит в таком изменении представления, при котором метрическая фигура стала бы окружностью. Мы осуществим это с помощью последовательности линейных преобразований, выполняемых одновременно над обеими фигурами. Важно, что одно и то же преобразование применяется к обеим фигурам. В этом смысл приставки «ко» в слове ковариантность. Цель соответствующих преобразований состоит в том, чтобы превратить параллелограмм в квадрат. В соответствующей системе координат метрическая фигура представляет собой окружность, и мы сможем увидеть непосредственно, действительно ли угол B прямой.

Прежде всего выполним преобразование сдвига так, чтобы параллелограмм A на рис. 3.4 стал прямоугольником, представленным на рис. 3.5. Затем сожмем рисунок вдоль оси x

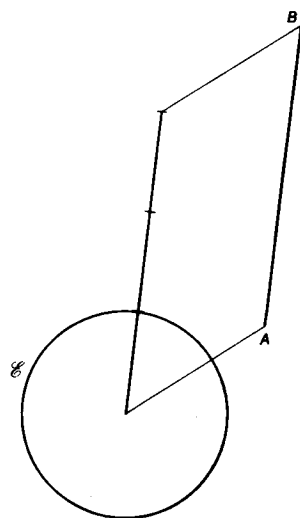


Рис. 3.3

Для измерения длины отрезка AB осуществляется его перенос и наложение на метрическую фигуру $\mathcal{C}$. Здесь длина AB примерно равна трем единицам длины.

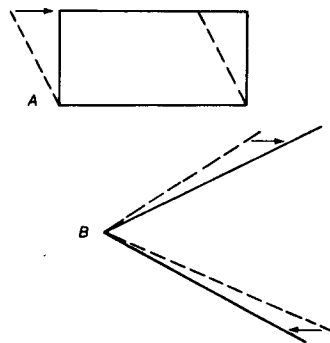


Рис. 3.5

Сдвиг вида $\begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$.

[Строго говоря, эти сдвиги включают в себя также растяжение и поворот.]

так, чтобы прямоугольник стал квадратом. Это также сдвиг. Результат перечисленных операций изображен на рис. 3.6. Угол B действительно прямой.

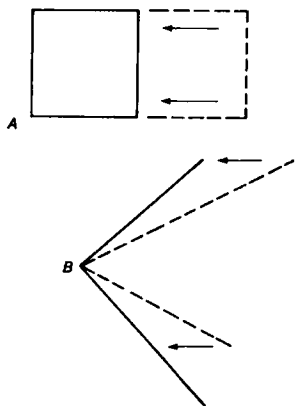


Рис. 3.6

Сдвиг вида $\begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} b & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$.

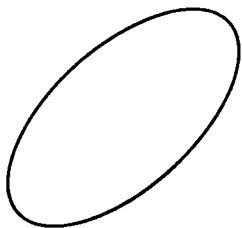


Рис. 3.7

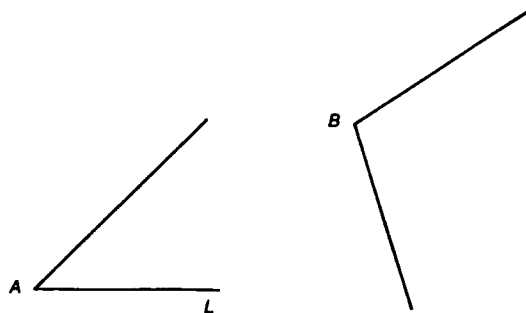


Рис. 3.8

ЗАДАЧИ

3.1. (17) В произвольной линейной системе координат, в которой единичная окружность имеет вид, изображенный на рис. 3.7, нарисуйте несколько квадратов различной ориентации.

3.2. (20) Задайтесь численными данными и повторите рассмотренную в примере задачу аналитически.

3.3. (22) Пусть оба угла, изображенные на рис. 3.8, прямые. Нарисуйте квадрат, одна сторона которого параллельна линии L . (Это построение выполняется не с помощью линейки и циркуля.)

3.4. (15) Выполните построение, изображенное на рис. 3.3, в произвольной линейной системе координат.

3.5. (25) Всегда ли существует система координат, в которой любые два угла типа рассмотренных в задаче 3.3 окажутся прямыми?

4. Инерциальные системы отсчета

Теперь мы готовы к тому, чтобы внимательно рассмотреть вопрос об инерциальной системе отсчета. Инерциальная система отсчета является отображением ψ событий физического мира на пространственно-временную диаграмму, отображением, которое обладает особыми свойствами по отношению к свободным частицам и часам. Строго говоря, мы можем рассматривать отображение просто как правило, позволяющее каждо-

му событию пространства-времени поставить в соответствие четыре числа некоторым систематическим способом. Примером такого сопоставления может служить изучение землетрясений. Пусть события, которые мы рассматриваем, представляют собой землетрясения. Чтобы определить положение такого события в пространстве и времени, нужно зарегистрировать время прихода данной сейсмической волны в четыре различные точки на поверхности Земли. Таким образом, каждому событию ставится в соответствие четыре числа, т. е. мы имеем пример одного из наших отображений ψ . Мы знаем, что пространство-время четырехмерно, так как для того чтобы различным событиям соответствовали различные наборы чисел, необходимо четыре числа и никогда не требуется более четырех чисел. Станции могут быть расположены где угодно, за исключением некоторых специальных случаев, например они не должны все лежать на одной и той же окружности большого радиуса.

Как обычно, хотелось бы преобразовать полученные координаты к другому, более удобному для дальнейшего использования виду. Предположим, что в нашем распоряжении имеется вычислительная машина, которая может выполнять над исходными координатами любое нужное нам преобразование. Выберем преобразование таким образом, чтобы по возможности упростить результирующую пространственно-временную картину. Здесь наш пример начинает выглядеть немного неестественным, поскольку не существует очевидного сейсмического аналога свободных частиц или часов с некоторой внутренней скоростью хода. Пример фантастической установки для получения пространственно-временных диаграмм изображен на рис. 4.1.

Прежде всего потребуем от нашего преобразования, чтобы мировые линии свободных частиц отображались в прямые линии. Если таких преобразований исходных координат не существует, то мы должны заключить, что имеются какие-то неизвестные силы, действующие на частицы, которые предполагались свободными. Если же это требование удовлетворено, остается все еще много преобразований, которые сохраняют прямолинейность мировых линий свободных частиц. До тех пор пока не наложены дополнительные требования, мы вправе воспользоваться любым из этих преобразований; в результате мы получим различные, но в равной степени удовлетворяющие нашим требованиям представления. Кроме линейных преобразований, которые были рассмотрены в разд. 3, существуют

Землетрясения

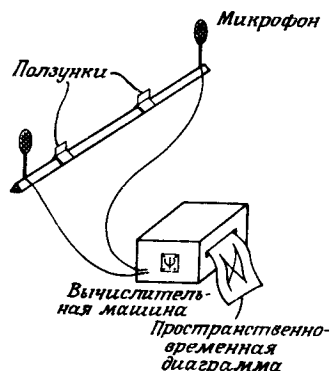


Рис. 4.1

Фантастическая схема установки для получения двумерных пространственно-временных диаграмм. В качестве событий можно рассматривать столкновение ползунков. Вычислительная машина вносит поправку на время распространения звука.

Инерциальные системы отсчета

другие преобразования, сохраняющие прямолинейность мировых линий, но не удовлетворяющие требованию равномерности хода часов. Это так называемые *проективные* преобразования, известные тем, кто занимается рисованием (выполнение рисунков в перспективе) и фотографией. Представим себе фотографию уходящего вдаль железнодорожного пути: на фотографии прямые линии остаются прямыми линиями, но тела одинакового размера, скажем железнодорожные шпалы, выглядят тем меньше, чем дальше они расположены. Примером проективного преобразования в случае двух измерений может служить преобразование

$$(x, t) \mapsto \left(\frac{x}{t-1}, \frac{t+1}{t-1} \right). \quad (4.1)$$

В качестве второго шага воспользуемся оставшейся свободой и с помощью проективного преобразования постараемся добиться того, чтобы временные интервалы удовлетворяли постулату о равномерности хода часов. В общем случае координаты, которые были выбраны, чтобы удовлетворить постулату свободных частиц, не будут удовлетворять постулату равномерности хода часов. Поэтому мы должны среди проективных преобразований отыскать преобразование, которое действительно делает временные интервалы удовлетворяющими этому постулату. После этого все еще остается свобода в проведении линейных преобразований. Вспомним наше определение.

Инерциальная система отсчета: определенное с точностью до линейных преобразований представление, изображаемое пространственно-временной диаграммой, на которой свободные частицы движутся по прямым линиям и отсчитываемые часами интервалы времени удовлетворяют постулату о равномерности хода часов.

Можно вполне успешно работать в представлении, которое определено только с точностью до линейных преобразований. Мы видели это на примере евклидовой геометрии. Однако можно накладывать и дополнительные ограничения на пред-

ставление, если нам нужна конкретная информация о поведении часов. Прделаем это для часов, измеряющих абсолютное время. Параллельно мы осветим ряд важных моментов, касающихся симметрии времени.

Расположим множество $\mathcal{A}$ горизонтально на расстоянии, равном единице, над началом координат. Используя линейные преобразования, это, конечно, можно сделать. Действительно, множество $\mathcal{A}$ должно представлять собой некоторую прямую линию, поэтому, осуществив простой поворот, а затем выполнив преобразование растяжения или сжатия, мы добьемся нужного результата. Инерциальные системы отсчета, выбор которых ограничен дополнительным требованием или соглашением, будем называть *каноническими* системами отсчета (слово «канон» означает «правило»). Описанная выше каноническая система отсчета все еще не определена однозначно; более того, однозначности нельзя добиться, накладывая дополнительные требования только на множество $\mathcal{A}$, представляющее хроноструктуру.

Оставшаяся свобода в выборе канонической системы отсчета для часов, измеряющих абсолютное время, определяется тем фактом, что множество $\mathcal{A}$, так же как и хроноструктура, которую оно представляет, обладает некоторой симметрией. Действительно, для любого фиксированного v преобразование

$$(x, t) \mapsto (x + vt, t) \quad (4.2)$$

переводит множество $\mathcal{A}$ в себя, т. е. заменяет одно каноническое представление другим. Преобразования такого вида называются преобразованиями Галилея.

Канонические системы отсчета

[Это определение не является общепринятым.]

Галилеева симметрия

[Хотя это преобразование включает в себя как чистый сдвиг, так и поворот, его называют сдвигом, и мы тоже будем называть его сдвигом.]

Симметрия: говорят, что объект обладает некоторой симметрией, если существует преобразование, которое оставляет его неизменным.

Если множество $\mathcal{A}$ не симметрично, а имеет выпуклости или перегибы, мы можем сформулировать правила, позволяющие расположить выпуклости и перегибы в определенных местах; это однозначно определяет представление. Симметрию абсолютного времени мы характеризуем словами, что время *инва-*

Инвариантность

риантно относительно преобразований Галилея. Мы собираемся использовать термины «инвариантность» и «ковариантность» в точном смысле этих слов, т. е. для обозначения разных понятий.

[Некоторые авторы не столь аккуратны и не различают эти два понятия.]

Инвариантность: симметрия некоторого объекта относительно преобразований.

Ковариантность: поведение представления некоторого объекта, не обязательно симметричного, при изменениях представлений пространства-времени.

Использование инвариантности

В качестве примера действия преобразований Галилея рассмотрим рис. 4.2. Отрезок мировой линии I с помощью некоторого преобразования Галилея можно совместить с отрезком I' . Следовательно, измеряющие абсолютное время часы, которые были перенесены вдоль отрезка I , и другие часы, которые были перенесены вдоль I' , должны отсчитать один и тот же интервал времени. Почему? Потому, что показания часов — это физическая величина, которая может быть вычислена в любом представлении. В частности, мы можем вычислить временной интервал, соответствующий отрезку I , в представлении, в котором I совпадает с I' . Правила нахождения показаний часов одинаковы в обоих представлениях. Следовательно, оба отрезка должны соответствовать одному и тому же интервалу времени.

Трансляции

Кроме симметрии, связанной с определенными линейными преобразованиями, существует также симметрия относительно трансляций. Это обусловлено тем, что если пространство-время однородно, то все события эквивалентны и в качестве начала отсчета можно выбрать любое событие. Такие неоднородные преобразования не включены в число линейных преобразований, и они не могут быть представлены с помощью матриц. Мы будем часто пользоваться трансляционной симметрией, чтобы упростить вычисления.

ЗАДАЧИ

4.1 (10) Какую некорректность мы допустили, предложив рас-

смотреть рис. 3.2 под углом, чтобы вместо эллипса увидеть окружность?

4.2. (12) Покажите, что показания часов, измеряющих абсолютное время, инвариантны относительно пространственных растяжений $(x, t) \rightarrow (kx, t)$,

где k — заданное число.

4.3. (18) Подробно рассмотрите проективные преобразования, которые приведены в примере на стр. 52.

4.4. (20) Пусть хроноструктура задается множеством $\mathcal{S}$, которое имеет вид (рассматривается случай двух измерений)

$$x + y = 1.$$

Найдите линейное преобразование, которое приводит $\mathcal{S}$ к каноническому виду для измеряющих абсолютное время часов.

4.5. (35) Объясните, почему серьезные ученые могут расходиться во мнении по поводу того, должна ли скорость света быть измерена экспериментально или постулирована.

4.6. (27) Прекрасный пример скрытой ковариантности дает интеграл

$$\oint \mathbf{r} \times d\mathbf{r}.$$

Покажите, что значение этого интеграла не зависит от положения начала координат. Считайте, что вы работаете в трехмерном евклидовом пространстве, где явный вид векторного произведения хорошо известен.

5. Физические часы

*Эллипс, бесспорно, хорош,
но — долой скромность!
Дайте мне гиперболы, чтобы
я мог насладиться истинной
красотой.*

Роджер Зелазни

Мы видели, каким образом в принципе можно задать инерциальную систему отсчета. На практике обычные временная и пространственные координаты, которые мы определяем с помощью часов и измерительных стержней, выделяют систему отсчета, которую с хорошей точностью можно считать инерциальной. В рамках этой системы отсчета мы можем изучать

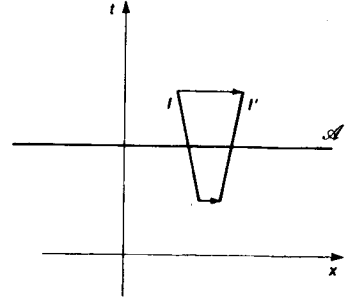


Рис. 4.2

Отрезок I преобразуется в отрезок I' с помощью преобразования Галилея. Здесь это преобразование имеет вид $(x, t) \mapsto (x + 2t/5, t)$.

[Для простоты мы продолжаем рассматривать случай только одного пространственного измерения.]

Эксперименты с часами

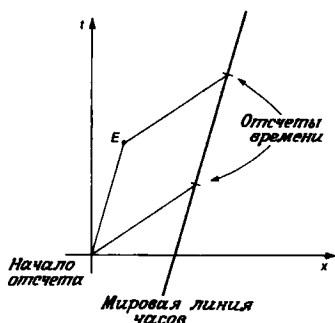


Рис. 5.1

Нахождение события E , принадлежащего множеству $\mathcal{S}$, с помощью рассмотрения соседних часов.

поведение реальных часов. Будут ли они описываться математической моделью, которую мы называли часами, измеряющими абсолютное время? На этот вопрос следует ответить отрицательно: только для малых скоростей модель часов, измеряющих абсолютное время, согласуется с поведением реальных часов.

Эксперимент заключается в том, что мы должны перемещать часы из события $(0,0)$ вдоль прямой мировой линии и отметить точку, которой часы достигли за единицу времени. Повторим такой эксперимент для часов, перемещающихся со всеми возможными скоростями. На пространственно-временной диаграмме таким часам соответствуют мировые линии, имеющие разный наклон. Все часы не могут одновременно проходить через одно и то же событие, поэтому чтобы перенести все отрезки мировых линий в начало координат, воспользуемся построением, изображенным на рис. 5.1. Найденное таким образом множество событий дает полное описание поведения физических часов.

Конечно, в реальных экспериментах мы можем получить только конечное число событий и перекрыть ограниченный диапазон изменения скоростей часов. Поэтому чтобы дать описание нашего множества некоторым математическим уравнением, следует произвести обобщение, воспользовавшись методом индукции. Оказывается, что если такой эксперимент действительно выполнен, то события, отстоящие от начала отсчета на единицу времени, ложатся (с точностью до ошибки измерения) на гиперболу

$$t^2 - \left(\frac{x}{c}\right)^2 = 1. \quad (5.1)$$

Результаты для реальных часов

Появляющаяся в этом выражении константа c зависит от выбора единиц измерения длины и времени. В СИ (Международная система единиц) она имеет численное значение, равное примерно $3 \cdot 10^8$ м/с. События, определяемые приведенным выше уравнением, образуют множество $\mathcal{S}$, которое описывает поведение физических часов. Оно изображено на рис. 5.2. В масштабе, в котором выполнен этот рисунок, множество $\mathcal{S}$ неотличимо от соответствующего множества, описывающего поведение часов, которые измеряют абсолютное время. Вот почему часы, измеряющие абсолютное время, служат хорошим приближением для описания поведения реальных часов. Мы действительно обнаружим отличие, только если будем рассматривать большие расстояния порядка 10^8 м для интервалов времени порядка секунд.

Чтобы увидеть, что происходит в действительности, необходимо сжать наш рисунок в направлении оси x ; тогда большие расстояния также будут отражены на диаграмме. Для это-

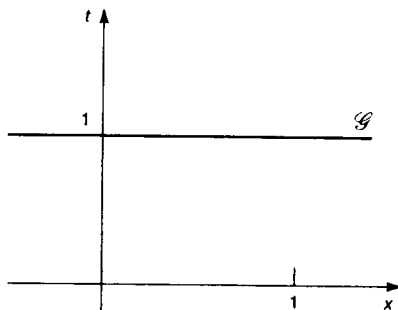


Рис. 5.2

Изображение результата эксперимента, выполненного с реальными часами. Время измеряется в секундах, расстояние — в метрах.

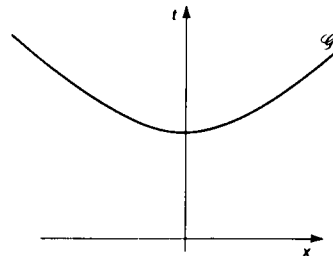


Рис. 5.3

Вид кривой $\mathcal{G}$, изображенной на рис. 5.2, в случае, когда в качестве единицы длины используется световая секунда.

го введем новую единицу длины; по причинам, которые станут понятными позже, назовем ее световой секундой. Она представляет собой расстояние, равное $2,998 \cdot 10^8$ м. Почти все расстояния в нашей книге будут измеряться в таких единицах. Кривая $\mathcal{G}$ в новых переменных изображена на рис. 5.3. Теперь легко увидеть отличие между реальными часами и часами, измеряющими абсолютное время. В качестве упражнения докажете, что для гиперболы, изображенной на рис. 5.3, часы, которые перемещаются с постоянной скоростью из события (x_0, t_0) в событие (x_1, t_1) , отсчитают интервал времени τ , определяемый выражением

$$\tau^2 = (t_0 - t_1)^2 - (x_0 - x_1)^2. \quad (5.2)$$

Что можно сказать о часах, которые перемещаются вдоль искривленных мировых линий? Многие часы, и в особенности часы, отличающиеся высокой точностью, не выдерживают сколько-нибудь значительных ускорений. Некоторые типы часов, например нестабильные частицы, по-видимому совершенно нечувствительны к ускорениям. Поэтому чувствительные к ускорениям часы следует использовать в эксперименте только для измерения прямых отрезков мировых линий. Если же часы не реагируют на ускорение, мы можем вычислять ин-

Световые секунды

[Здесь x_0 и x_1 измеряются в световых секундах.]

Ускорение

тервалы времени вдоль искривленных путей, представляя кривую в виде последовательности большого числа прямолинейных отрезков и переходя к пределу, в точности так же, как вычисляется длина кривой в курсе интегрального исчисления.

История

Излагая развитие теории, мы не придерживались здесь исторической последовательности. На самом деле уравнение (5.1) было угадано при рассмотрении симметрии теории электромагнетизма Максвелла и только позже было проверено экспериментально. Мы используем очевидный и непосредственный подход к проблеме; чтобы понять его, не нужно обладать какими-либо выдающимися способностями или ждать, чтобы вас озарило вдохновение. Самые первые эксперименты заключались в измерении времени жизни нестабильных частиц, μ -мезонов, которые рождаются в верхних слоях атмосферы при прохождении через нее космических лучей. Было известно, что эти частицы перемещаются со скоростью, очень близкой к скорости света, а их среднее время жизни было известно из других экспериментов. Соответствующая такому эксперименту

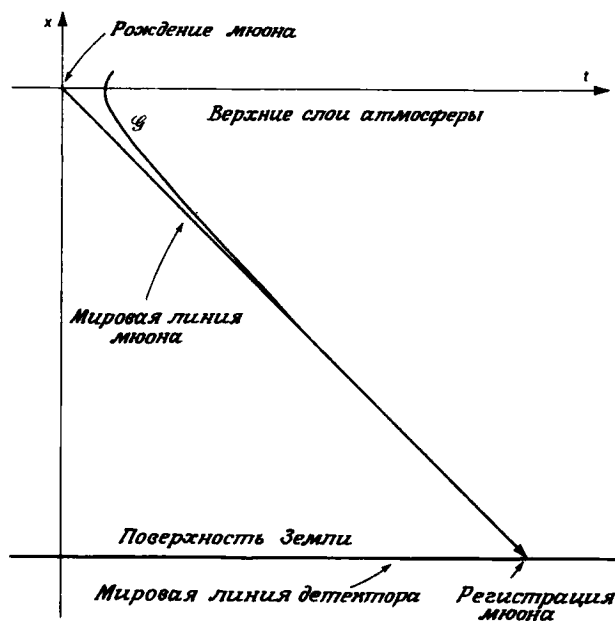


Рис. 5.4

Наблюдение распада мюона. Кривая g задается уравнением $t^2 - x^2 = \tau^2$, где расстояние x измеряется в световых секундах, τ — время жизни мюона, равное $2 \cdot 10^{-6}$ с (масштаб не соблюден).

пространственно-временная диаграмма изображена на рис. 5.4. Мы повернули рисунок таким образом, чтобы его верхняя часть соответствовала верхним слоям атмосферы. Наивные люди (знающие, что такое μ -мезон, но не имеющие представления о специальной теории относительности) могут ожидать, что мезон, который движется со скоростью, близкой к скорости света и имеет время жизни всего $2 \cdot 10^{-6}$ с, должен пройти до распада расстояние, равное всего $2 \cdot 10^{-6}$ световым секундам. Однако фактически мезон проходит сквозь всю атмосферу и покрывает расстояние порядка $40 \cdot 10^{-6}$ с. (Обратите внимание, что эксперимент занимает очень короткое время. Для таких малых времен влиянием гравитации можно пренебречь и связанную с Землей систему отсчета рассматривать как инерциальную. Мы обсудим этот вопрос подробнее, когда перейдем к описанию гравитационного поля.)

Конечно же, потребовались дополнительные эксперименты. Двух точек не достаточно для определения всей кривой. Многочисленные подтверждения справедливости соотношения (5.1) дают эксперименты на ускорителях частиц высоких энергий.

ЗАДАЧИ

5.1. (10) Если рис. 5.4 выполнен в соответствующем масштабе, то на какой высоте над поверхностью Земли родился мезон?

5.2. (16) Часы перемещаются с постоянной скоростью из точки (0, 0) в точку (1, 2) и затем в точку (0, 4). Опишите на словах, что произойдет. Какие показания часов будут соответствовать точке (0, 4), если их показания в точке (0, 0) были приняты за нуль? Расстояния измеряются в световых секундах, время — в секундах. Что показали бы часы, которые остались в точке с координатой $x = 0$?

5.3. (28) Рассмотрите гипотетические часы, которые удовлетворяют постулатам универсальности и равномерности хода и для которых множество $\mathcal{S}$ задается соотношением

$$t^4 = x^4 + Ax^2 + 1,$$

где A — некоторая заданная константа. Покажите, что интервал времени τ , прошедший по часам, которые переместились с постоянной скоростью из начала отсчета в точку с координатами (x, t) , определяется выражением

$$\tau^2 = \frac{2t^4(1 - v^4)}{Ax^2 + [4t^4 + (A^2 - 4)x^4]^{1/2}}.$$

5.4. (16) В наших единицах ускорение измеряется в с^{-1} . Покажите, что ускорение свободного падения на Земле примерно равно $(1 \text{ год})^{-1}$. Случайно ли это?

6. Световые сигналы и канонические системы отсчета

Свет играет важную роль как в теории относительности, так и в космологии. Конфликт между механикой и электродинамикой породил специальную теорию относительности. Информацию о Вселенной приносит главным образом электромагнитное излучение, которое воспринимается нашими телескопами. Абстрагируясь от конкретных деталей, приходим к понятию светового сигнала, под которым понимается «короткий импульс света, распространяющийся в определенном направлении»; такой импульс можно создать при помощи лазера и быстродействующего затвора. Не принимая во внимание пространственные размеры светового импульса, будем рассматривать световой сигнал как особую частицу. Мы отказались от мысли называть ее фотоном, поскольку последний термин закреплен за вполне определенным квантовомеханическим понятием. В нашей модели световые сигналы будут представляться мировыми линиями. Фундаментальные закономерности, управляющие поведением световых сигналов, могут быть сформулированы в виде следующих двух постулатов.

Постулат инерциальности для светового сигнала: мировые линии светового сигнала представляют собой прямые линии в любой инерциальной системе отсчета.

Постулат неизменности скорости света: в инерциальной системе отсчета мировые линии световых сигналов, распространяющихся в одном и том же пространственном направлении, параллельны.

Основная идея, которая отражена в этих постулатах, заключается в том, что распространение света не зависит от того, каким источником он испускается и как источник движется. Здесь слова «пространственное направление» используются в обыч-

ном смысле. Вдоль прямой существуют два пространственных направления: вправо и влево. На плоскости имеется однопараметрическое семейство направлений: север, юг и т. д. В случае наших $1 + 1$ -мерных пространственно-временных диаграмм мировые линии световых сигналов образуют два семейства параллельных прямых. Договоримся в дальнейшем изображать мировые линии световых сигналов штриховыми линиями. Мы помним, что одних фундаментальных постулатов, касающихся поведения часов, недостаточно для построения теории; к ним необходимо добавить точный вид хроноструктуры $\mathcal{S}$, которая может быть получена из эксперимента. Точно так же наши постулаты о световых сигналах описывают только их наиболее фундаментальные свойства; необходим эксперимент, чтобы сказать, как в действительности проходят их мировые линии.

На рис. 6.1 изображена возможная ситуация в $1 + 1$ -мерном случае, а на рис. 6.2 — в случае $2 + 1$ измерений. На

[Чтобы избежать недоразумений, мы называем пространственно-временную диаграмму в случае одного пространственного и одного временного измерения $1 + 1$ -мерной диаграммой.]

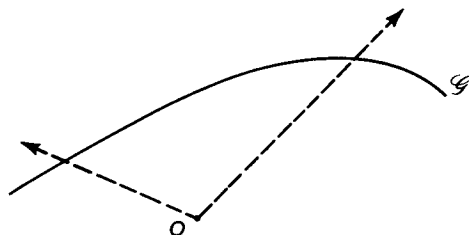


Рис. 6.1

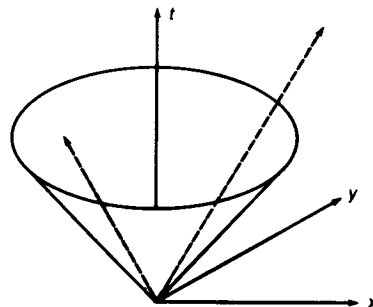


Рис. 6.2.

Световые сигналы в случае двух пространственных и одного временного измерений.

рис. 6.2 мы изобразили все возможные световые сигналы, выходящие из данного события E . Они образуют коническую поверхность, которую называют *световым конусом*. Каким образом можно мысленно представить себе ситуацию в $3 + 1$ -мерном случае? Один из используемых здесь приемов заключается в рассмотрении сечений диаграммы. Это почти то же самое, что делает архитектор, когда вычерчивает план каждого этажа здания по отдельности. Для упражнения рассмотрим

2 + 1- мерную диаграмму. На рис. 6.3 изображены два ее сечения, выполненные в различные моменты времени. Они напоминают два кадра в кино. На каждом кадре изображено кольцо из световых сигналов, которое с течением времени расширяется. Как будут выглядеть следующие один за другим кадры, соответствующие разрезам 3 + 1-мерной диаграммы? На них, очевидно, будут изображены сферы, радиус которых увеличивается с течением времени.

Проверка постулатов

Сформулированные в этом разделе постулаты постоянно перепроверяются. Например, в одном обсуждаемом в настоящее время эксперименте будет проводиться сравнение давно испущенного («старого») света, который приходит к нам от отдаленных галактик, с «молодым» светом, испущенным в данном месте и в данный момент. Чтобы эксперимент был чистым, надо проводить сравнение вне атмосферы, так как прохождение через атмосферу может привести к постепенному превра-

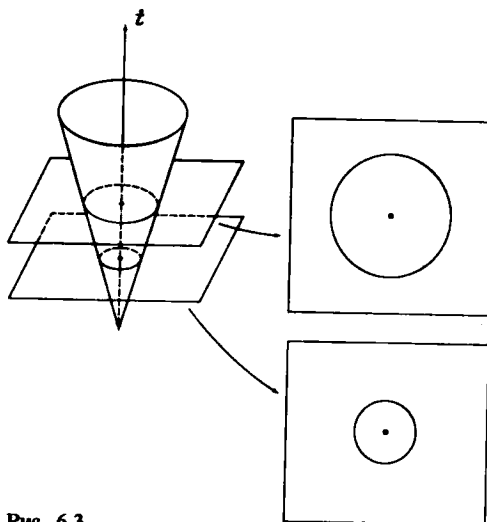


Рис. 6.3

щению «старого» света в «молодой». Сейчас никто не ожидает обнаружить какое-либо отличие между «старым» и «молодым» светом, но никто и не думает, что наша современная точка зрения сохранится и в дальнейшем.

Результирующая пространственно-временная диаграмма

Наши представления о структуре пространства-времени могут быть отображены на одной пространственно-временной диаграмме, на которой указано начало отсчета, изображено множество $\mathcal{S}$ и направления, по которым могут распростра-

няться световые сигналы. Пример такой диаграммы дает рис. 6.1. Другой пример изображен на рис. 6.4; это хорошо знакомые нам часы, измеряющие абсолютное время, в произвольной инерциальной системе отсчета.

Дополнительная информация о пространстве-времени, которую мы получаем при исследовании распространения световых сигналов, позволяет нам продвинуться дальше в рассмотрении канонических инерциальных систем отсчета. Наиболее прост случай часов, измеряющих абсолютное время. Это очень важный и поучительный пример, даже если такие часы и не соответствуют реальной действительности. Один из способов специализированного определения инерциальной системы отсчета состоит в том, чтобы известным образом ориентировать некоторую выделенную мировую линию. Часто мировую линию наблюдателя, с которым связана система отсчета, на пространственно-временной диаграмме располагают вертикально. Мы будем говорить о такой связанной с наблюдателем системе отсчета как о системе отсчета, в которой наблюдатель «покоится». Тогда существуют два пути дальнейшей специализации системы отсчета. Можно либо упростить описание часов, либо упростить описание световых сигналов. Как правило, нельзя сделать одновременно и то и другое.

Специализированная по времени система отсчета — это такая система отсчета, в которой множество $\mathcal{A}$ представляется горизонтальной прямой, перпендикулярной мировой линии наблюдателя. Системой отсчета, специализированной по световым сигналам, мы будем называть такую систему отсчета, где световые сигналы на пространственно-временной диаграмме изображаются прямыми линиями, тангенс угла наклона которых равен $+1$ или -1 . На рис. 6.5 показано, каким образом ситуацию, изображенную на рис. 6.4, можно преобразовать к виду, соответствующему системе отсчета, специализированной по времени. Буквой W обозначена мировая линия выделенного наблюдателя. Как показано на рисунке, сначала мы выполняем поворот, затем — преобразование Галилея. Рис. 6.6 соответствует переходу к системе отсчета, специализированной по световым сигналам.

Какая из этих систем отсчета правильна? Надеюсь, читателю ясно, что не имеет смысла так ставить вопрос. Обе системы отсчета являются правильными и эквивалентными представлениями *одной и той же* физической ситуации. Можно использовать ту или другую систему в зависимости от того, что вы собираетесь делать дальше, но любая задача может ре-

[Канонические системы отсчета обсуждались ранее в разд. 4.]

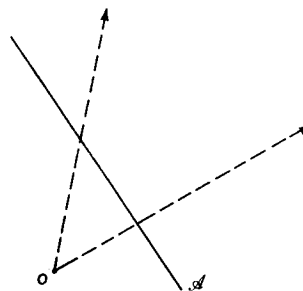


Рис. 6.4

Системы отсчета, специализированные по времени, и системы отсчета, специализированные по световому сигналу

[О преобразованиях Галилея см. стр. 53.]

Можно использовать любую систему отсчета

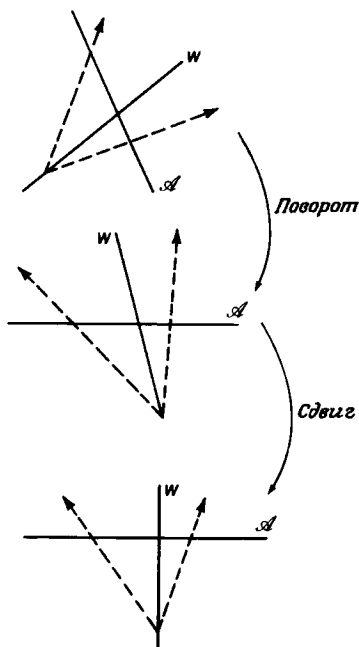


Рис. 6.5

Построение системы отсчета, специализированной по времени, для ситуации, изображенной на рис. 6.4. Сдвиг представляет собой преобразование Галилея.

шаться в любой системе отсчета, включая исходную, не специализированную. Обычно ошибки возникают из-за того, что представлению придается слишком большое значение. Мы даже сами часто не замечаем, как используемый нами язык заставляет нас работать во вполне определенной системе отсчета. Наблюдатель, который рассматривает происходящее с точки зрения системы, специализированной по времени, мог бы

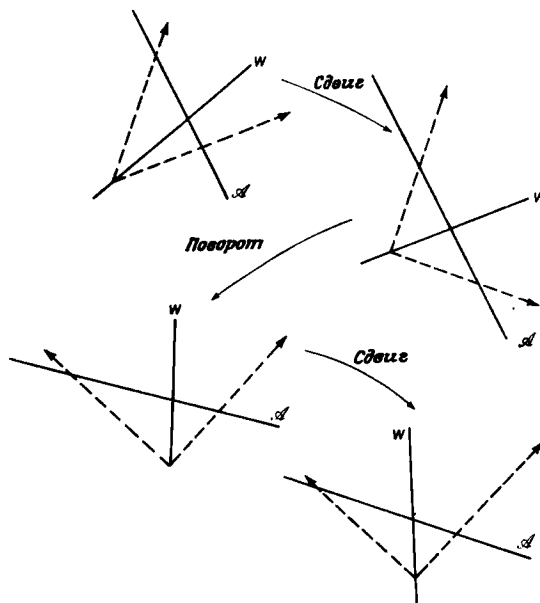


Рис. 6.6

Построение системы отсчета, специализированной по световому сигналу. Первый сдвиг осуществляется вдоль направления распространения левого светового сигнала. Второй сдвиг является преобразованием сжатия вдоль направления распространения правого светового сигнала.

сказать, что скорость света различна для различных направлений распространения света. Другой наблюдатель, работающий в системе, специализированной по световому сигналу, будет считать первого безумцем, поскольку для него скорость света постоянна по определению. Внимательно подумайте, насколько оба этих наблюдателя правы и не правы в одно и то же время. Тот, кто связан путами ньютоновой философии, считает время более фундаментальным понятием, чем свет, и выбор любой системы отсчета, кроме системы, специализированной

по времени, кажется ему странным. *Революционная идея, которую внесла специальная теория относительности, состоит в том, что свет играет более важную роль, чем время.*

В специальной теории относительности очень красиво разрешается конфликт между часами и световыми сигналами. Там его просто не существует. В одной и той же системе отсчета выглядят проще и поведение часов, и распространение световых сигналов. Оказывается, что мировые линии световых сигналов являются асимптотами для гиперболы $\mathcal{G}$, которая изображает поведение часов (рис. 6.7). Мы не смогли бы должным образом оценить это, если бы прежде не рассмотрели часы, измеряющие абсолютное время.

Спецрелятивистские часы

Связанная с мировой линией W каноническая система отсчета — это инерциальная система отсчета, где световые сигналы распространяются вдоль прямых, тангенс угла наклона которых равен единице, а мировая линия W расположена вертикально.

[Не существует общепринятого определения, аналогичного данному. Это определение принадлежит автору.]

В гипотетическом случае часов, измеряющих абсолютное время, можно специализировать выбор системы отсчета. Существует только одна каноническая система отсчета, в которой изображающая хроноструктуру линия $\mathcal{A}$ перпендикулярна выделенной мировой линии. Мы будем называть ее *привилегированной системой отсчета*. В такой системе отсчета физические закономерности выглядят особенно просто.

Нашу теорию, в которой частицы и световые сигналы представляют собой первичные элементы, не следует считать наиболее фундаментальной из существующих в настоящее время теорий. Свет лучше описывается электродинамикой Максвелла, а частицы — квантовой механикой. Но было бы неправильно начинать с изучения этих теорий, так как их язык покоится на понятиях инерциальной системы и лоренцевой симметрии хода часов и световых сигналов. В гл. III мы подойдем ближе к этим более последовательным теориям, выбрав волновой пакет в качестве первичного понятия. Соответствующая теория называется полуклассической.

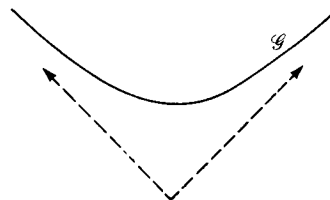


Рис. 6.7

Кривая, изображающая поведение реальных часов в специальной теории относительности. Оси координат не изображены, чтобы подчеркнуть независимость полученного представления от выбора системы координат.

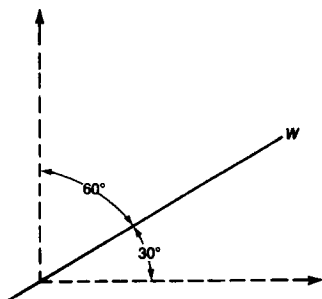


Рис. 6.8

ЗАДАЧИ

6.1. (25) Найдите матрицы, которые соответствуют преобразованиям, изображенным на рис. 6.5. Углы равны 30° , 45° и 60° .

6.2. (25) Прodelайте то же самое для преобразований, изображенных на рис. 6.6.

6.3. (20) Графически и аналитически найдите преобразование к канонической системе отсчета для световых сигналов и мировой линии W , изображенных на рис. 6.8.

6.4. (30) Найдите каноническую систему отсчета, специализированную по световому сигналу, для гипотетических часов, для которых множество $\mathcal{S}$ имеет вид

$$t^4 - x^4 = 1,$$

если мировые линии световых сигналов описываются уравнением

$$t^2 = x^2.$$

Мировая линия W составляет угол 30° с вертикалью.

6.5. (35) В случае гипотетической ситуации, когда часы измеряют абсолютное время, найдите способ измерения скорости относительно привилегированной системы отсчета, описанной на стр. 65, используя только часы и световые сигналы.

7. Специальная теория относительности

Мне не хотелось бы переубеждать Вас, сударь, но должен заметить, что это неверно.

Г. Лихтенберг

Мы завершили обсуждение основных представлений специальной теории относительности. Частности, речь о которых пойдет в оставшихся разделах настоящей главы, вытекают из этих представлений. Причем вплоть до последнего раздела, где будут рассмотрены энергия и импульс, нам не придется сталкиваться с какими-либо новыми физическими идеями. Но прежде чем двигаться дальше, вернемся еще раз к основным понятиям.

В моделях, рассматриваемых в специальной теории относительности, физическая реальность описывается с помощью таких первичных понятий, как события, свободные частицы, часы и световые сигналы. Согласно нашим основным постулатам, существует такое представление физических процессов с

[Обратите внимание на этот перечень первичных понятий: не время, а часы являются одним из основных элементов специальной теории относительности.]

помощью математических объектов в векторном пространстве, при котором мировые линии свободных частиц и световых сигналов изображаются прямыми линиями, ход часов равномерен, а световые сигналы распространяются с одинаковой скоростью, не зависящей от скорости их источника. Такое представление мы ассоциируем с инерциальной системой отсчета. Можно выбрать и более специализированное представление, если положить скорость света равной единице и выделить некоторую мировую линию в качестве линии неподвижного наблюдателя. Это специализированное представление мы называем канонической системой отсчета для выделенной мировой линии. Существует каноническая система отсчета, в которой поведение часов, согласно наблюдениям, описывается множеством $\mathcal{S}$ событий, отстоящих от начала координат на единичный интервал времени. Координаты этих событий удовлетворяют уравнению

$$t^2 - x^2 = 1, \quad (7.1)$$

а в 3 + 1-мерном случае — уравнению

$$t^2 - x^2 - y^2 - z^2 = 1. \quad (7.2)$$

Все это схематически изображено на рис. 7.1.

В следующих разделах на основе изложенных представлений мы обсудим одновременность, синхронизацию часов, замедление хода времени, относительную скорость, лоренцеву симметрию и 4-векторы. Мы покажем, что, если хроноструктура $\mathcal{S}$ изображается гиперболой в какой-нибудь одной канонической системе отсчета, то она будет изображаться такой же гиперболой во всех канонических системах отсчета. В большинстве случаев при изложении СТО такую эквивалентность рассматривают как основное исходное предположение. Мы же, напротив, избрали подход, в котором эквивалентность вытекает из основных представлений. Подобный подход необходим для обсуждения экспериментов, подтверждающих эквивалентность. На деле такое обсуждение было бы довольно формальным, и мы не будем на нем останавливаться. Однако следует отдавать себе отчет, что для более полного понимания СТО необходимо хотя бы в принципе представлять ход этого обсуждения.

[Как инерциальные, так и канонические системы отсчета являются одинаково корректными представлениями. Единственная причина перехода к канонической системе отсчета — ее удобство.]

ЗАДАЧА

7.1. (30) В качестве одного из законов общей теории систем Джеральд Вейнберг [41] предложил так называемый принцип

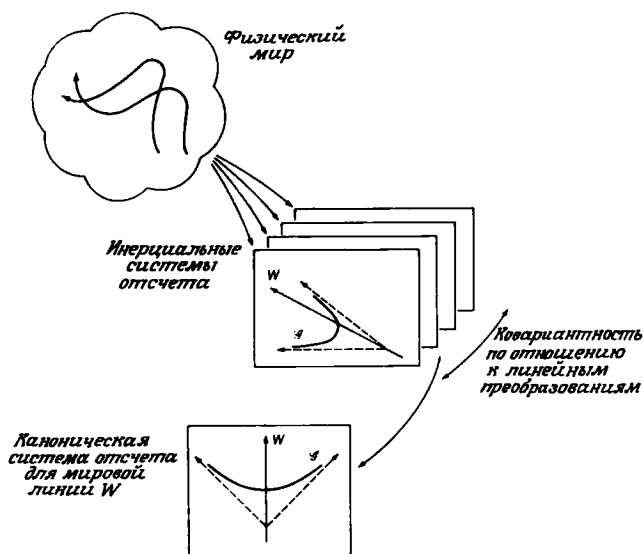


Рис. 7.1

«Считай до трех»: «Если вы не можете придумать три способа использования инструмента не по прямому назначению, то вы не понимаете, как им пользоваться.» Примените его к СТО.

8. Одновременность

Поскольку основное внимание в СТО уделяется не времени, а часам, концепция одновременности играет здесь менее важную роль, нежели в ньютоновой физике. Такая парадоксальная ситуация привела к тому, что в большинстве учебников теории относительности проводится пространное обсуждение одновременности. Я думал было сделать радикальный шаг и вообще не рассматривать вопрос об одновременности, тем самым подчеркнув, что без него вполне можно обойтись, но его обсуждение дает массу полезных идей и дополнительных примеров, которые требуют так импонирующего мне геометрического способа рассуждений.

Мы не будем рассматривать одновременность как еще одно первичное понятие, а, напротив, будем считать, что эта концепция определяется в рамках уже разработанной теории. Не существует физического процесса, который позволил бы определить одновременность неким «внутренним» образом — нет

Одновременность: понятие, подлежащее определению

естественной меры одновременности, — если не считать, что это понятие можно определить с помощью часов и световых сигналов. Для нас одновременность — это всего лишь понятие, подлежащее определению. И вот что примечательно: наиболее разумное, с нашей точки зрения, определение одновременности в СТО оказывается непригодным. Такая несостоятельность интуиции проявляется, в частности, в парадоксе, который принято называть парадоксом близнецов.

Наше интуитивное представление о времени связано с часами, измеряющими абсолютное время. Это совершенно естественно, ибо, как мы видели, такие часы при малых скоростях ведут себя почти как «настоящие». В связи со сказанным возникает вопрос: как можно было бы использовать часы, измеряющие абсолютное время, для определения одновременности? На этот вопрос можно ответить следующим образом: измеренный такими часами интервал времени между двумя событиями не зависит от пути, по которому перемещались часы между этими событиями. Следовательно, как мы видим, можно дать определение одновременности, основанное на процедуре переноса часов в пространстве-времени, как показано на рис. 8.1.

Хронотранспортная одновременность: если для двух событий A и B , а также для происшедшего ранее события E часы, перенесенные из E в A , показывают тот же интервал времени, как и часы, перенесенные из E в B , то события A и B называются хронотранспортно одновременными.

Ход реальных часов, которые описываются спецрелятивистской гиперболой, существенно зависит от пути, вдоль которого они переносятся в пространстве-времени; следовательно, мы вынуждены отказаться от определения одновременности, основанного на процедуре переноса часов. На что же нам опереться? Пытаясь дать определение физической величины, следует обратиться к ее предполагаемым применениям. К сожалению, одновременность в основном нужна для задания системы отсчета, но, оказывается, для этого можно найти и другие способы. Таким образом, применения не дают нам ключа к разрешению поставленного вопроса.

Определение одновременности, основанное на понятии абсолютного времени

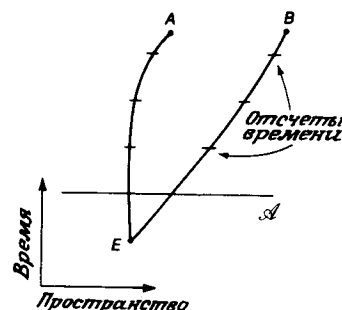


Рис. 8.1
Построение двух одновременных событий A и B с помощью переноса часов, измеряющих абсолютное время, из какого-то другого события E .

Определение одновременности в СТО

[Зависимость показаний часов от пути переноса обсуждается в задаче 5.2.]

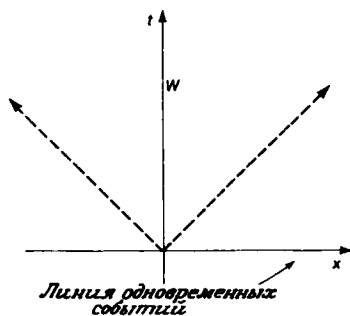


Рис. 8.2

Линия одновременных событий в канонической системе отсчета, связанной с мировой линией W .

При обсуждении лоренц-инвариантности будет доказано, что невозможно определить одновременность, опираясь лишь на первичные понятия, входящие в нашу теорию. Следовательно, здесь требуется кое-что еще. Чтобы сдвинуться с мертвой точки, можно, например, определить одновременность по отношению к особой канонической системе отсчета. В этой особой системе отсчета события, расположенные вдоль горизонтальной линии, считаются по определению одновременными, как показано на рис. 8.2. Собственно, такое определение — это лучшее из того, что можно было сделать. Все можно несколько упростить, если дать операциональное определение, эквивалентное приведенному выше. Каноническая система отсчета зависит от выбора мировой линии наблюдателя, которого следует считать покоящимся. Можно дать операциональное определение одновременности, основанное на выделенной мировой линии и на таких понятиях, как часы и световые сигналы.

Прибегнем снова к стандартному способу рассуждений. Пусть имеется определение, справедливое в канонической системе отсчета. Если его можно сформулировать с помощью понятий, не относящихся к какой-нибудь конкретной системе отсчета, то этим определением можно пользоваться в любой системе. Такое определение называется *ковариантным*. Это означает, что оно обладает такой же степенью независимости от представления, как и вся остальная теория. Ход рассуждений здесь такой же, как и при обсуждении определения перпендикулярности в евклидовой геометрии, ковариантного по отношению к линейным преобразованиям. Чтобы привести определение, идея которого отражена на рис. 8.2, к ковариантной форме, нужно лишь сказать, что интересующие нас события A и B лежат на линии, перпендикулярной выделенной мировой линии W . Однако это высказывание должно опираться только на такие понятия, как прямые линии, световые сигналы и показания часов. Сделать это несложно. Взгляните на рис. 8.3, где изображены все интересующие нас световые сигналы. Ясно, что перпендикулярность гарантирована только в том случае, если равны интервалы времени τ_1 и τ_2 , отмеренные вдоль мировой линии W . Теперь у нас есть все необходимое для формулировки искомого определения.

Световая одновременность: Два события A и B называются одновременными по отношению к мировой линии W , если линии световых сигналов, проведенные из A и B в прошлое и в будущее, отсекают на W одинаковые интервалы времени.

Как показывает рис. 8.4, такое определение применимо в любой системе отсчета. Совершенно очевидно, что наше определение зависит от выбора мировой линии W , ибо, как следует из рис. 8.5, соотношение между интервалами на мировой ли-

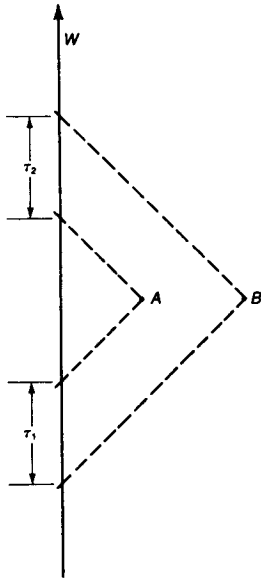


Рис. 8.3
Ковариантные операции,
определяющие световую од-
новременность.

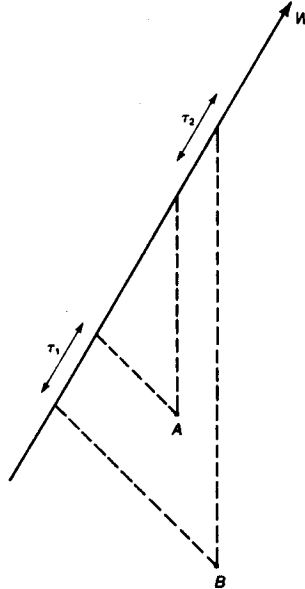


Рис. 8.4
Определение световой одно-
временности в произвольной
инерциальной системе отсче-
та.

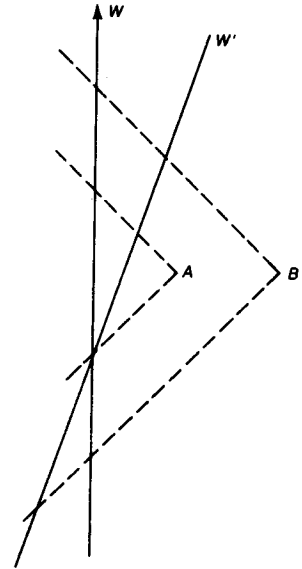


Рис. 8.5

нии движущегося наблюдателя W' не согласуется с соотношением между соответствующими интервалами на мировой линии наблюдателя W , который предполагается неподвижным по отношению к данной системе отсчета.

Заметьте, что здесь промежутки времени сравнивались только вдоль каждой отдельной мировой линии. Это означает, что наше определение одновременности опирается лишь на равномерность хода часов и не зависит от выбора какой-то конкретной хроноструктуры $\mathcal{S}$. Именно по этой причине мы назвали ее световой, а не спецрелятивистской одновременностью. В принципе такое определение вполне пригодно и для ча-

Одновременность зависит только
от световых сигналов

сов, измеряющих абсолютное время. Мы не стали им пользоваться по той простой причине, что в тех случаях, когда можно оперировать понятием абсолютного времени, зависимость от мировой линии W вносит излишние усложнения, которых вполне можно избежать. И только несостоятельность понятия хронотранспортной одновременности в общем случае вынуждает нас принять это качественно более слабое определение световой одновременности. В дальнейшем любое упоминание об одновременности будет относиться только к световой одновременности.

Графики

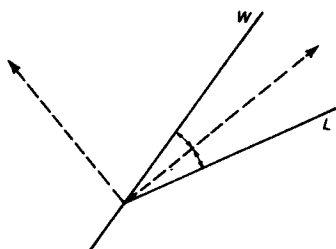


Рис. 8.6

Если показанные углы равны в канонической системе отсчета, то все события, лежащие на прямой L , удовлетворяют определению световой одновременности по отношению к мировой линии W .

Существует простая связь между взаимным расположением мировой линии W и прямой L , являющейся геометрическим местом событий, которые по отношению к W удовлетворяют определению световой одновременности: в любой канонической системе отсчета мировая линия W и прямая L составляют одинаковые евклидовы углы с мировой линией светового сигнала, как показано на рис. 8.6. Вернемся немного назад и оценим то, что мы только что делали. Модель геометрии пространства-времени строится на основе евклидовой геометрии. Некоторые концепции евклидовой геометрии имеют смысл и в пространстве-времени, но не все. В частности, в СТО имеет смысл говорить только о равенстве углов, составляемых прямыми с мировыми линиями световых сигналов и лишь по отношению к канонической системе отсчета. Корректность выполненного на рис. 8.6 построения будет доказана в разделе, посвященном преобразованиям Лоренца. Описанный там подход позволит выполнить это доказательство в очень компактной форме.

Синхронизация часов

Сформулировав определение одновременности, мы можем пойти дальше и применить его при обсуждении процедуры синхронизации часов. Поскольку наше определение зависит от наклона некоторой конкретной мировой линии, синхронизация часов A с часами B даст результат, отличный от исхода синхронизации B с A , если только эта пара часов не описывается параллельными мировыми линиями. Таким образом, корректная процедура синхронизации имеет место только для семейства часов, движущихся вдоль параллельных мировых линий. Рассматриваемая процедура, опирающаяся на определение световой одновременности, представлена на рис. 8.7. В момент времени t_1 по часам A от них посылается световой сигнал к часам B . Затем этот сигнал сразу же отправляется обратно к часам A , которые в момент его получения показывают время t_2 . Синхронизировать часы B с часами A — значит настроить часы B

так, чтобы в момент отражения сигнала они показывали время

$$t_0 = \frac{1}{2}(t_1 + t_2). \quad (8.1)$$

К этой процедуре приводят те же рассуждения, какие использовались нами ранее. В канонической системе отсчета, по отношению к которой все рассматриваемые часы покоятся, координатное время должно согласовываться с временем, отмеряемым часами. Прделанные выше операции гарантируют выполнение этого требования, к тому же они могут быть описаны в *любой* системе отсчета. Таким образом, если дана диаграмма пространства-времени с несколькими мировыми линиями движущихся часов и поставлена задача синхронизировать часы, то вовсе нет необходимости в переходе к системе покоя часов, их синхронизации в этой системе отсчета и в последующем преобразовании к исходной системе отсчета. Вместо всего этого достаточно выполнить указанные выше ковариантные операции. Теперь вы, наверное, можете по достоинству оценить всю полезность ковариантного формализма, позволяющего избежать многократных преобразований систем отсчета.

Процедура синхронизации играет решающую роль в обсуждении скорости хода движущихся часов. Большинство недоразумений по поводу обсуждаемого вопроса (собственно вопрос, который мы собираемся обсуждать, и важен лишь из-за связанной с ним довольно большой путаницы) вызывается предположением о возможности сравнения скоростей хода движущегося и неподвижного хронометров. Однако это предположение неверно. Дело в том, что можно сравнивать показания только тех часов, которые находятся (почти) в одной и той же мировой точке, соответствующей некоторому событию. Но для измерения скорости хода нужно сделать два отсчета. В такой ситуации мы можем только сравнить скорость хода движущихся часов со скоростью хода *семейства* синхронизированных неподвижных часов, а для этого нужно иметь как минимум два неподвижных хронометра. Только если указанное условие не выполнено, ситуация, связанная со сравнением скорости хода часов, будет казаться симметричной и как следствие этого парадоксальной. Взгляните на рис. 8.8. На нем C_1 и C_2 — мировые линии неподвижных, а C — движущихся часов. Вычисления проще всего проделать в канонической системе отсчета, где часы C_1 и C_2 покоятся. Поскольку все наши операции ковариантны, в принципе годится любая инерциальная система отсчета (см. задачу 8.3).

[Будьте осторожны, если вы не знакомы с таким способом рассуждений. Мы неожиданно вводим определение и лишь затем доказываем, что оно корректно. Каноническая система отсчета используется здесь только для обоснования определения.]

Скорость хода движущихся часов

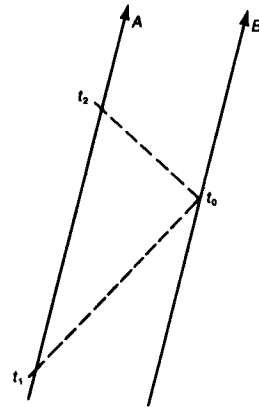


Рис. 8.7

Синхронизируем неподвижные часы с помощью описанной выше процедуры. Для простоты совместим начало координат с событием E , которое соответствует встрече часов C с часами C_1 . Поставим стрелки часов C и C_1 так, чтобы в момент события E они показывали нуль. Тогда часы C_2 будут показывать нуль в момент события G одновременно с E . Мы сможем сравнить скорости хода часов, если нам удастся определить показания часов C и C_2 в момент события F . Сначала найдем координаты события F . Уравнение мировой линии движущихся часов имеет вид

$$x = vt, \quad (8.2)$$

а мировой линии часов C_2 —

$$x = L, \quad (8.3)$$

где L — некоторая постоянная, являющаяся мерой расстояния между двумя неподвижными часами. Событие F лежит на обеих мировых линиях, поэтому его координаты должны удовлетворять обоим уравнениям. Следовательно, они имеют значения $(L, L/v)$.

Теперь можно вычислить интервалы времени между событиями E и F и между событиями G и F . Время τ_2 , которое показывают часы C_2 , в силу их неподвижности в точности равно координатному интервалу времени между интересующими нас событиями:

$$\tau_2 = \frac{L}{v}. \quad (8.4)$$

Движущиеся часы покажут время

$$\tau = \sqrt{\frac{L^2}{v^2} - L^2} = \frac{L}{v} \sqrt{1 - v^2}. \quad (8.5)$$

Тогда отношение скоростей хода часов равно частному

$$\frac{\tau}{\tau_2} = \sqrt{1 - v^2} \quad (8.6)$$

и не зависит от L . Итак, движущиеся часы отстают по сравнению с неподвижными. Надеюсь, вы понимаете, что, несмотря на простоту, с которой мы пришли к этому выводу, результат нельзя считать тривиальным.

С помощью нашего определения одновременности можно дать операциональное определение расстояния между точками, в которых произошли два события. Посмотрите на рис. 8.9. Договоримся, что расстояние между двумя точками, в которых

[Не забывайте, что мы измеряем расстояние в световых секундах, так что скорости измеряются в долях скорости света, т.е. безразмерны.]

[Вычисление промежутков времени между двумя известными событиями основано на применении формулы (5.2).]

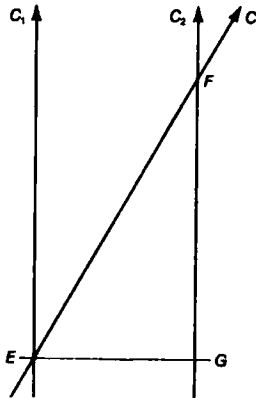


Рис. 8.8
Измерение скорости хода движущихся часов.

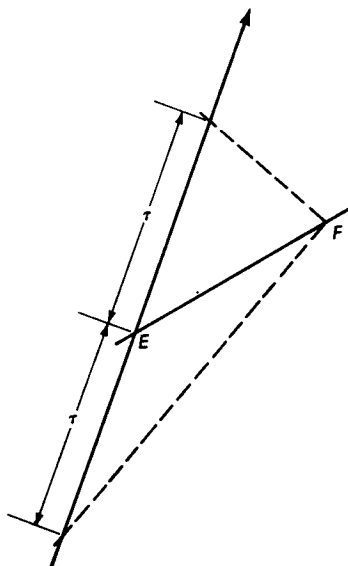


Рис. 8.9
Измерение расстояния между точками, в которых произошли два события E и F . В данном случае это расстояние равно τ световым секундам.

произошли события, определяется величиной, измеренной в канонической системе отсчета, где эти события одновременны. Здесь снова можно перейти к ковариантной форме операционального определения, напоминающего часто используемый на практике радарный способ измерения расстояний. Выберем любую мировую линию, по отношению к которой события E и F можно считать одновременными. Для этого воспользуемся описанным ранее методом равных углов, образуемых этой мировой линией с мировой линией светового сигнала. Определим расстояние между точками, в которых произошли события E и F , как половину интервала времени, прошедшего между отправлением и возвращением светового сигнала, совершившего «путешествие» туда и обратно. Тогда координаты событий, происшедших в точках, отстоящих от начала на единичном пространственном расстоянии, должны удовлетворять уравнению

$$x^2 - t^2 = 1. \quad (8.7)$$

Пространственное расстояние

Длина — не фундаментальная величина

Возможно, вы удивлены, что мы не включили в число наших первичных понятий пространственное расстояние, с которым связано физическое представление о твердом стержне. Разумеется, это можно было бы сделать. Однако часы являются более простым понятием. На пространственно-временной диаграмме часы описываются единственной мировой линией с нанесенными на ней отсчетами, тогда как твердый стержень описывается целой двумерной поверхностью. Кроме того, нет оснований считать, что невозможно создавать все более и более точные часы, тогда как сама природа материалов не позволяет делать все более и более твердые стержни.

Если принять такое хронометрическое определение длины, то возникает необходимость в ином логическом истолковании экспериментов по измерению, скажем, скорости света, например эксперимента Майкельсона — Морли, в котором сравнивались процессы распространения света в двух взаимно перпендикулярных направлениях. С точки зрения хронометрического толкования в этих экспериментах выявляется лишь свойство упругих тел сохранять при определенных условиях неизменные геометрические размеры, измеренные с помощью световых сигналов. Так как при построении системы отсчета мы пользуемся только часами и свободными частицами, твердые стержни не будут играть в этой книге никакой роли в любой проблеме.

[Отличие этой точки зрения от общепринятой полезно сравнить с отличием систем отсчета, специализированных по времени, от систем отсчета, специализированных по световому сигналу; см. стр. 63.]

ЗАДАЧИ

8.1. (08) Докажите, что на рис. 8.4 все события, лежащие на прямой AB , одновременны.

8.2. (10) Один из двух близнецов — служащий, а другой — водитель автобуса. При прочих равных условиях кто из них будет старше и на сколько, когда они выйдут на пенсию? Дайте количественную оценку.

8.3. (10) Преобразуйте чертеж на рис. 8.8 так, чтобы на нем была изображена каноническая система отсчета, в которой часы C покоятся.

8.4. (12) Докажите, что геометрическое построение на рис. 8.6 действительно согласуется с нашим предыдущим определением световой одновременности.

8.5. (13) Найдите, контрпример, показывающий, что в инерциальной системе отсчета общего типа метод равных углов не пригоден для построения линии одновременных событий.

8.6. (12) Убедитесь, что приведенное на рис. 8.7 построение га-

рантирует идентичность показаний обоих часов для одновременных событий.

8.7. (24) Две одинаковые неподвижные ракеты расположены друг от друга на расстоянии L и ориентированы вдоль соединяющей их прямой, как показано на рис. 8.10. В результате одновременного включения двигателей, обладающих одинаковой тягой, ракеты начинают ускоряться вдоль этой прямой. Нарисуйте соответствующую пространственно-временную диаграмму. На каком расстоянии друг от друга ракеты окажутся позднее при переходе в режим движения по инерции? Натянется ли соединяющий их трос длиной L ?

8.8. (23) Докажите, что высказывание «событие A произошло одновременно с событием B », которое мы условимся писать как $A =_s B$, является отношением эквивалентности; иными словами, докажите, что удовлетворяются следующие три условия:

- (1) $A =_s A$ (рефлексивность);
- (2) из $A =_s B$ следует $B =_s A$ (симметрия);
- (3) из $A =_s B$ и $B =_s C$ следует $A =_s C$ (транзитивность).

В условии (3) следует считать, что события A , B и C не лежат на одной прямой. Для упрощения доказательства воспользуйтесь симметрией относительно отражения времени $t \rightarrow -t$.

8.9. (30) Можно исключить зависимость показаний спецрелятивистских часов от пути в пространстве-времени, если договориться переносить часы только с постоянной скоростью, т. е. уподобить их в указанном смысле свободным частицам. Докажите с помощью рис. 8.11, что даже в таком случае хроно-

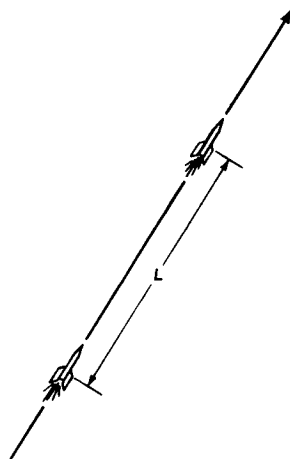


Рис. 8.10

Здесь представлено пространственное изображение, а не пространственно-временная диаграмма.

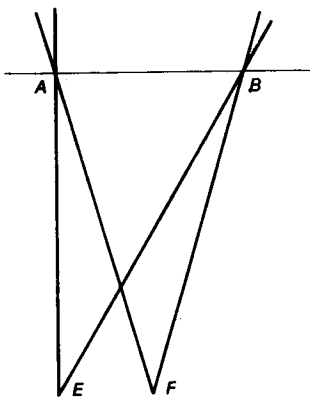


Рис. 8.11

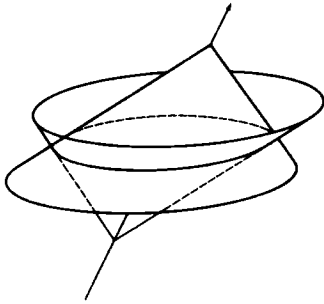


Рис. 8.12

транспортная одновременность остается неудовлетворительной.

8.10. (28) Расчитайте замедление хода гипотетических часов с хроноструктурой $\mathcal{S}$ вида

$$t^4 - x^4 = 1.$$

8.11. (21) На рис. 8.12 показано построение одновременных событий в $2 + 1$ -мерном пространстве-времени. Найдите уравнение линии пересечения изображенных там двух световых конусов и докажите, что это плоская кривая.

8.12. (35) С помощью предыдущего построения докажите, что с точки зрения неподвижного наблюдателя движущийся диск сокращается в направлении движения.

9. Относительная скорость

Перейдем теперь к обсуждению относительной скорости, которое будет проводиться в духе предыдущего раздела. Пусть известно, что следует понимать под скоростью частицы (описываемой некоторой мировой линией) по отношению к наблюдателю, покоящемуся в канонической системе отсчета. Тогда можно дать ковариантное операциональное определение этого понятия, пригодное в любой инерциальной системе отсчета. Заметьте, что можно определить относительную скорость движения двух частиц до обсуждения преобразований Лоренца. В большинстве книг по СТО относительная скорость рассматривается в связи с обсуждением результатов преобразований Лоренца. Это явное недоразумение, ибо определение относительной скорости не зависит от особенностей хроноструктуры $\mathcal{S}$.

В любой канонической системе отсчета применимо хорошо знакомое нам определение скорости как расстояния, пройденного за единицу времени. Чтобы найти скорость, достаточно сравнить наклон мировой линии частицы с наклоном мировой линии светового сигнала. На рис. 9.1 изображена мировая линия W частицы, движущейся относительно данной системы отсчета со скоростью

$$v = a/b. \quad (9.1)$$

Как превратить это определение в ковариантное операциональное определение? Сначала откажемся от представления о канонической системе отсчета, кроме ассоциированной с ней выделенной мировой линии, и будем теперь вести речь об относительной скорости частиц, которые описываются двумя миро-

Знакомое определение

[Не забывайте, что длины измеряются в световых секундах, а скорости — в долях скорости света.]

выми линиями. Нам нужно перейти от измерений наклонов к измерениям интервалов времени. Поскольку мировые линии световых сигналов в канонической системе отсчета составляет угол 45° , длины отрезков a и b можно перенести на выделен-

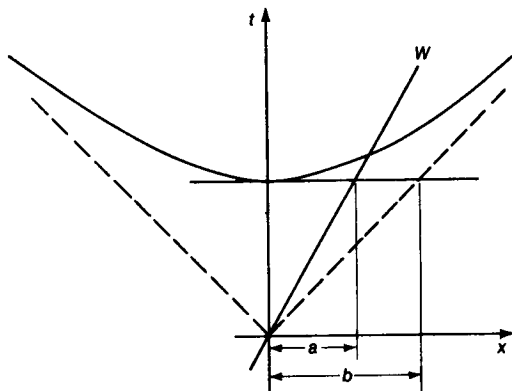


Рис. 9.1

Измерение скорости в канонической системе отсчета. Если мировая линия W составляет угол 30° с вертикалью, то соответствующая скорость равна $1/\sqrt{3}$.

ную мировую линию покоящегося наблюдателя, как показано на рис. 9.2. Затем можно измерить интервалы времени τ_1 и τ_2 , отмеченные на рис. 9.3; они связаны с a и b соотношениями

$$\tau_1 = b - a, \quad (9.2)$$

$$\tau_2 = b + a, \quad (9.3)$$

из которых следует, что

$$v = \frac{a}{b} = \frac{\tau_2 - \tau_1}{\tau_2 + \tau_1}. \quad (9.4)$$

Таким образом, нам удалось выразить относительную скорость через интервалы времени, определяемые световыми сигналами. Но это уже ковариантное определение, которым можно пользоваться в любой системе отсчета. С другой стороны, нетрудно заметить, что здесь требуется сравнение интервалов времени, измеренных только вдоль единственной мировой линии. Следовательно, относительная скорость зависит лишь от пространственно-временной структуры инерциальной системы

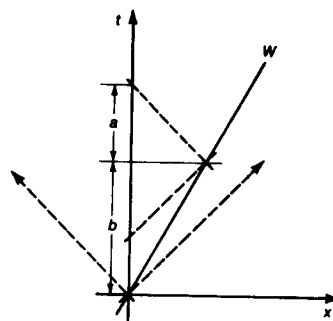


Рис. 9.2.

С помощью световых сигналов длины отрезков a и b , изображенных на рис. 9.1, можно измерять как промежутки времени.

отсчета, но не от хроноструктуры $\mathcal{S}$. Эту скорость вполне можно было бы назвать свето-сигнальной относительной скоростью.

Сложение скоростей

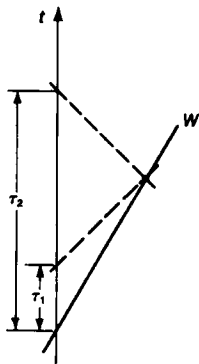


Рис. 9.3

Весьма примечателен закон сложения относительных скоростей. На рис. 9.4 схематически изображена ситуация, в которой частица с мировой линией A движется со скоростью $1/2$ относительно покоящейся в данной системе отсчета частицы с мировой линией W . Пусть частица с мировой линией B движется со скоростью $1/2$ относительно частицы с мировой линией A . Тогда оказывается, что скорость частицы B относительно частицы W будет равна не $1/2 + 1/2 = 1$, а, как явствует из диаграммы, примерно $4/5$. Такой своеобразный закон сложения относительных скоростей является следствием свето-сигнального определения относительной скорости, а не каким-то специфическим свойством СТО. Само собой разумеется, что если бы физические часы представляли собой часы, измеряющие абсолютное время, то нам пришлось бы принять иное определение относительной скорости.

Только что проделанные геометрические построения можно описать алгебраически. Если две частицы движутся по отношению к наблюдателю со скоростями v_1 и v_2 , то их относительная скорость равна

$$v = \frac{v_2 - v_1}{1 - v_1 v_2}. \quad (9.5)$$

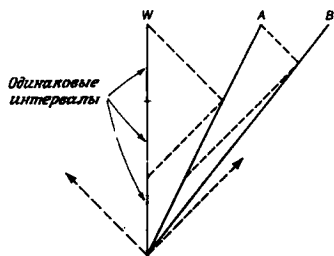


Рис. 9.4

Последнее выражение легко получается с помощью 4-векторов и четырехмерного аналога скалярного произведения, что будет показано в разд. 13 после того, как мы познакомимся с этими математическими объектами.

Как уже отмечалось, величина скорости неразрывно связана с наклоном мировой линии на пространственно-временной диаграмме. Поскольку в евклидовой геометрии наклон прямой характеризуется тангенсом угла, который эта прямая составляет с осью, вас не должно удивлять, что закон сложения (в данном случае его скорее следовало бы именовать законом вычитания) скоростей имеет некоторое сходство с формулой для тангенса разности двух углов:

$$\operatorname{tg}(A - B) = \frac{\operatorname{tg} A - \operatorname{tg} B}{1 + \operatorname{tg} A \operatorname{tg} B}. \quad (9.6)$$

Это, конечно, не полный аналог выражения (9.5), ибо геометрия пространства-времени отличается от евклидовой.

Внешнее сходство формул (9.5) и (9.6) явилось причиной упорных поисков величины, подобной углу в евклидовой геометрии. Дело в том, что в обычной геометрии никто не будет заниматься такой неблагодарной работой, как нахождение тангенсов углов наклона с помощью формул типа (9.6), когда можно воспользоваться элементарным правилом сложения углов, которое сводится к простому суммированию. Естественно возникает вопрос: нельзя ли найти такую функцию скорости, обозначим ее ψ , чтобы закон сложения скоростей свелся к простому суммированию этих функций? Это можно сделать; функция ψ , обладающая нужным нам свойством, называется *быстротой*¹⁾. По аналогии с формулой евклидовой геометрии

$$\mu = \operatorname{tg} \theta, \quad (9.7)$$

можно написать

$$v = \operatorname{th} \psi, \quad (9.8)$$

что и дает искомую связь. С помощью теории гиперболических функций (обратите внимание на название) можно показать, что такое определение приводит к закону

$$\psi = \psi_2 - \psi_1. \quad (9.9)$$

Для построений в евклидовой геометрии очень удобен транспортир — полукруг с нанесенными на нем градусными делениями. Удобно иметь аналогичный инструмент и для построения наглядных изображений в геометрии пространства-времени. Для этого возьмем единичную гиперболу $t^2 - x^2 = 1$ и нанесем на нее различные наклоны в единицах быстроты. Такой транспортир, схематически изображенный на рис. 9.5, может найти много применений. Поработайте немного с транспортиром. Это поможет вам глубже понять взаимосвязи в геометрии пространства-времени. Заметьте, что к транспортиру на рис. 9.5 можно добавить шкалу скоростей, нанесенную на прямую

$$t = 1. \quad (9.10)$$

Единица измерения ψ сходна с радианом, поэтому для малой быстроты

$$\psi \sim v. \quad (9.11)$$

Быстрота

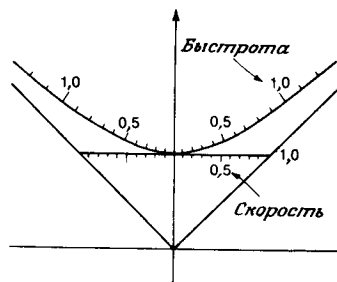


Рис. 9.5
Спецрелятивистский транспортир.

Спецрелятивистский транспортир

[Мы снова встретимся с быстротой в разд. 44, когда займемся изучением псевдосферы. Пространство всевозможных скоростей во всех направлениях — это и есть псевдосфера, а быстрота — удобная координата на ней.]

¹⁾ Употребляется также термин «параметр скорости». — Прим. перев.

ЗАДАЧИ

9.1. (10) Пусть в некоторой общей инерциальной системе отсчета задана гипербола $\mathcal{H}$ и точка на ней, соответствующая какому-то произвольному значению скорости. Используя евклидову геометрию, найдите способ построения на этой гиперболе точки, соответствующей удвоенному значению скорости. Объясните выполненные операции.

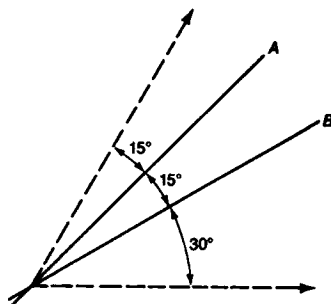


Рис. 9.6

9.2. (14) Убедитесь, что определение относительной скорости симметрично.

9.3. (18) Инвариантна ли относительная скорость по отношению к проективным преобразованиям?

9.4. (19) Чему равна относительная скорость частиц с мировыми линиями A и B на пространственно-временной диаграмме, показанной на рис. 9.6? Решите задачу графически и аналитически.

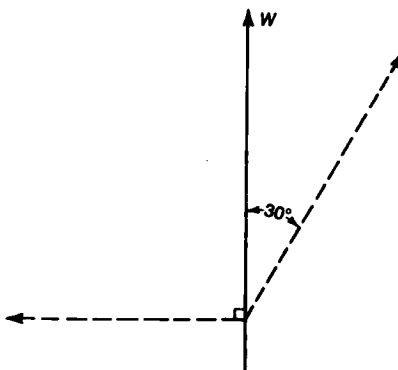


Рис. 9.7

9.5. (11) Начертите на пространственно-временной диаграмме, показанной на рис. 9.7, мировые линии частиц, движущихся относительно частицы с мировой линией W со скоростью $\pm \frac{1}{2}$.

9.6. (27) Сформулируйте правила работы со спецрелятивистским транспортером.

10. Лоренц-инвариантность

Мы осознаем изменение, только когда обнаруживаем нечто инвариантное, а неизменное — только когда имеем дело с чем-то преобразующимся.

Дж. Вейнберг

Теперь мы готовы исследовать исключительно важное свойство наших спецрелятивистских часов. Поскольку мы считаем хроноструктурную гиперболу $\mathcal{S}$ результатом экспериментальных наблюдений, можно утверждать лишь, что она характеризует часы в отдельной канонической системе отсчета, в которой описывались эксперименты с часами. Сейчас мы постараемся доказать, что если хроноструктура $\mathcal{S}$ является гиперболой в одной системе отсчета, то она будет такой же гиперболой и в любой другой канонической системе отсчета.

Последовательность доводов, которой мы здесь воспользуемся, может служить примером общего метода доказательства существования симметрии тех или иных объектов. Собственно, речь идет о преобразованиях, оставляющих эти объекты неизменными. Сфера обладает симметрией, она не изменяется при поворотах. Сфера с нанесенной на нее точкой обладает уже меньшей симметрией, ибо уменьшается число преобразований, не вызывающих никаких изменений. Такого рода доводы использовались нами при обсуждении вращательной симметрии евклидовой геометрии; мы вновь прибегнем к ним здесь при поиске лоренцевой симметрии спецрелятивистских часов. Кроме того, мы обратимся к ним в дальнейшем при обсуждении симметрии волн на воде и псевдосфер.

Начнем с представления о канонической системе отсчета, которое опирается на свойства световых сигналов, равномерность хода часов и выбор выделенной мировой линии. Это представление не зависит от конкретных особенностей поведе-

[Мы не будем делать различия между $\mathcal{S}$ и « $\mathcal{S}$ », т.е. между хроноструктурой и геометрическим местом точек на диаграмме пространства-времени, описывающим хроноструктуру.]

Доказательство существования симметрии

Операциональное определение канонической системы отсчета

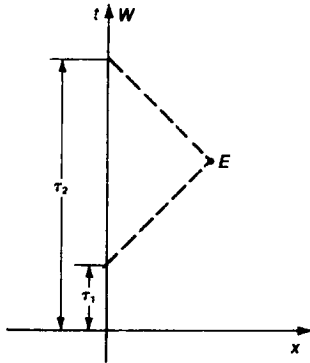


Рис. 10.1

Промежутки времени τ_1 и τ_2 являются светоподобными координатами события E по отношению к мировой линии W .

[Следует помнить, что ковариантные операции — это физические операции, которые имеют смысл вне зависимости от выбора какого-нибудь частного представления. Примерами могут служить испускание светового сигнала или измерение промежутка времени.]

ния часов. Здесь тем же способом, что и в предыдущих разделах, можно найти ковариантное операциональное определение канонической системы отсчета. Это определение позволяет преобразовать одну каноническую систему отсчета в другую. Если в исходной канонической системе отсчета хроноструктура $\mathcal{S}$ имеет вид гиперболы

$$t^2 - x^2 = 1, \quad (10.1)$$

то, как будет доказано, $\mathcal{S}$ останется той же самой гиперболой и в любой другой канонической системе отсчета. Это характерное свойство именно гиперболической хроноструктуры. В задачах в конце раздела вы столкнетесь с примерами отсутствия такой симметрии. Подобного рода симметрия хроноструктуры $\mathcal{S}$ называется лоренц-инвариантностью, а соответствующие преобразования — преобразованиями Лоренца.

Инерциальные системы отсчета и их особая разновидность — канонические системы отсчета — играют в СТО центральную роль. До сих пор у нас было лишь, так сказать, внутреннее определение инерциальных систем отсчета. Они определялись через свои свойства, а это более абстрактный и, возможно, менее доступный подход, нежели определение, основанное на четкой последовательности операций. Теперь, в духе предыдущих разделов, мы в состоянии дать ковариантное операциональное определение канонической системы отсчета, ассоциированной с мировой линией W . Для простоты будем по-прежнему работать в $1 + 1$ -мерном пространстве-времени. Способ рассуждения здесь полностью соответствует хорошо знакомой схеме: мы знаем, что следует делать в некоторой конкретной канонической системе отсчета; тогда, описав эту процедуру с помощью ковариантных операций, мы получаем определение, пригодное в любой инерциальной системе отсчета.

Мы ищем ковариантные операции, позволяющие определить координаты любого события E , происшедшего в произвольной мировой точке пространства-времени. В нашем распоряжении есть лишь световые сигналы, выделенная мировая линия W , связанная с интересующей нас конкретной канонической системой отсчета, и умение измерять интервалы времени. Все это показано на рис. 10.1. На нем изображены только световые сигналы, связанные с событием E , и только те интервалы времени τ_1 и τ_2 , которые выделяются этими сигналами. Интервалы τ_1 и τ_2 связаны с координатами события в канонической

системе отсчета следующими соотношениями:

$$\tau_1 = t - x, \quad (10.2)$$

$$\tau_2 = t + x. \quad (10.3)$$

Отсюда нетрудно получить выражения координат через эти интервалы:

$$x = \frac{1}{2}(\tau_2 - \tau_1), \quad (10.4)$$

$$t = \frac{1}{2}(\tau_2 + \tau_1), \quad (10.5)$$

а это и есть искомое ковариантное операциональное определение координат x и t канонической системы отсчета, связанной с мировой линией W .

Фактически, числа τ_1 и τ_2 сами являются прекрасными координатами. Будем называть их светоподобными координатами. Они образуют систему координат, которая получается в результате поворота на 45° обычных осей координат (x, t) .

Здесь возникает естественный вопрос: как связаны между собой координаты в двух различных канонических системах отсчета? Теперь, когда у нас есть четкое определение координат, ответить на этот вопрос не составляет особого труда.

Найдем преобразование от заданной канонической системы отсчета, в которой хроноструктура $\mathcal{S}$ имеет вид

$$t^2 - x^2 = 1, \quad (10.6)$$

к другой канонической системе отсчета, связанной с мировой линией наблюдателя, движущегося со скоростью v относительно наблюдателя с выделенной мировой линией в исходной системе отсчета. Ситуация изображена на рис. 10.2. Пусть τ_1 и τ_2 — светоподобные координаты в системе отсчета, связанной с W , а τ'_1 и τ'_2 — соответствующие координаты в системе, связанной с W' . Будем считать, что значения τ_1 и τ_2 известны, и найдем выражения координат τ'_1 и τ'_2 через τ_1 , τ_2 , и v — скорость движения наблюдателя W' по отношению к W . Теперь нам нужно найти координаты события A , задающего значения координат в системе отсчета, ассоциированной с мировой линией W' . Можно найти как (τ_1, τ_2) -координаты этого события, так и его (x, t) -координаты. Светоподобные координаты проще, но менее нам знакомы. Поэтому сначала проделаем расчет в координатах (x, t) , а затем повторим его для координат (τ_1, τ_2) .

Событие A лежит на мировой линии W' , откуда следует, что его координаты, которые мы обозначим $\bar{x}$ и $\bar{t}$, должны

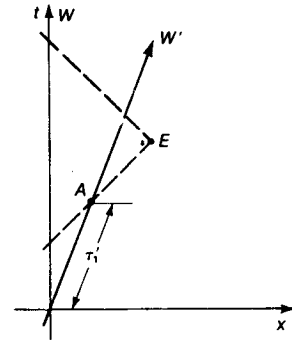


Рис. 10.2
Светоподобная координата τ'_1 события E по отношению к мировой линии W' .

Светоподобные координаты

[Вернитесь немного назад и вспомните все, что мы делали выше. Тогда вопросы, подобные этому, придут вам на ум сами собой.]

[Определение и обсуждение относительной скорости вы найдете в разд. 9.]

удовлетворять уравнению

$$\bar{x} = v\bar{t}. \quad (10.7)$$

Но оно лежит и на мировой линии светового сигнала, так что

$$\bar{t} = \bar{x} + \tau_1. \quad (10.8)$$

Таким образом, событие A имеет координаты

$$\bar{x} = \frac{v\tau_1}{1-v}, \quad (10.9)$$

$$\bar{t} = \frac{\tau_1}{1-v}. \quad (10.10)$$

[Вычисление промежутков времени между двумя известными событиями основано на применении формулы (5.2).]

Величина τ_1' — это просто интервал времени, прошедшего от начала отсчета до события A . Тогда

$$(\tau_1')^2 = \bar{t}^2 - \bar{x}^2, \quad (10.11)$$

$$(\tau_1')^2 = \frac{(\tau_1)^2}{(1-v)^2} (1-v^2); \quad (10.12)$$

т. е.

$$\tau_1' = \sqrt{\frac{1+v}{1-v}} \tau_1. \quad (10.13)$$

Аналогичные выкладки дают

$$\tau_2' = \sqrt{\frac{1-v}{1+v}} \tau_2. \quad (10.14)$$

С помощью наших определений

$$\tau_1 = t - x, \quad (10.15)$$

$$\tau_2 = t + x, \quad (10.16)$$

$$\tau_1' = t' - x', \quad (10.17)$$

$$\tau_2' = t' + x', \quad (10.18)$$

в результате довольно утомительных алгебраических преобразований выражения (10.13) и (10.14) можно записать в виде уравнений для координат x , t и x' , t' . Решив эти уравнения относительно t' и x' , получим

Преобразования Лоренца

$$t' = \frac{1}{\sqrt{1-v^2}} (t - vx), \quad (10.19)$$

$$x' = \frac{1}{\sqrt{1-v^2}}(x - vt). \quad (10.20)$$

Эти преобразования от координатной канонической системы отсчета к другой называются преобразованиями Лоренца.

Цель этой книги — дать ковариантную трактовку максимально возможного числа интересующих нас объектов. Поскольку ковариантность позволяет работать в любой системе отсчета, мы не станем приспособливать преобразования Лоренца к специальному классу систем отсчета, как это обычно делается. В принципе можно было бы вообще обойтись без преобразований Лоренца.

В качестве примера полезности преобразований Лоренца приведем упоминавшееся ранее доказательство корректности правила равенства углов при построении линии событий L , одновременных по отношению к наперед заданной мировой линии W . Начнем с канонической системы отсчета, связанной с этой мировой линией. В ней L и W взаимно перпендикулярны и на самом деле составляют одинаковые углы с мировыми линиями световых сигналов. Теперь обратимся к преобразованиям Лоренца для светоподобных координат, т. е. к формулам (10.13) и (10.14). Согласно этим формулам, одна светоподобная координата сжимается, а вторая в том же отношении растягивается. Очевидно, однако, что растяжение одной светоподобной координаты не изменяет углов, поскольку они, как показано на рис. 10.3, расположены по отношению к ней симметрично. Следовательно, при преобразованиях Лоренца углы, образуемые прямыми W и L с мировыми линиями световых сигналов, остаются равными. Таким образом, правило равенства углов остается в силе в любой канонической системе отсчета, к которой можно перейти в результате преобразований Лоренца. Кроме того, никогда ни одна мировая линия какой-либо частицы не наблюдалась вне светового конуса, а все мировые линии внутри конуса можно преобразовать друг в друга с помощью преобразований Лоренца. Рис. 10.4 поясняет последнее утверждение.

Обратите внимание на простоту преобразований Лоренца для светоподобных координат. Читатель, знакомый с линейной алгеброй, несомненно скажет, что светоподобные координаты полезны потому, что они диагонализуют преобразования Лоренца. Чтобы хорошо разобраться в СТО, необходимо научиться строить пространственно-временные диаграммы для одной и той же ситуации в различных канонических системах

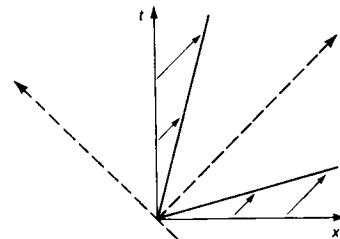


Рис. 10.3

Одновременность

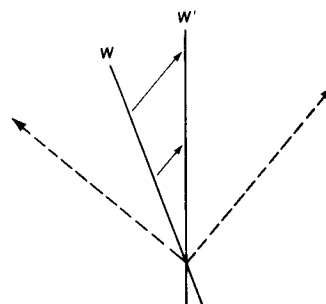


Рис. 10.4

[Световой конус описан в разд. 6 на стр. 61.]

отсчета. Для этого требуется выполнять преобразования Лоренца, что проще всего делать с помощью светоподобных координат. В качестве упражнения повторим проделанный выше вывод преобразований Лоренца, но теперь уже с помощью светоподобных координат. Как вы помните, мы искали координаты события A , показанного на рис. 10.2. Сейчас мы найдем светоподобные координаты этого события, которые обозначим $\tilde{\tau}_1$ и $\tilde{\tau}_2$. Из (10.7) при учете (10.4) и (10.5) следует

$$(1 - v)\tilde{\tau}_2 = (1 + v)\tilde{\tau}_1. \quad (10.21)$$

Условие того, что события A и E лежат на одной мировой линии светового сигнала, выражается равенством

$$\tilde{\tau}_1 = \tau_1, \quad (10.22)$$

где τ_1 — одна из светоподобных координат события E . В силу приведенных соотношений существует следующая связь между координатами событий A и E :

$$\tilde{\tau}_2 = \frac{1 + v}{1 - v} \tau_1. \quad (10.23)$$

Для дальнейших расчетов необходимо переписать в светоподобных координатах правило, связывающее интервалы времени. В таких координатах оно сводится к весьма компактному выражению

$$\tau^2 = \tau_1 \tau_2, \quad (10.24)$$

где τ — промежуток времени между началом системы отсчета и событием (τ_1, τ_2). Отсюда видно, что промежуток времени между началом координат и событием A , т. е. новая координата τ'_1 , определяется выражением

$$\tau'_1 = \sqrt{\frac{1 + v}{1 - v}} \tau_1 \quad (10.25)$$

в согласии с результатами предыдущих вычислений.

Обратите внимание на простоту формулы для хроноструктуры $\mathcal{S}$ в светоподобных координатах:

$$\tau_1 \tau_2 = 1. \quad (10.26)$$

С помощью преобразований Лоренца (10.13) и (10.14) можно показать, что уравнение кривой, являющейся геометрическим местом тех же самых событий в новой системе отсчета, име-

ет вид

$$\tau_1' \tau_2' = 1, \quad (10.27)$$

а это все та же гипербола. Таким образом, описание часов в движущейся канонической системе отсчета подчиняется тому же самому правилу. Иными словами, канонические системы отсчета равноправны. В этом и заключается лоренц-инвариантность, которую мы обещали продемонстрировать. В качестве упражнения постарайтесь доказать, что уравнение светового конуса имеет вид

$$\tau_1 \tau_2 = 0 \quad (10.28)$$

и что оно тоже инвариантно. Собственно, последнее вам уже должно быть известно, поскольку условие постоянства скорости света с самого начала введено в определение канонической системы отсчета.

Выше мы обещали доказать, что невозможно дать определение одновременности, пользуясь только часами и световыми сигналами. Все дело в том, что и часы и световые сигналы лоренц-инвариантны. То, что можно сделать в одной канонической системе отсчета, можно проделать и в любой другой, но беда в том, что прямая, на которой лежат одновременные события, не является лоренц-инвариантной. Нельзя дать непротиворечивое ковариантное определение одновременности только с помощью часов и световых сигналов. Чтобы выйти из тупика, необходимо каким-то способом выделить некоторую конкретную каноническую систему отсчета. Мы это сделали, потребовав, чтобы определение одновременности помимо всего прочего зависело еще и от выбора некоторой выделенной мировой линии.

После того как мы установили путем экспериментальных наблюдений лоренц-инвариантность часов, у нас возникает естественное желание обобщить этот результат. Наиболее широким, но тем не менее правдоподобным обобщением следует считать утверждение, что все физические процессы лоренц-инвариантны. Ни одна физическая система не позволяет измерить абсолютную скорость. На самом деле имеет смысл говорить лишь о скорости наблюдателя по отношению к некоторой системе отсчета. Вселенная заполнена микроволновым излучением, и мы можем измерить такую важную величину, как наша скорость по отношению к этому электромагнитному фону.

В педагогических целях мы часто будем рассматривать теории, в которых лоренц-инвариантность нарушается. Хочу на-

Важное обобщение

помнить читателю, что нет никакой необходимости в подобных модификациях СТО. Специальная теория относительности позволяет дать адекватное объяснение всем экспериментальным наблюдениям. И все же для более полного понимания истинного смысла СТО полезно рассмотреть такие альтернативные теории. Они играют роль пробного камня, на котором можно отточить логику наших рассуждений, касающихся СТО.

ЗАДАЧИ

[Вы должны уметь самостоятельно выдумывать для себя шаблонные задачи, подобные этим двум.]

10.1. (12) Рассмотрите простое преобразование Лоренца, в результате которого τ_1 удваивается, а τ_2 уменьшается в два раза. Какой быстроте это соответствует?

10.2. (13) Прodelайте опущенные в тексте вычисления интервала τ_2' .

10.3. (13) Восстановите недостающие звенья в цепочке выкладок, ведущих от формул (10.13), (10.14) к формулам (10.19), (10.20).

10.4. (23) Постройте три пространственно-временные диаграммы, описывающие наблюдателя, который измеряет замедление хода времени движущихся часов: а) для системы отсчета, неподвижной по отношению к наблюдателю; б) для системы, неподвижной по отношению к часам; в) для системы отсчета, по отношению к которой наблюдатель и часы движутся с одинаковыми, но направленными в противоположные стороны скоростями. Покажите на каждой диаграмме, почему скорость хода движущихся часов кажется меньше.

10.5. (20) Найдите преобразование Лоренца, превращающее изображенный на рис. 10.5 параллелограмм в евклидов квадрат на пространственно-временной диаграмме. Это всего лишь упражнение, ибо такой квадрат не имеет физического смысла. Какая быстрота ψ соответствует такому преобразованию?

10.6. (26) Прочтите раздел, посвященный релятивистскому сокращению длин, в любом учебнике специальной теории относительности. Объясните его на нашем языке. Воспользуйтесь тремя диаграммами пространства-времени, описанными в задаче 10.4.

10.7. (20) Перепишите уравнение

$$t^4 - x^4 = 1$$

в светоподобных координатах.

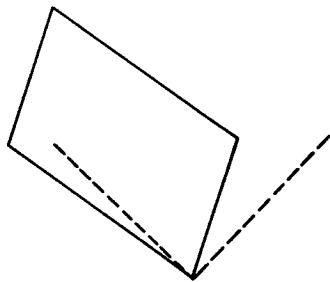


Рис. 10.5

10.8. (15) Докажите, что гипотетические часы, которые описываются хроноструктурой $\mathcal{S}$ вида

$$t^4 - x^4 = 1$$

не обладают никакой симметрией, подобной лоренцевой. Покажите, что не существует преобразования от одной системы отсчета к другой, движущейся с иной скоростью, которое оставило бы неизменной хроноструктуру $\mathcal{S}$.

10.9. (36) Рассмотрите канонические координаты и преобразования Лоренца для гипотетических часов из задачи 10.8. Преобразования Лоренца полезны даже в том случае, когда поведение часов не является лоренц-инвариантным!

10.10. (28) Неподвижный стержень длиной L расположен под углом θ к оси x . Опишите мировые линии этого стержня и его расположение в системе отсчета, движущейся вдоль оси x .

10.11. (33) Найдите и проанализируйте преобразования, которые сохраняют структуру мировых линий световых сигналов, но не спецрелятивистскую хроноструктуру $\mathcal{S}$.

10.12. (22) Покажите, что четырехмерные преобразования

$$t' = \frac{1}{\sqrt{1-v^2}}(t - vx),$$

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

$$y' = y,$$

$$z' = z,$$

сохраняют 4-мерную хроноструктуру вида

$$t^2 - x^2 - y^2 - z^2 = 1.$$

10.13. (24) Обладает ли какой-либо симметрией хроноструктура $\mathcal{S}$, показанная на рис. 10.6?

10.14. (30) Ответьте на тот же вопрос для хроноструктуры $\mathcal{S}$, которая представляет собой параболу, проходящую через начало координат.

10.15. (30) Хроноструктура $\mathcal{S}$ представляет собой кривую

$$t^n = x^n,$$

где n — целое число. Обладает ли она какой-нибудь симметрией?

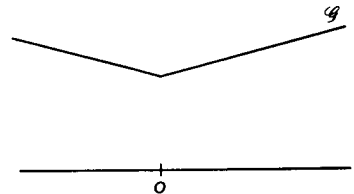


Рис. 10.6

10.16. (30) Обладает ли какой-нибудь симметрией хроноструктура $\mathcal{L}$, если она представляет собой кривую

$$t^n = x^m,$$

где n и m — целые числа?

11. Непротиворечивость специальной теории относительности

[Этот эпиграф взят из прекрасного рассказа [19].]

Что поделаешь, если люди не хотят понимать Относительность.

Урсула Легуин

Из всех разделов физики только термодинамика привлекает большее число чудаковатых ниспровергателей истин, чем специальная теория относительности. Правильно ли поступают ученые, игнорируя их критику? Не следует ли проанализировать каждое критическое замечание в соответствии с его научной ценностью, чтобы из-за присущего нам консерватизма не прозевать следующую научную революцию? Нет, не следует.

Причина такого категорического ответа заключается в том, что большая часть критики СТО касается непротиворечивости этой теории, а не подтверждающих ее экспериментов. Можно игнорировать все подобные заявления о непоследовательности специальной теории относительности, поскольку имеются все основания считать, что она логически непротиворечива. Хотя всякая физическая теория устанавливает соответствие между объектами физического мира и математическими структурами, вопрос о непротиворечивости — это вопрос лишь о непротиворечивости математической модели, и, как и на многие вопросы математического характера, на него можно дать убедительный ответ.

Парадоксы

Наша модель СТО построена на концепциях, заимствованных из евклидовой геометрии, но с иными правилами обращения с ними. Если эти правила непротиворечивы, то наша модель должна быть столь же непротиворечивой, как и евклидова геометрия. Другими словами, любой логический парадокс специальной теории относительности можно превратить в логический парадокс, затрагивающий евклидову геометрию.

Здесь мы используем тот же самый ход рассуждений, какого придерживался Анри Пуанкаре для разрешения вопроса о непротиворечивости неевклидовой геометрии. Неевклидова гео-

Неевклидова геометрия

метрия является обобщением евклидовой геометрии, в котором отвергается аксиома о параллельных. Согласно этой аксиоме, через точку, не лежащую на данной прямой, проходит только одна прямая, лежащая в одной плоскости с данной прямой и не пересекающая ее. Предпринималось множество попыток доказать эту аксиому на основе других аксиом, т. е. доказать, что отказ от нее алогичен¹⁾. Пуанкаре разрешил эту проблему, построив основанную на евклидовой геометрии модель неевклидовой геометрии²⁾. Фактически он построил две модели, но одна из них представляет для нас особый интерес. Как выяснится в дальнейшем, она является моделью открытой расширяющейся Вселенной.

Эта модель была построена на верхней половине плоскости (x, y) (т.е. для точек $y > 0$) и называется моделью Пуанкаре на полуплоскости. Точки в неевклидовой геометрии — это просто точки на плоскости. Однако прямые в неевклидовой геометрии уже не являются прямыми линиями на плоскости, поэтому их название будет заключено в кавычки. «Прямая» — это любая евклидова окружность с центром на оси x . Такие «прямые» удовлетворяют всем аксиомам евклидовой геометрии, за исключением аксиомы о параллельных. Через любую точку, лежащую вне данной «прямой», можно провести множество не пересекающих ее «прямых», каждая из которых теперь может считаться параллельной данной (рис. 11.1). Если бы аксиома о параллельных была необходимой, такая модель была бы противоречивой, что в конечном счете свелось бы к различного рода противоречиям, касающимся понятия окружности в евклидовой геометрии. Таким образом, Пуанкаре показал, что неевклидова геометрия непротиворечива в такой же степени, как и евклидова, которая непротиворечива настолько, насколько мы можем об этом судить. Наша модель специальной теории относительности в точности подобна вышеупомянутой модели неевклидовой геометрии. Она отличается от евклидовой геометрии не новым определением прямой линии, а использованием другой кривой для измерений — гиперболы вместо окружности.

Полуплоскость Пуанкаре

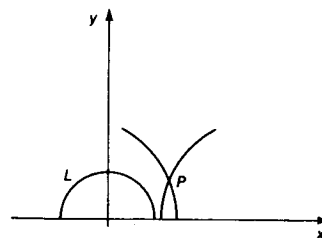


Рис. 11.1

¹⁾ С драматической историей построения неевклидовой геометрии читатель может познакомиться по книге: Ливанова А. Три судьбы. Постигание мира. — М.: Знание, 1969. — Прим. ред.

²⁾ Подробнее об этом написано в книгах: Кокстер Г. С. Введение в геометрию. — М.: Наука, 1966; Гильберт Д., Кон-Фоссен С. Наглядная геометрия. — М.: Наука, 1981. — Прим. ред.

Релятивистские парадоксы

Типичный спецрелятивистский парадокс связан со словесным описанием ситуации и ее обсуждением в двух различных инерциальных системах отсчета. Хотя мы отдаем себе отчет в том, что нет никакой необходимости в обсуждении ситуации более чем в одной системе отсчета, тем не менее такой анализ часто весьма поучителен и, конечно же, не приносит никакого вреда. Мы показали, что все первичные понятия специальной теории относительности — события, свободные частицы, часы и световые сигналы — подчиняются одним и тем же правилам в любой канонической системе отсчета. Интервал времени между фиксированными событиями можно вычислять различными способами и в различных системах отсчета, но результат будет получаться один и тот же. Чтобы в разных системах отсчета получились различные результаты, в чем обычно и заключается парадокс, где-то должна вкратце вкрасться ошибка. Возможно, что ошибка заключается в неверном математическом описании рассматриваемой ситуации. Операция, с помощью которой словесное описание ситуации приводится к математической форме, довольно сложна и требует соблюдения определенных правил. Это наиболее вероятное место, где следует искать ошибку.

Другим благодатным источником релятивистских парадоксов является использование концепций, смысл которых хорошо знаком, но которые не имеют корректного определения в рамках СТО. В качестве простых примеров «правонарушителей» такого рода можно привести понятия твердого стержня и потенциальной энергии.

Вы должны во всеоружии математики бесстрашно сражаться с любыми релятивистскими парадоксами. Опорой в этой борьбе должна служить непротиворечивость теории относительности. Однако все это ни в коем случае не означает, что по моему мнению СТО совершенно свободна от неясностей или не нуждается в дальнейшем экспериментальном исследовании. Я очень далек от такой точки зрения и постарался изложить СТО так, чтобы читателю были ясны способы ее экспериментальной проверки, и вместе с тем вооружить его формализмом, пригодным для описания даже тех ситуаций, когда в СТО следует внести некоторые малые поправки. Описанная здесь пространственно-временная структура позволяет рассмотреть ситуацию и в том случае, если СТО окажется неверной; при этом не потребуется радикального изменения представлений.

Модель волн на воде

Еще большую уверенность в непротиворечивости СТО дает обсуждение волн на воде, которым мы займемся в разд. 28.

Такая знакомая физическая система, как волны на глубокой воде, служит еще одной моделью реализации логической структуры теории относительности. В этой реализации часы, световые сигналы и свободные частицы вновь играют роль первичных понятий, удовлетворяющих таким же постулатам, что и в СТО. Однако хроноструктура будет иной, но такой, что в силу присущей ей симметрии движущиеся наблюдатели вновь окажутся равноправными. Многие спецрелятивистские парадоксы, если они не касаются твердых стержней, могут быть непосредственно описаны с помощью этой модели.

Моделирование теории относительности путем рассмотрения волн на воде позволяет выявить изъян в обычных рассуждениях, касающихся движения быстрее света. Действительно, движение со сверхсветовой скоростью не согласуется со СТО — это верно, но не существенно. Всякого рода затруднения в математических рассуждениях подчас связаны с тем, что небольшие изменения в исходных посылках не обязательно приводят только к малым изменениям в окончательных выводах. Модель волн на воде можно усовершенствовать, если учесть в ней слабое влияние поверхностного натяжения. Это оказывается достаточным, чтобы модель допускала некий тип движения, эквивалентного движению быстрее света. Аналогично в спецрелятивистскую хроноструктуру $\mathcal{S}$ можно было бы внести такие малые изменения, чтобы она, оставаясь в согласии с имеющимися экспериментами, тем не менее допускала движение быстрее света.

Еще не перевелись авторы, которые опаснее даже чудачков, воюющих с теорией относительности. Они «выводят» СТО и преобразования Лоренца из, как им кажется, совершенно безобидных предположений, что приводит к опасной путанице в логической структуре теории. Если вас прельщают их доводы, то обязательно прочтите историю неевклидовой геометрии: такие с виду безобидные предположения, как, скажем, возможность существования треугольников произвольной площади, на деле оказываются противоречащими аксиомам неевклидовых геометрий. Суть не в том, чтобы свести основы теории к наименьшему числу добротных предположений, а в том, чтобы свести их к предположениям, которые можно проверить экспериментально.

Отличие принятого здесь подхода от общепринятых лучше всего раскрывается на примере парадокса близнецов. Мы попытались здесь показать, как поведение часов можно исследовать с помощью экспериментов, результаты которых вначале

Движение быстрее света

[Эти мысли, касающиеся движения быстрее света, отличаются от представлений о тахионе, которые иногда обсуждаются. Кое-что о тахионах вы найдете в разд. 29.]

«Доказательства» специальной теории относительности

Парадокс близнецов

[Ответ на задачу 5.2 получил название парадокса близнецов, или парадокса часов, поскольку некоторые считают его совершенно немыслимым. Этот парадокс отлично проанализирован в книге Тейлора и Уилера [40], что нельзя сказать о многих других книгах, где он обсуждается. Однако поучительны обсуждения и того и другого типа.]

кажутся совершенно неожиданными. Нет никакого преимущества в аксиоматическом разрешении парадокса близнецов. Аксиоматический подход — это всего лишь более искусный, но туманный способ переформулировать эффект. В нашем подходе он заложен в основы специальной теории относительности.

ЗАДАЧА

11.1. (28) На подоконнике окна, раскрытого на ширину L , сидит мышь. Мимо на ракете проносится бессердечный мальчишка и бросает в мышь палку длиной L . Мышь в силу лоренцева сокращения длины видит более короткую палку и приходит в отчаяние. Мальчишка же видит сократившийся промежуток между створками окна и приходит к выводу, что мышь осталась цела. Однако следует думать, что смерть мыши — понятие инвариантное. В чем ошибка в этих рассуждениях? (Scientific American, April 1975, p. 126).

12. 4-векторы

Векторная алгебра

Чтобы получить не только качественные, но и количественные решения тех или иных задач, нужны более совершенные математические инструменты, в особенности когда число пространственных измерений больше одного. Собственно, нам нужен векторный анализ, пригодный для работы в пространстве-времени. Векторная алгебра уже была приспособлена к евклидовой геометрии; теперь нужно приспособить векторную алгебру и к специфическим свойствам спецрелятивистской геометрии. С помощью такой векторной алгебры в этом разделе дается простой вывод закона сложения скоростей, в разд. 13 обсуждается доплеровское смещение и абберрация света, а в разд. 14 — импульс и энергия.

Модель векторного пространства для пространства-времени

Векторы здесь используются двумя различными способами. С одной стороны, когда в пространстве-времени нужно задать события, применяются радиус-векторы, выходящие из некоторой произвольно выбранной начальной точки. С другой стороны, для описания смещений от одних событий к другим используются свободные векторы. Линейная структура векторного пространства позволяет переносить вектор к любому событию с помощью правила параллелограмма. Свободные векторы в пространстве-времени мы будем называть 4-векторами.

Свободные векторы в пространстве-времени

Для количественного описания векторов выделим систему базисных векторов и будем записывать любой вектор в виде их

линейной комбинации. В СТО наиболее удобна система базисных векторов, связанная с канонической системой отсчета. В каждой канонической системе отсчета есть своя система базисных векторов. Примем для этого особого класса базисных векторов обозначения $\hat{x}$, $\hat{y}$, $\hat{z}$ и $\hat{t}$. 4-вектор $\hat{x}$ описывает смещение от начала системы отсчета к событию $(t, x, y, z) = (0, 1, 0, 0)$. Кроме того, он описывает смещение от любого события (t, x, y, z) к событию $(t, x + 1, y, z)$. Все это по аналогии распространяется и на остальные три 4-вектора. Базис, связанный с канонической системой отсчета именно таким способом, мы будем называть ортонормированным базисом.

Любой 4-вектор можно разложить по этому базису. Коэффициенты разложения входят в соответствующую формулу следующим образом:

$$a = a^x \hat{x} + a^y \hat{y} + a^z \hat{z} + a^t \hat{t}. \quad (12.1)$$

Чтобы умножить вектор на масштабный множитель k , следует умножить на k все его компоненты. Чтобы сложить два 4-вектора, нужно сложить соответствующие компоненты.

Пример

Вектор b , описывающий смещение от начала системы отсчета к событию $(1, 1, 0, 0)$, записывается в виде

$$b = \hat{x} + \hat{t} \quad (12.2)$$

и имеет компоненты

$$\begin{aligned} b^x &= 1, & b^z &= 0, \\ b^y &= 0, & b^t &= 1. \end{aligned} \quad (12.3)$$

Мы можем указать направление мировой линии, задав касательный к ней 4-вектор. Касательным вектором к прямой мировой линии является любой вектор, соединяющий различные события, лежащие на этой прямой.

Пример

Мировой линии, описываемой уравнением

$$x = y = z = 0, \quad (12.4)$$

касается 4-вектор $\hat{t}$, а 4-вектор b из предыдущего примера касает

[Чтобы отличать эти 4-векторы от 3-векторов, мы не будем выделять их жирным шрифтом, а просто договоримся, что в этом разделе буквы a , b , σ , всевозможные λ и символы $\hat{x}$, $\hat{y}$, $\hat{z}$, $\hat{t}$ используются только для обозначения 4-векторов.]

Базисные векторы

Компоненты

[Здесь a^x обозначает x -компоненту вектора a , а не x -ю степень некоторого числа a . В формуле (12.1) a , $\hat{x}$, $\hat{y}$, и $\hat{t}$ представляют собой 4-векторы.]

Касательные векторы

ется мировой линии светового сигнала

$$x = t, \quad y = z = 0. \quad (12.5)$$

Ковариантность по отношению к преобразованиям Лоренца

В различных канонических системах отсчета события описываются по-разному. Это означает, что один и тот же 4-вектор тоже должен иметь разные представления в разных канонических системах отсчета. В различных канонических системах отсчета используются различные ортонормированные базисы. Поскольку мы определили 4-векторы через события, закон их преобразования должен быть таким же, как закон преобразования событий, т. е. представлять собой преобразования Лоренца. Вектор, описываемый в одной канонической системе отсчета выражением

$$a = a^x \hat{x} + a^y \hat{y} + a^z \hat{z} + a^t \hat{t}, \quad (12.6)$$

в другой системе отсчета, движущейся относительно первой системы вдоль оси x со скоростью v , будет описываться выражением

[См. задачу 10.11.]

$$a = \frac{1}{\sqrt{1-v^2}} \left[(a^x - va^t) \hat{x}' + (a^t - va^x) \hat{t}' \right] + a^y \hat{y}' + a^z \hat{z}', \quad (12.7)$$

где $(\hat{x}', \hat{y}', \hat{z}', \hat{t}')$ — базисные векторы в движущейся системе отсчета.

Скалярное произведение

Специфические свойства евклидовой геометрии хорошо отражаются в операции скалярного произведения. Мы можем воспользоваться аналогичным приемом и в спецрелятивистской геометрии. Дадим следующее определение скалярного произведения 4-векторов, которое будем записывать как $a \cdot b$:

Скалярным произведением 4-векторов называется выражение

$$a \cdot b \equiv a^x b^x + a^y b^y + a^z b^z - a^t b^t, \quad (12.8)$$

где $a^x, b^x, \dots, a^t, b^t$ — компоненты векторов в ортонормированном базисе.

Промежутки времени

Нетрудно понять, почему скалярное произведение следовало определить именно таким образом. Это определение позволяет очень просто написать формулу для промежутка времени в СТО.

Промежуток времени в СТО: промежуток времени τ , отмеренный по часам, которые переносятся вдоль отрезка мировой линии, задаваемого векторами a , определяется формулой

$$\tau^2 = -a \cdot a. \quad (12.9)$$

Это скалярное произведение, подобно евклидову скалярному произведению, ковариантно. Его значение зависит только от самих векторов и может быть вычислено в любой канонической системе отсчета. Ясно, что такая ковариантность обуславливается ковариантностью промежутков времени. Можно непосредственным расчетом показать, что преобразования Лоренца не изменяют значение скалярного произведения. Если бы мы следовали традиционной схеме и постулировали преобразования Лоренца, то ход рассуждений нужно было бы изменить на противоположный. Мы должны были бы выбрать такое определение скалярного произведения, чтобы оно было инвариантно по отношению к преобразованиям Лоренца. Правомерен и тот и другой подход.

Преобразования Лоренца

Скалярное произведение позволяет дать компактную формулировку различных законов специальной теории относительности. Например, мировая линия светового сигнала описывается касательным 4-вектором, удовлетворяющим условию

Световые сигналы

$$\sigma \cdot \sigma = 0. \quad (12.10)$$

Согласно нашему определению, касательный вектор мировой линии может иметь любую длину. Можно так подобрать длину касательного вектора, чтобы он удовлетворял некоторым удобным условиям; такой особый касательный вектор мы будем называть 4-скоростью.

4-скорость

4-скорость — это касательный 4-вектор λ , нормированный так, что

$$\lambda \cdot \lambda = -1, \quad (12.11)$$

$$\lambda' > 0. \quad (12.12)$$

[Везде, где возможно, мы будем обозначать 4-скорость для мировой линии часов буквой λ , а касательный 4-вектор мировой линии светового сигнала — буквой σ .]

Непосредственно из определения промежутка времени в СТО следует, что если величина λ представляет собой 4-скорость, то отрезку мировой линии, который описывается вектором $\tau\lambda$, соответствует промежуток времени τ .

[Под квадратом вектора v здесь понимается скалярное произведение $v \cdot v$.]

Векторы, квадраты которых положительны, касаются мировых линий, описывающих движение со скоростями больше скорости света. Они называются пространственноподобными векторами. Векторы, квадраты которых отрицательны, называются времениподобными. Наконец, векторы, квадраты которых равны нулю, называются изотропными. Это касательные векторы мировых линий световых сигналов. Несмотря на то что их квадраты равны нулю, сами векторы нулю не равны.

Примеры

Какая 4-скорость соответствует мировой линии, описывающей движение со скоростью v ? Пусть уравнения мировой линии имеют вид

$$\begin{aligned}x &= vt, \\ y &= z = 0.\end{aligned}\tag{12.13}$$

Касательный вектор a запишем в форме

$$a = v\hat{x} + \hat{t}.\tag{12.14}$$

Его нормировка

$$a \cdot a = v^2 - 1\tag{12.15}$$

отличается от нормировки 4-вектора, но вектор

$$\lambda = \frac{1}{\sqrt{1-v^2}}(v\hat{x} + \hat{t})\tag{12.16}$$

параллелен a и нормирован как надо; следовательно, λ — это и есть искомая 4-скорость. Мировой линии покоящейся частицы соответствует 4-скорость $\hat{t}$.

Скалярное произведение двух любых 4-скоростей — величина инвариантная. Кроме того, оно должно быть связано с относительной скоростью движения частиц, описываемых соответствующими мировыми линиями, или, в явном виде,

Относительная скорость

$$\lambda_1 \cdot \lambda_2 = -\frac{1}{\sqrt{1-v^2}}.\tag{12.17}$$

Откуда взялась эта формула? Скалярное произведение можно

вычислять в любой системе отсчета, поэтому вычислим его в простой системе отсчета, где одна из интересующих нас 4-скоростей соответствует мировой линии неподвижной частицы. Обратите внимание, этот 4-вектор не равен нулю! Теперь формулы двух 4-скоростей можно заимствовать из только что рассмотренного примера. Чтобы найти скалярное произведение этих 4-скоростей, следует воспользоваться спецрелятивистским определением (12.8). Так как скалярное произведение ковариантно, расчет, выполненный в любой другой канонической системе отсчета, даст тот же самый результат.

С помощью этого результата можно дать очень короткий вывод закона сложения скоростей. Предположим, что наблюдатель следит за двумя частицами, движущимися со скоростями v_1 и v_2 . Согласно (12.16), их 4-скорости в канонической системе отсчета, по отношению к которой наблюдатель неподвижен, имеют вид

$$\lambda_1 = \frac{1}{\sqrt{1 - (v_1)^2}} (v_1 \hat{x} + \hat{t}), \quad (12.18)$$

$$\lambda_2 = \frac{1}{\sqrt{1 - (v_2)^2}} (v_2 \hat{x} + \hat{t}). \quad (12.19)$$

Относительная скорость v , с которой вторая частица движется по отношению к первой, удовлетворяет соотношению (12.17):

$$-\frac{1}{\sqrt{1 - v^2}} = \lambda_1 \cdot \lambda_2 \quad (12.20)$$

или

$$-\frac{1}{\sqrt{1 - v^2}} = \frac{v_1 v_2 - 1}{\sqrt{1 - (v_1)^2} \sqrt{1 - (v_2)^2}}, \quad (12.21)$$

откуда после элементарных алгебраических преобразований (включающих правильный выбор знака квадратного корня) получим

$$v = \frac{v_2 - v_1}{1 - v_1 v_2}, \quad (12.22)$$

что совпадает с формулой, приведенной в разд. 9.

Представление о касательном векторе можно распространить и на искривленные мировые линии. Для этого воспользуемся предельным переходом к столь коротким отрезкам мировой линии, что их можно считать прямыми, для которых легко

[Упражнением в рассуждениях подобного сорта в рамках евклидовой геометрии может служить задача 12.8.]

Закон сложения скоростей

Касательные к искривленным мировым линиям

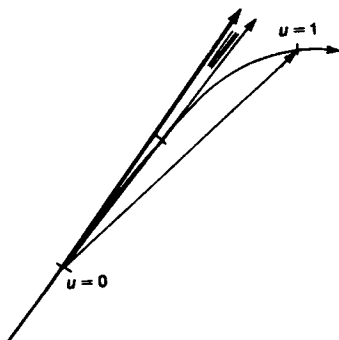


Рис. 12.1

Процесс предельного перехода при построении касательного вектора в точке $u = 0$.

[Здесь опущено довольно много промежуточных шагов, но все они просты и не очень поучительны, так что вы можете спокойно принять этот результат на веру. Более полный вывод дан в разд. 26.]

Собственное время

построить касательные векторы. Проще всего работать с кривой, заданной в *параметрической форме*; другими словами, кривую можно задать в виде четырех функций, являющихся координатами событий, которые лежат на этой кривой:

$$x = X(u), \quad z = Z(u), \quad (12.23)$$

$$y = Y(u), \quad t = T(u).$$

На языке отображений (см. разд. 1, стр. 36) параметризованная кривая определяется таким отображением P , что

$$P: \mathbb{R} \rightarrow V; u \mapsto P(u). \quad (12.24)$$

Тогда касательный вектор v определяется формулой

$$v \equiv \lim_{\epsilon \rightarrow 0} \frac{P(u + \epsilon) - P(u)}{\epsilon}. \quad (12.25)$$

Каждое значение $P(u)$ представляет собой событие, а различие в положениях событий в пространстве-времени описывается вектором смещения, так что в правой части формулы (12.25) на самом деле стоит вектор. Процесс предельного перехода схематически изображен на рис. 12.1. Касательный 4-вектор v можно разложить по базисным векторам

$$v = \frac{dX}{du} \hat{x} + \frac{dY}{du} \hat{y} + \frac{dZ}{du} \hat{z} + \frac{dT}{du} \hat{t}. \quad (12.26)$$

Следует иметь в виду, что одну и ту же кривую можно параметризовать различными способами. При изменении параметризации изменяется и длина касательного вектора. Однако если

$$v \cdot v = -1, \quad (12.27)$$

то касательный вектор представляет собой 4-скорость. Параметр кривой, для которого выполняется указанное условие, называется *собственным временем*, ибо промежутки времени в этом случае численно равны изменению параметра.

ЗАДАЧИ

12.1. (14) Убедитесь, что спецрелятивистское скалярное произведение ковариантно по отношению к преобразованиям Лоренца.

12.2. (20) Проверьте справедливость закона преобразования компонент 4-вектора при преобразованиях Лоренца, вычислив касательный вектор мировой линии до и после выполнения преобразований.

12.3. (22) Покажите, как в пространстве-времени разложить 2-векторы по изотропному базису, соответствующему светоподобным координатам (τ_1, τ_2) .

12.4. (34) По аналогии с тем, как это делалось для 4-скорости, дайте определение 4-ускорения, чтобы оно описывало изменение 4-скорости по отношению к собственному времени. Найдите 4-ускорения частиц (при $u = 0$) с мировыми линиями

$$u \mapsto (\cos \Omega u, 0, 0, u),$$

$$u \mapsto (0, \sin \Omega u, 0, u),$$

$$u \mapsto (\cos \Omega u, \sin \Omega u, 0, u).$$

[Внимание! Параметр u — не обязательно собственное время.]

12.5. (18) Докажите, что скалярное произведение 4-ускорения и 4-скорости, ассоциированных с данной мировой линией, тождественно равно нулю.

12.6. (30) Какова мировая линия частицы, движущейся вдоль прямой с постоянным по величине 4-ускорением? Как сильно «состарится» мир к моменту вашей кончины, если вы всю жизнь будете двигаться с постоянным ускорением, равным $1g$? (Для упрощения вычислений используйте гиперболические синус и косинус.)

12.7. (16) Докажите, что ракета, движущаяся все время с постоянным ускорением относительно мгновенно сопутствующей ей инерциальной системы отсчета, способна неограниченно долго опережать световой сигнал. Насколько раньше светового сигнала она должна стартовать, чтобы это стало возможным?

12.8. (31) Определим диздрическое произведение трех векторов в евклидовой геометрии следующим образом:

$$Q(a, b, c) \equiv \frac{(b \cdot b)(a \cdot c) - (b \cdot a)(b \cdot c)}{[(b \cdot b)(a \cdot a) - (b \cdot a)^2]^{1/2} [(b \cdot b)(c \cdot c) - (b \cdot c)^2]^{1/2}}.$$

Докажите, что оно обладает следующими свойствами:

1) инвариантно по отношению к изменениям длин векторов a , b и c ;

2) инвариантно по отношению к изменениям вектора a , оставляющим a в плоскости, содержащей векторы a и b , т. е.

$$Q(a + kb, b, c) = Q(a, b, c).$$

Таким образом, диэдрическое произведение зависит только от двух пересекающихся плоскостей, которые определяются этими векторами. Единственная геометрическая величина, связанная с двумя такими плоскостями, — это образуемый ими двугранный угол θ . Докажите с помощью расчетов в частной системе координат, что имеет место равенство

$$\cos \theta = Q(a, b, c).$$

13. Доплеровское смещение

Мы теперь достаточно подготовлены, чтобы дать простой и полный анализ доплеровского смещения. Это не только хороший пример работы с 4-векторами, но и хорошая практика. Позднее нам придется выполнять аналогичные расчеты в искривленном пространстве-времени. Систематическое красное смещение, наблюдаемое в спектрах далеких галактик, свидетельствует о динамически эволюционирующей Вселенной — важнейшем представлении современной космологии. Мы также кратко остановимся на явлении абберации света — смещении видимых положений объектов, обусловленном движением наблюдателя. Обсуждение абберации является хорошим упражнением, хотя оно и не имеет прямого отношения к космологии.

Доплеровское смещение

Доплеровское смещение — это изменение частоты периодического сигнала вследствие относительного движения источника и приемника. Рассмотрим доплеровское смещение частоты только для света, считая периодический сигнал последовательностью световых сигналов. Собственно, нужны лишь два сигнала. Предположим, что это два следующих друг за другом гребня волны, которая и является периодическим сигналом. Такая идеализированная картина схематически изображена на рис. 13.1.

Геометрия пространства-времени

[Следует помнить, что изотропным называется 4-вектор, удовлетворяющий условию $\sigma \cdot \sigma = 0$, и что это не нулевой вектор.]

Для описания рассматриваемой ситуации нужны три различных 4-вектора. Движения источника и приемника определяются двумя векторами 4-скорости, а их относительное расположение в пространстве-времени — изотропным вектором σ , который соединяет события, соответствующие испусканию и приему светового сигнала. Упрощенная геометрическая схема эксперимента показана на рис. 13.2. Будем считать, что период

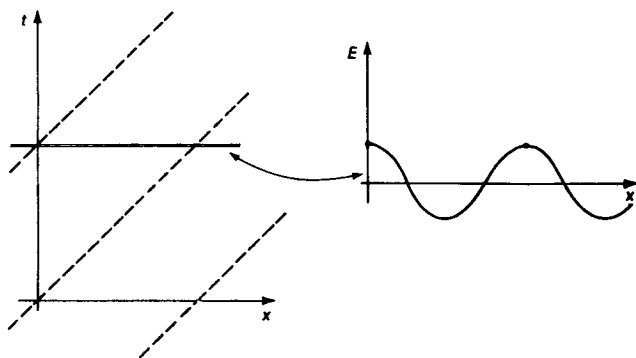


Рис. 13.1

Слева изображена пространственно-временная диаграмма для электромагнитной волны, распространяющейся вдоль оси x вправо. Штриховые мировые линии соответствуют гребням волн. Справа показано изменение амплитуды волны E как функции x в сечении пространственно-временной диаграммы, отмеченном двойной чертой.

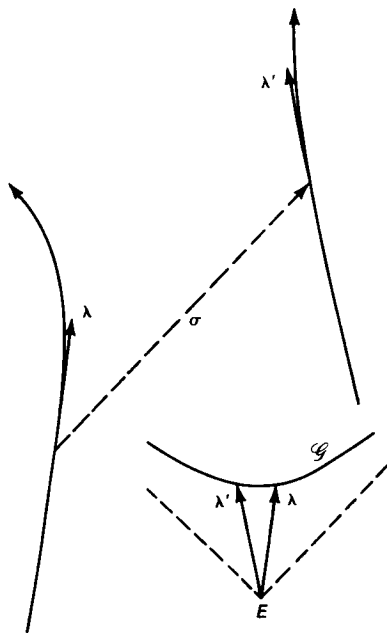


Рис. 13.2

Геометрия эксперимента по измерению доплеровского смещения. Метрическая фигура $\mathcal{G}$, заданная в момент события E , показывает, что векторы λ и λ' нормированы как подобает 4-скоростям.

рассматриваемых колебаний столь мал, что источник и приемник не успевают сильно сместиться за соответствующий промежуток времени. Пусть τ — промежуток собственного времени источника, равный его периоду, а τ' — соответствующий интервал собственного времени наблюдателя, расположенного рядом с приемником. Мы хотим найти отношение τ' к τ при условии, что эти величины очень малы. Геометрическое построение для двух следующих друг за другом световых сигналов схематически представлено на рис. 13.3. В пространстве-времени мы имеем четырехугольник, две противоположные стороны которого образованы отрезками мировых линий световых сигналов. Две другие стороны — это отрезки мировых линий источника и приемника. Следует иметь в виду, что диаграмма не может быть плоской, за исключением случая, когда источник и приемник движутся только вдоль соединяющей их прямой.

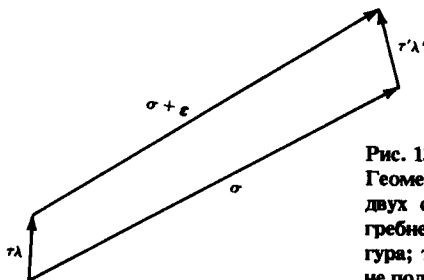


Рис. 13.3
Геометрическое построение для двух следующих друг за другом гребней волны. Это не плоская фигура; так, световые сигналы идут не под углом в 45° к вертикали.

Из чертежа видно, что интересующая нас фигура будет замкнутой, если

$$\tau\lambda + \sigma + \epsilon = \tau'\lambda' + \sigma, \quad (13.1)$$

[Напоминаем: если λ представляет собой 4-скорость, то отрезок мировой линии, соответствующий промежутку времени τ , определяется вектором $\tau\lambda$. См. стр. 100.]

где σ — изотропный 4-вектор, касательный к мировой линии первого светового сигнала, а $\sigma + \epsilon$ — изотропный 4-вектор мировой линии второго светового сигнала, отличающийся от σ на малый вектор ϵ , который стремится к нулю по мере неограниченного уменьшения величин τ' и τ . Итак, мы имеем две 4-скорости

$$\lambda \cdot \lambda = -1, \quad (13.2)$$

$$\lambda' \cdot \lambda' = -1 \quad (13.3)$$

и два изотропных вектора

$$\sigma \cdot \sigma = 0, \quad (13.4)$$

$$(\sigma + \varepsilon) \cdot (\sigma + \varepsilon) = 0. \quad (13.5)$$

Спецрелятивистское скалярное произведение подчиняется обычному закону дистрибутивности операции умножения, так что последнюю формулу можно записать в виде

$$(2\sigma \cdot \varepsilon) + (\varepsilon \cdot \varepsilon) = 0. \quad (13.6)$$

В пределе, при стремлении ε к нулю, квадрат вектора ε становится меньше других членов, и мы имеем

$$\sigma \cdot \varepsilon \rightarrow 0. \quad (13.7)$$

Если обе части равенства (13.1) умножить скалярно на σ , то получим

$$\tau(\sigma \cdot \lambda) = \tau'(\sigma \cdot \lambda'), \quad (13.8)$$

откуда сразу следует искомое отношение

$$\frac{\tau'}{\tau} = \frac{(\sigma \cdot \lambda)}{(\sigma \cdot \lambda')}. \quad (13.9) \quad \text{Результат}$$

Как мы и ожидали, в это выражение входят только три вектора λ , λ' и σ .

[В разд. 23 дан другой вывод формулы.]

Пример 1

Найдем доплеровское смещение в частном случае движения по прямой. Этот случай настолько прост, что позволяет получить даже графическое решение. Будем работать в канонической системе отсчета, связанной с источником. Соответствующая пространственно-временная диаграмма приведена на рис. 13.4. На каждом краю диаграммы начерчена гипербола

Движение вдоль прямой

$$t^2 - x^2 = \tau^2, \quad (13.10)$$

где τ — период источника. На рис. 13.5 дано увеличенное изображение области пространства-времени вблизи наблюдателя и отмечены промежутки времени τ и τ' . Как видно из чертежа, скорость наблюдателя примерно равна 0,43, а соответствующее доплеровское смещение $\tau'/\tau = 1,6$.

Теперь решим эту задачу аналитически. 4-скорость источника — это просто $\hat{t}$, а 4-скорость движущегося наблюдателя —

$$\lambda' = \frac{1}{\sqrt{1-v^2}}(v\hat{x} + \hat{t}). \quad (13.11)$$

[По поводу этой формулы см. стр. 100.]

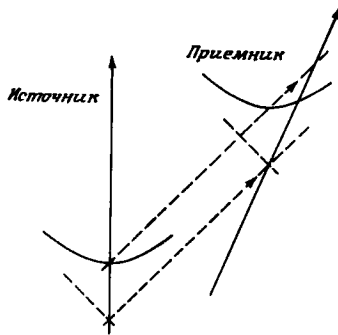


Рис. 13.4

Здесь изображены только те мировые линии и метрические фигуры, которые имеют непосредственное отношение к рассматриваемому физическому процессу. Оси координат не показаны умышленно, поскольку в них нет необходимости. Чтобы найти времениподобное направление, достаточно взглянуть на метрические фигуры.

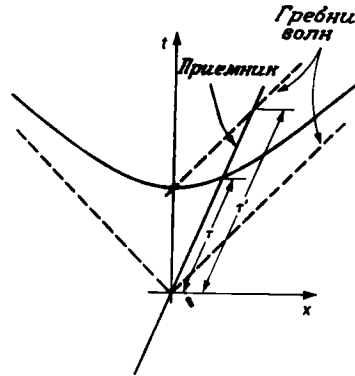


Рис. 13.5

Увеличенное изображение области пространства-времени, прилегающей на рис. 13.4 к мировой линии приемника.

Обратите внимание, что, согласно формуле (13.9), величина доплеровского смещения зависит от направления вектора σ , но не от его длины. Правильно направленный вектор описывается формулой

$$\sigma = \dot{t} + \dot{x}, \quad (13.12)$$

откуда следует, что

$$\frac{\tau'}{\tau} = \frac{\sigma \cdot \dot{t}}{\sigma \cdot \lambda'} = \frac{\sqrt{1-v^2}}{1-v} = \sqrt{\frac{1+v}{1-v}}. \quad (13.13)$$

Вращающийся диск

Пример 2

Заслуживает внимания еще одна задача, а именно расчет доплеровского смещения частоты источника, расположенного на ободе маховика, когда приемник находится на его оси. Пусть маховик радиусом a вращается с угловой скоростью Ω по отношению к канонической системе отсчета. Такого рода эксперименты были выполнены на самом деле, и их результаты подтвердили выводы СТО.

Нам снова нужно найти три вектора λ , λ' и σ . Соответствующая геометрическая конфигурация приведена на рис. 13.6. Мировая линия приемника — это просто прямая

$$x = y = z = 0. \quad (13.14)$$

Мировой линией источника является пространственно-временная винтовая линия

$$\begin{aligned}x &= a \cos \Omega t, \\y &= a \sin \Omega t, \\z &= 0.\end{aligned}\quad (13.15)$$

Рассмотрим световой сигнал, испущенный в момент времени $t = 0$. Уравнения его мировой линии имеют вид

$$\begin{aligned}x &= a - t, \\y &= z = 0.\end{aligned}\quad (13.16)$$

Это правильная мировая линия, поскольку она удовлетворяет следующим трем условиям: 1) ее касательный вектор изотропен; 2) она проходит через событие $(a, 0, 0, 0)$, соответствующее послышке сигнала; 3) она пересекает мировую линию приемника. Мы можем найти касательные векторы к этим мировым линиям, преобразовав уравнения последних к очевидным параметрическим формам. Уравнения мировой линии источника можно записать в виде

$$\begin{aligned}x &= a \cos \Omega u, \\y &= a \sin \Omega u, \\z &= 0, \\t &= u,\end{aligned}\quad (13.17)$$

и т. д.

В результате получаются следующие формулы для касательных векторов:

$$\lambda' \propto \hat{t}, \quad (13.18)$$

$$\lambda \propto \hat{t} - a\Omega \sin \Omega t \hat{x} + a\Omega \cos \Omega t \hat{y}, \quad (13.19)$$

$$\sigma = \hat{t} - \hat{x}. \quad (13.20) \quad [\text{См. формулу (12.26).}]$$

Здесь стоят знаки пропорциональности вместо знаков равенства, поскольку нормировка пока не обсуждалась. Нормированные 4-скорости имеют вид

$$\lambda' = \hat{t}, \quad (13.21)$$

$$\lambda = \frac{1}{\sqrt{1 - a^2 \Omega^2}} (\hat{t} + a\Omega \hat{y}). \quad (13.22)$$

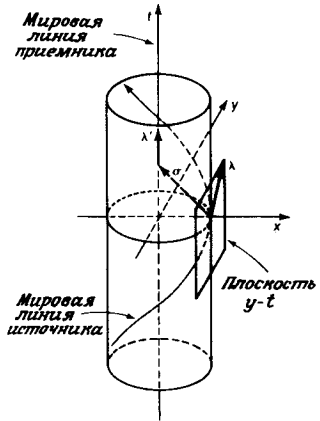


Рис. 13.6
Трехмерная пространственно-временная диаграмма.

Вектор σ не нуждается в нормировке. В результате несложных расчетов, связанных в основном с вычислениями скалярных произведений, получается следующая формула для доплеровского смещения частоты:

$$\frac{\tau'}{\tau} = \frac{\sigma \cdot \lambda}{\sigma \cdot \lambda'} = \frac{1}{\sqrt{1 - a^2 \Omega^2}}. \quad (13.23)$$

Итак, единственная трудность, которая остается при использовании нашего 4-векторного формализма, связана с построением пространственно-временной диаграммы, адекватной рассматриваемой ситуации.

Аберрация

Общая стратегия, которой мы старались следовать в предыдущих разделах, сводилась к поиску ковариантных описаний, пригодных в любой системе отсчета. Такая стратегия исключает необходимость в выполнении множества преобразований Лоренца, что обычно практикуется в различных руководствах по СТО. Вывод формулы доплеровского смещения (13.9) — прекрасный пример такой стратегии. И раз уж мы занимаемся светом, не будет лишним рассмотреть еще один пример. Каков угол между направлениями движения двух световых сигналов с точки зрения того или иного наблюдателя? Наблюдатели, движущиеся с различными скоростями, обнаружат разные углы между двумя одними и теми же лучами света. Такое явление называется *абберацией* света. Чтобы получить не только качественный, но и количественный ответ на поставленный вопрос, необходимо рассмотреть соответствующую задаче геометрическую конфигурацию. Вся необходимая нам геометрическая информация исчерпывается заданием трех векторов: 4-скорости наблюдателя и двух изотропных векторов, касательных к мировым линиям световых сигналов. Следующий ниже расчет служит хорошим упражнением в использовании 4-векторов для получения корректной формулы интересующего нас явления.

Будем рассуждать в том же духе, что и раньше. Сначала обсудим поставленную задачу в некоторой частной канонической системе отсчета. Затем с помощью 4-векторов сформулируем полученный результат в таком виде, чтобы он не зависел от выбора системы отсчета. На рис. 13.7 приведена диаграмма рассматриваемой ситуации в канонической системе отсчета, связанной с наблюдателем. На диаграмме показаны два вектора σ и σ' , касательные к мировым линиям световых сигналов.

Геометрия пространства-времени

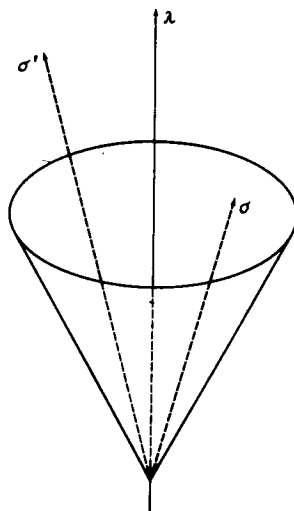


Рис. 13.7

Трехмерная пространственно-временная диаграмма. Конус образован мировыми линиями выходящими из начала координат световых сигналов, к двум из которых относятся векторы σ и σ' .

Что следует понимать под углом между ними? Если посмотреть вдоль транспорта в направлении двух источников света, то в принципе можно точно измерить интересующий нас угол, хотя, конечно, это довольно грубый метод. Иллюстрирующая такой метод пространственная диаграмма и соответствующая трехмерная пространственно-временная диаграмма показаны на рис. 13.8. Направления распространения световых сигналов

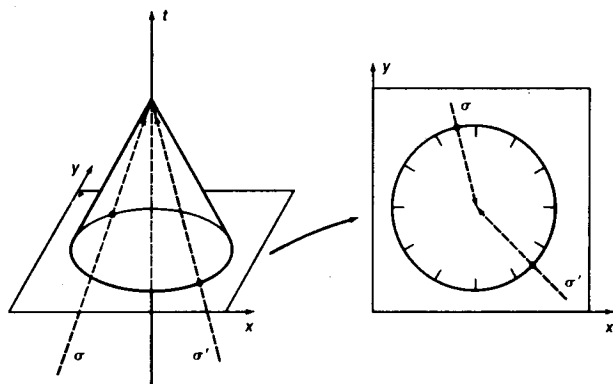


Рис. 13.8

Измерение углов между изотропными векторами σ и σ' с помощью проецирования этих 4-векторов на плоскость $x - y$. Отдельно показано сечение плоскостью проекции.

определяются пространственными составляющими касательных векторов. Составляющая вектора σ , лежащая вдоль 4-скорости наблюдателя, имеет вид $-(\sigma \cdot \lambda)\lambda$. Чтобы найти пространственную составляющую, нужно из вектора σ вычесть его составляющую, лежащую вдоль оси времени. Если обозначить пространственную составляющую символом $\sigma_{\perp}$, то

$$\sigma_{\perp} = \sigma + (\sigma \cdot \lambda)\lambda. \quad (13.24)$$

[Обратите внимание на знак минус. Он возникает из-за нормировки $\lambda \cdot \lambda = -1$.]

Этот вектор удовлетворяет условию

$$\sigma_{\perp} \cdot \lambda = 0. \quad (13.25)$$

Угол θ , который, по мнению наблюдателя, образуют между собой направления распространения двух световых сигналов, должен быть равен углу между векторами $\sigma_{\perp}$ и $\sigma'_{\perp}$. Но для векторов, у которых нет пространственных составляющих, спецрелятивистское скалярное произведение совпадает с евклидовым, так что можно воспользоваться формулой евклидовой геометрии (3.4):

$$\cos \theta = \frac{\sigma_{\perp} \cdot \sigma'_{\perp}}{\sqrt{(\sigma_{\perp} \cdot \sigma_{\perp})(\sigma'_{\perp} \cdot \sigma'_{\perp})}}. \quad (13.26)$$

После упрощений мы приходим к искомому ковариантному выражению

$$(\cos \theta - 1) = \frac{\sigma \cdot \sigma'}{(\sigma \cdot \lambda)(\sigma' \cdot \lambda)}, \quad (13.27)$$

с различными применениями которого вы познакомитесь в процессе решения задач.

ЗАДАЧИ

13.1. (20) Наблюдаемая яркость источника света зависит как от количества фотонов, поступающих в единицу времени, так и от энергии каждого фотона. Энергия фотона пропорциональна его частоте. Как яркость движущегося источника связана с его яркостью в состоянии покоя?

13.2 (21) С помощью графического метода, использованного в примере на стр. 107, нанесите на диаграмме против каждого значения скорости соответствующее значение доплеровского смещения для случая, когда источник и приемник движутся вдоль соединяющей их прямой.

13.3. (19) Выведите с помощью графического метода формулу доплеровского смещения в пределе малых относительных скоростей источника и приемника, движущихся вдоль одной прямой.

13.4.(10) Докажите, что в пределе малых относительных скоростей источника и приемника, движущихся вдоль одной прямой, часы, измеряющие абсолютное время, обнаружат одно и то же доплеровское смещение.

13.5. (14) Докажите, что в случае, когда источник и приемник движутся с постоянной относительной скоростью вдоль одной прямой, рассуждения на стр. 106 не нуждаются в предположении о малости τ .

13.6.(14) Некоторый наблюдатель, движущийся со скоростью v по направлению

$$\hat{x} + \hat{y} + \hat{z},$$

измеряет угол между направлениями движения двух фотонов, летящих к нему вдоль базисных векторов $\hat{x}$ и $\hat{y}$. Каков результат его измерений?

13.7.(24) Найдите величину угла θ' , измеренного движущимся наблюдателем между направлениями на звезду, лежащую на линии его движения, и на звезду, которую неподвижный наблюдатель видит под углом θ к этой линии (рис. 13.9). С помощью этого расчета покажите, что закон преобразования координат на небесной сфере

$$(\theta, \phi) \mapsto (\theta', \phi'),$$

принадлежащих двум различным системам отсчета, имеет вид

$$\operatorname{tg} \frac{\theta'}{2} = \sqrt{\frac{1-v}{1+v}} \operatorname{tg} \frac{\theta}{2},$$

$$\phi = \phi',$$

13.8.(25) Какова связь между формулой (13.26) и диэдрическим произведением, определенным в задаче 12.8?

13.9.(33) Докажите, что отображение

$$(\theta, \phi) \mapsto (\theta', \phi'),$$

определенное в задаче 13.7, конформно. Иными словами, покажите, что оно сохраняет углы и переводит бесконечно малые окружности в другие бесконечно малые окружности.

13.10. (35) Докажите, что в результате того же самого отобра-

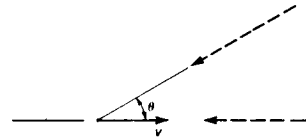


Рис. 13.9

Это не пространственно-временная диаграмма, а диаграмма в обычном пространстве.

жения конечные окружности переходят снова в конечные окружности.

13.11.(25) Получите то же самое отображение с помощью преобразований Лоренца.

13.12.(39) Допустим, что вам надоела наша Галактика и вы решили ее покинуть, двигаясь перпендикулярно ее плоскости с ускорением $1g$. Опишите, как со временем будет меняться вид на удаляющуюся Галактику. (См. [16].)

14. 4-импульс

Далеко идущее обобщение

В том виде, как мы ее до сих пор излагали, специальная теория относительности — это теория, предназначенная для описания часов, световых сигналов и свободных частиц. Было показано, что выбранный нами способ описания всех этих объектов обладает лоренцевой симметрией, в силу которой все состояния движения равноправны. Не существует выделенного состояния, которое можно было бы считать абсолютным покоем. Теперь можно рискнуть сделать следующее далеко идущее обобщение: возможно, вся физика обладает такой симметрией. Это смелый шаг, и нужно отдать дань восхищения проныцательности и решительности Эйнштейна, который сделал его, имея на то гораздо меньше оснований, чем мы имеем теперь. На сегодняшний день нет ни одного факта, противоречащего этому всеобъемлющему обобщению, поэтому его можно положить в основание любой физической теории. Это обобщение претерпевает небольшие изменения в общей теории относительности, где лоренцева симметрия имеет лишь локальный характер. Для больших промежутков времени и расстояний постулаты СТО следует модифицировать. Но СТО по-прежнему должна давать правильное локальное описание — это утверждение часто называют принципом эквивалентности.

Столкновения частиц

Описание динамики столкновения частиц можно без особых затруднений привести в соответствие с существованием лоренцевой симметрии. В ньютоновой механике столкновения подчиняются законам сохранения энергии и импульса. А как это выглядит с пространственно-временной точки зрения? К счастью, обобщение производится элементарно. Будем описывать частицу 4-вектором, определяемым направлением ее мировой линии, но таким, что его пространственные компоненты связаны с обычным импульсом, а временная компонента — с энергией. Этот 4-вектор называется 4-импульсом и представляет собой

4-импульс

физическую величину с лоренцевой симметрией. Следовательно, это ковариантная величина. Как и раньше, для построения ковариантных описаний легче всего пользоваться 4-векторными соотношениями. Если 4-импульс p связан только с 4-скоростью λ , то единственное соотношение, которое мы можем написать, имеет вид

$$p = m\lambda, \quad (14.1)$$

где постоянный коэффициент пропорциональности m не зависит от λ . Выясним смысл этого соотношения в случае, когда скорость частицы мала по сравнению со скоростью света. 4-скорость такой частицы

$$\lambda = \frac{1}{\sqrt{1-v^2}} (v + \hat{t}) \quad (14.2)$$

с хорошей степенью точности можно записать в виде

$$\lambda \approx v + \left(1 + \frac{v^2}{2}\right)\hat{t}, \quad (14.3)$$

откуда следует, что

$$p = mv + (m + \frac{1}{2}mv^2)\hat{t}. \quad (14.4)$$

Если m — масса частицы, то, согласно определению 4-импульса, его пространственная часть p в данном приближении — это обычный 3-импульс, а временная компонента представляет собой сумму кинетической энергии и массы покоя. При столкновениях с малыми скоростями сохраняются полные 3-импульс, энергия, а также масса. Следовательно, при таких столкновениях будет сохраняться и полный 4-импульс рассматриваемой системы.

Чтобы получить лоренц-инвариантное обобщение динамики столкновений, достаточно потребовать сохранения полного 4-импульса системы при любых столкновениях. Это более слабый закон сохранения, чем в ньютоновой механике, ибо здесь не требуется раздельного сохранения массы и энергии. Наверняка каждый найдет убедительные подтверждения этого.

Пример

При малых скоростях кинетическая энергия и 3-импульс связаны соотношением

$$E_{\text{KINETIC}} = \frac{p^2}{2m}. \quad (14.5)$$

Масса покоя

[Мы обозначили эту постоянную буквой m . В нескольких следующих строчках будет показано, что m следует считать массой частицы. Ее обычно называют *массой покоя* частицы.]

[Здесь три пространственные компоненты 4-вектора записаны в виде обычного 3-вектора. Таким образом, v — это 4-вектор, у которого отсутствует составляющая, параллельная базисному вектору $\hat{t}$, а его $\hat{x}$ -составляющая имеет вид $v\hat{x}$. То же относится к $\hat{y}$ - и $\hat{z}$ -составляющим. Кроме того, принято сокращение $v^2 \equiv v \cdot v$.]

Сохранение 4-импульса

[Надеюсь, вы уже заметили, что основной способ получения информации о 4-векторах — это вычисление всех скалярных произведений, которые попали в ваше поле зрения.]

Найдем соответствующее релятивистское выражение, которое следует из равенств

$$p \cdot p = m^2 \lambda \cdot \lambda = -m^2. \quad (14.6)$$

Развернем эту формулу:

$$p^x p^x + p^y p^y + p^z p^z - p^t p^t = -m^2. \quad (14.7)$$

Будем называть полной энергией E величину

$$E \equiv p^t, \quad (14.8)$$

а 3-импульсом — величину

$$\mathbf{p} \equiv p^x \hat{x} + p^y \hat{y} + p^z \hat{z}, \quad (14.9)$$

$$p^2 \equiv \mathbf{p} \cdot \mathbf{p}. \quad (14.10)$$

Тогда релятивистское обобщение нашего закона имеет вид

$$E^2 = p^2 + m^2. \quad (14.11)$$

откуда в пределе малых скоростей получим

$$E \simeq m + \frac{p^2}{2m}. \quad (14.12)$$

Фотоны

Можно определить понятие 4-импульса даже для частиц, движущихся со скоростью света. Для этого снова будем считать, что 4-импульс параллелен соответствующей мировой линии, т. е.

$$p \propto \sigma, \quad (14.13)$$

откуда, принимая во внимание, что теперь σ' — изотропный вектор, получим

$$E^2 = p^2. \quad (14.14)$$

[Мы, конечно, здесь злоупотребили буквой "p", но что поделаешь — все эти выражения принято писать в такой форме.]

Вот почему говорят, что такие частицы не имеют массы покоя. Нормировка 4-вектора p должна быть выбрана так, чтобы его $\hat{t}$ -компонента была полной энергией, а пространственная часть — полным импульсом, удовлетворяющими закону сохранения полного 4-импульса.

Пространство импульсов

Представьте себе на миг множество всех возможных 4-импульсов. Пространство 4-импульсов — это векторное пространство, сходное с пространством-временем, но такое, что координаты в нем не x и t , а p^x и E . Геометрия соответствующей диаграммы, которую мы будем называть диаграммой импульсов, характеризуется не спецрелятивистской гиперболой

$t^2 - x^2 = 1$, а полученным в предыдущем примере соотношением для энергии-импульса

$$E^2 - p^2 = m^2. \quad (14.15)$$

Наверное, не следует удивляться, что мы здесь вновь встречаемся с такой уже ставшей привычной нам кривой, как гипербола.

Пример

Предположим, что в результате столкновения двух частиц с равными массами образовалась одна новая частица. На рис. 14.1 представлена пространственно-временная диаграмма и диаграмма импульсов рассматриваемого процесса. Обратите внимание, что в момент столкновения все 4-импульсы параллельны мировым линиям интересующих нас частиц. Кроме того, на рисунке изображена гипербола

$$E^2 - p^2 = m^2, \quad (14.16)$$

где m — величина массы каждой из сталкивающихся частиц. Совершенно очевидно, что концы векторов 4-импульсов сталкивающихся частиц должны лежать на этой гиперболе. Внимательное изучение диаграмм показывает, что новая частица удаляется с быстротой порядка 0,12 и обладает массой, которая примерно в 2,18 раз больше массы m . Последнее число получается в результате сравнения длины вектора 4-импульса новой частицы с длиной, определяемой гиперболой (14.16), которая задает в пространстве импульсов стандарт массы подобно тому, как спецрелятивистская гипербола задает в пространстве-времени стандарт времени. В результате можно было бы сказать, что некоторая часть кинетической энергии сталкивающихся частиц перешла в массу. Наблюдаемое здесь несохранение массы — это аналог в пространстве импульсов пространственно-временного парадокса близнецов.

Когда физики изобретают все более изощренные методы описания частиц, совместимые с квантовой механикой, они всегда убеждены, что частицы в создаваемых ими теориях ведут себя в соответствии с требованиями лоренц-инвариантности. Нет никаких факторов, позволяющих усомниться в корректности такой точки зрения, и ни одна теория, не согласующаяся с требованием лоренц-инвариантности, не будет принята всерьез до тех пор, пока такие факты не будут обнаружены.

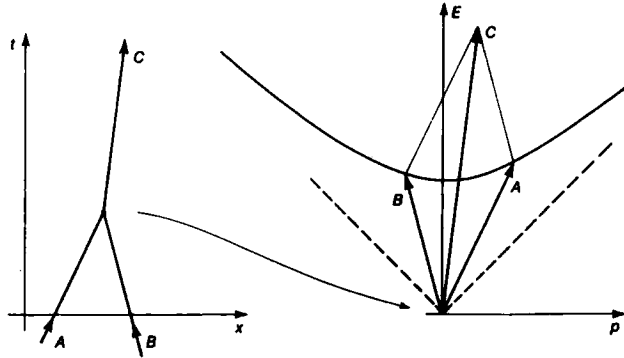


Рис. 14.1

Слева пространственно-временная диаграмма, а справа — диаграмма импульсов для одного и того же процесса столкновения. Заметьте, что выбор начала координат на левой диаграмме совершенно произволен, чего нельзя сказать о диаграмме импульсов. Обратите внимание на параллельность линий, принадлежащих разным диаграммам. Гипербола на диаграмме импульсов описывается уравнением $E^2 - p^2 = m^2$.

ЗАДАЧИ

14.1. (18) Объясните, какие единицы используются в формуле

$$E^2 = p^2 + m^2.$$

14.2. (14) Проведите подробный численный расчет величин, характеризующих диаграммы на рис. 14.1.

14.3. (27) Предположим, что обнаружена частица, которая в состоянии покоя распадается на две новые идентичные частицы, разлетающиеся в противоположных направлениях, каждая с удвоенной скоростью света. Начертите соответствующую пространственно-временную диаграмму. Существует ли способ определения 4-импульса новых частиц, гарантирующий сохранение полного 4-импульса? Если предположить, что рассматриваемый процесс лоренц-инвариантен, то эти частицы можно вернуть в их собственное прошлое. Покажите, как это сделать.

II. Геометрия

При построении специальной теории относительности мы сосредоточили внимание на изучении хроноструктуры события. Эта структура графически изображается в виде семейства гипербол, подобной приведенной на рис. II. 1. В настоящей главе мы сделаем некоторое отступление и изложим математический аппарат, который приспособлен для описания таких структур и является составной частью теории гравитации, которую называют общей теорией относительности.

Можно провести следующую аналогию. Рассмотрим в качестве примера поток жидкости. Вектор скорости жидкости, вообще говоря, является сложной функцией пространственных координат и времени. Однако для малых интервалов времени и в малых областях пространства скорость оказывается примерно постоянной. Это схематически изображено на рис. II. 2. Точно так же пространство-время общей теории относительности обладает хроноструктурой, которая сложным образом зависит от координат события. Но в достаточно малых областях пространство-времени хроноструктура однородна, и пространство-время локально совпадает с пространством-временем специальной теории относительности—пространством Минковского. Эта глава посвящена изучению аппарата, который позволяет понять ситуации, аналогичные изображенной на рис. II. 3. Сделаем еще одно замечание. В то время как локальным приближением для векторного поля служит постоянное векторное поле (поле свободных векторов), имеющее сравнительно простую структуру, локальным приближением пространства-времени общей теории относительности служит пространство Минковского, которое, как мы видели в предыдущей главе, само по себе обладает довольно сложной геометрией.

Математические объекты, которые мы изображаем гиперболами, называются тензорами. В настоящей главе рассматриваются свойства тензоров и тензорных полей. Без знания тен-

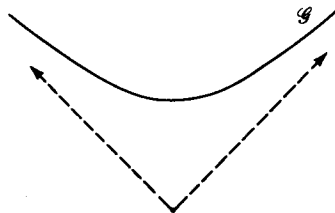


Рис. II.1

Специальная теория относительности как локальная теория

[В разд. 51 мы увидим, что рис. II.3 соответствует черной дыре.]

Тензоры

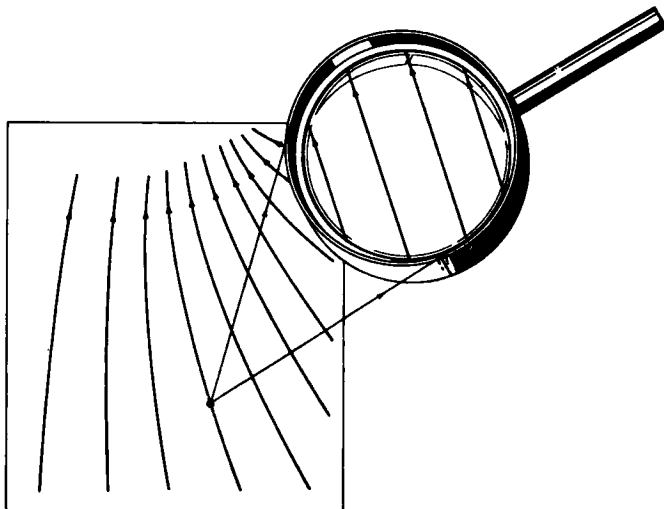


Рис. II.2
Поле скоростей при сильном увеличении.

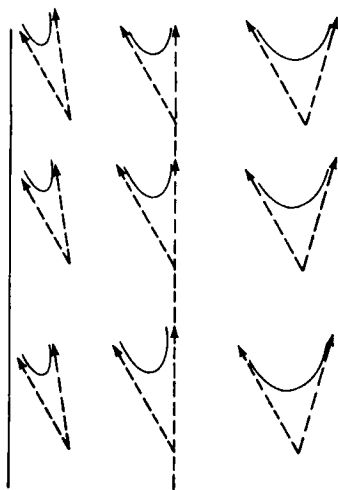


Рис. II.3
Метрические фигуры для девяти событий в пространстве-времени. Световые сигналы отклоняются, поэтому соответствующая система отсчета не является инерциальной; более того, такой системы отсчета не существует.

зорного анализа нельзя освоить теорию гравитации, точно так же как без умения обращаться с векторами невозможно изучать электромагнетизм. Мы будем придерживаться наиболее простого и эффективного способа изложения тензорного анализа, при котором по возможности используются ковариантные, т. е. не зависящие от выбора системы координат, обозначения. С типичным примером основанного на таком подходе расчета мы встречались в разд. 13, когда проводили вычисление доплеровского сдвига. Чтобы подчеркнуть, что рассматриваемые объекты существуют независимо от выбора системы координат, мы намеренно не изображали координатные оси на рис. II. 1. и II.2.

В настоящей главе мы будем изучать поведение тензорных полей в непосредственной близости от некоторой выбранной точки в столь малой окрестности, что всюду в ней поле можно считать постоянным. Соответствующая теория называется тензорной алгеброй. В гл. III мы перейдем к изучению переменных тензорных полей, что составляет предмет тензорного анализа. Применение аппарата тензорной алгебры позволит нам изложить специальную теорию относительности как геометрию некоторого простейшего тензорного поля.

15. Векторы и ковекторы

Тензорная алгебра занимается изучением векторных пространств и связанных с ними линейных операторов. Тензорную алгебру можно построить на любом векторном пространстве. Для нас наибольший интерес представляет алгебра на пространстве векторов, касательных к мировым линиям в пространстве-времени.

Для любого векторного пространства V самыми простыми и наиболее важными являются операторы, действие которых на векторы из V дает действительные числа. В разд. 16 мы увидим, что если векторы V представляют собой векторы скорости, то соответствующие операторы будут определять градиенты функций. Пусть ω — оператор (или отображение; мы используем термин «оператор», так как это отображение специального вида, с полным основанием его можно было бы называть также функцией), такой, что

$$\omega: V \rightarrow \mathbb{R}; a \mapsto \omega \cdot a. \quad (15.1)$$

Оператор ω называется линейным, если для любых векторов a и b и для любого числа k справедливо соотношение

$$\omega \cdot (ka + b) = k(\omega \cdot a) + \omega \cdot b. \quad (15.2)$$

Обратите внимание на то, что в левой части равенства мы складываем векторы, в то время как в правой части складываются числа. Такие линейные операторы будут называться *ковекторами*; как мы увидим, они также образуют линейное пространство, которое обычно обозначается символом V^* . Используемое нами условное обозначение для операции действия ковектора на вектор имеет иной смысл, чем соответствующее обозначение для скалярного произведения евклидовой геометрии и геометрии пространства Минковского. Однако это не должно привести к путанице, поскольку последние в оставшейся части книги будут использоваться крайне редко. Такое обозначение мы ввели, поскольку связь между векторами и ковекторами симметрична. Это значит, что мы можем рассматривать вектор a как оператор, действующий в пространстве ковекторов.

Введенные объекты, векторы и ковекторы, допускают знакомое физикам представление в виде векторов-столбцов и векторов-строк, используемых в матричной алгебре. Существует также графическое представление этих объектов, которое

Ковекторы

[Условные обозначения, используемые для отображений, были пояснены на стр. 36.]

Графическое представление

весьма удобно и полезно, так как позволяет рассматривать эти объекты вне связи с той или иной системой координат. Напомним, что представление структуры события $\mathcal{S}$ мы построили непосредственно из элементов самого пространства V . Поступим аналогичным образом.

Рассмотрим множество всех векторов v , принадлежащих пространству V , таких, что

$$\omega \cdot v = 1. \quad (15.3)$$

Обозначим это множество символом « ω »; оно и будет представлением ковектора ω . В последующих разделах мы будем опускать кавычки и обозначать множество тем же символом, что и сам ковектор.

В двумерном векторном пространстве множество « ω » представляет собой прямую линию. Чтобы убедиться в этом, возьмем два различных вектора a и b , таких, что

$$\omega \cdot a = 1, \quad (15.4)$$

$$\omega \cdot b = 1. \quad (15.5)$$

Как a , так и b принадлежат множеству « ω ». Из свойства линейности (15.2) следует, что

$$\omega \cdot (a - b) = 0; \quad (15.6)$$

поэтому для любого k

$$\omega \cdot (k(a - b) + a) = 1. \quad (15.7)$$

Таким образом, если векторы a и b принадлежат множеству « ω », то ему принадлежит и вся определяемая ими прямая линия (рис. 15.1).

В двумерном случае « ω » не содержит других векторов. Действительно, предположим, что есть вектор, который удовлетворяет условию (15.3), но не принадлежит указанной прямой. В этом случае существовала бы вторая прямая, которая пересекала бы первую и также принадлежала множеству « ω ». Но тогда множество « ω » совпадало бы со всем двумерным пространством. С другой стороны, из (15.2) следует, что

$$\omega \cdot 0 = 0; \quad (15.8)$$

полученное противоречие показывает, что исходное предположение неверно. В пространстве n измерений ковектор может быть представлен $(n - 1)$ -мерным линейным подпространством. Например, в случае трех измерений « ω » будет плоскостью

Случай двух измерений

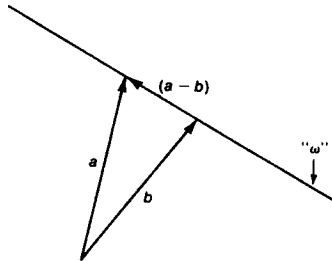


Рис. 15.1

и т. д. Если не существует векторов, удовлетворяющих условию (15.3), то для любого вектора v должно выполняться равенство

$$\omega \cdot v = 0. \quad (15.9)$$

Такой оператор мы будем называть нулевым оператором.

Мы могли бы нарисовать полный набор линий уровня для нашего оператора. Линии уровня определяются из условия

$$\omega \cdot v = \dots -1, 0, 1, 2, \dots \quad (15.10)$$

Нетрудно показать, что это набор прямых, которые параллельны « ω » и расположены на равных расстояниях друг от друга. Линия, соответствующая нулевому значению оператора, проходит через нулевой вектор.

По аналогии с понятием свободного вектора можно ввести понятие свободного ковектора. Закрепленный вектор полностью задается одной точкой; в этом случае множество « ω » состоит из таких точек. Чтобы задать свободный вектор, необходимы две точки и способ различить начало и конец; обычно для этих целей используется стрелка. Поэтому для свободного ковектора мы должны нарисовать две линии уровня и указать положительное направление. Принятый нами способ обозначения показан на рис. 15.2 и 15.3 для векторов и ковекторов соответственно в пространстве двух и трех измерений.

Полученное представление позволяет легко вычислять результат действия ковектора на вектор. Нужно только подсчитать число (учитывая нецелую часть) линий уровня семейства « ω », которые пересекаются вектором. При этом знак зависит от взаимной ориентации вектора и ковектора. Иногда такую операцию называют «свертыванием».

Нулевой оператор

[Поскольку мы рассматриваем случай двух измерений, я употребляю слово "линии", в то время как следовало бы говорить "($n - 1$)-мерные линейные подпространства".]

Линии уровня

[О свободных векторах см. разд. 1; см. также рис. 16.5.]

Свободные ковекторы



Рис. 15.2

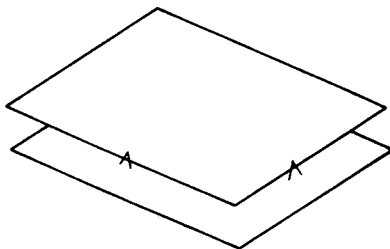


Рис. 15.3

Следуя обычному соглашению, мы будем использовать первые буквы латинского алфавита для обозначения векторов, а последние буквы греческого алфавита — для обозначения ковекторов; единственное исключение — использование буквы v для обозначения вектора.]

Линейность пространства ковекторов

Умножение на число

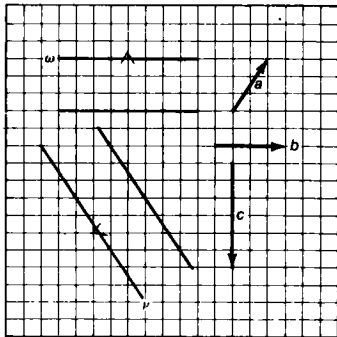


Рис. 15.4

Пример

На рис. 15.4, мы изобразили два свободных ковектора и три свободных вектора. Нетрудно видеть, что результаты действия ковекторов на векторы таковы:

$$\begin{aligned}\omega \cdot a &= +1, & \nu \cdot a &= -1, \\ \omega \cdot b &= 0, & \nu \cdot b &= -1, \\ \omega \cdot c &= -2, & \nu \cdot c &= +1.\end{aligned}$$

Множество ковекторов само является линейным пространством. Назовем ковектором $k\omega$ оператор, действие которого определено соотношением

$$(k\omega) \cdot v = k(\omega \cdot v). \quad (15.11)$$

Аналогично сумма двух ковекторов ω и ν представляет собой такой ковектор $(\omega + \nu)$, что

$$(\omega + \nu) \cdot v = \omega \cdot v + \nu \cdot v. \quad (15.12)$$

Принятые определения делают ненужным употребление скобок в выражениях типа (15.11) или (15.12). Возникает вопрос: что представляет собой множество « $k\omega$ »? Предположим, что вектор a принадлежит « ω », т. е.

$$\omega \cdot a = 1. \quad (15.13)$$

Тогда, используя свойство линейности (15.2), получаем

$$(k\omega) \cdot \left(\frac{a}{k}\right) = 1. \quad (15.14)$$

Таким образом, линии уровня « $k\omega$ » параллельны линиям « ω », но расположены в k раз ближе друг к другу. Для случая $k = 2$ это изображено на рис. 15.5. Обратите внимание на то, что поведение векторов и ковекторов взаимно дополняет друг друга. Это необходимо, поскольку представление векторного пространства не единственно. Как и в случае инерциальных систем отсчета, мы можем осуществить произвольное линейное преобразование. В частности, наше представление должно быть ковариантным относительно растяжений, и, как мы видим, такая ковариантность действительно имеет место. Такие преобразования соответствуют растяжению вектора, но пропорциональному сжатию ковектора; при этом величина $\omega \cdot v$ остается неизменной (рис. 15.6).

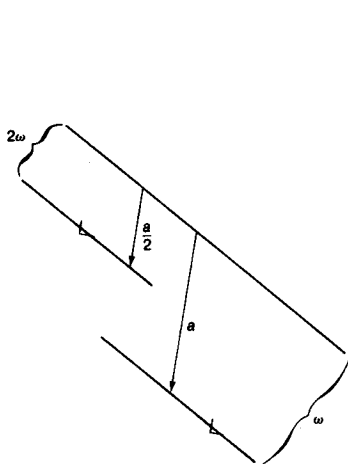


Рис. 15.5

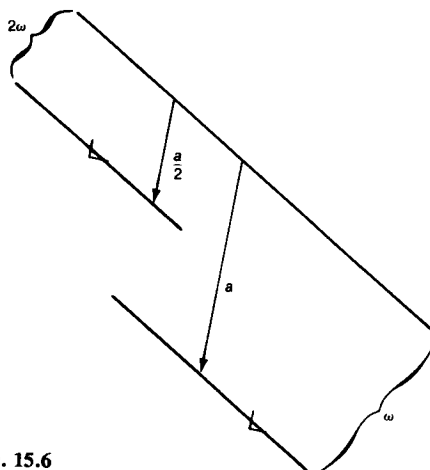


Рис. 15.6

Рис. 15.5 после выполнения операции растяжения.

Что можно сказать относительно операции сложения двух ковекторов? Как построить сумму « $\omega + \nu$ », если заданы два произвольных ковектора « ω » и « ν »? Снова воспользуемся свойством линейности. Выберем два вектора a и b , такие, что

$$\begin{aligned}\omega \cdot a &= 1, \\ \nu \cdot a &= 0, \\ \omega \cdot b &= 0, \\ \nu \cdot b &= 1.\end{aligned}\tag{15.15}$$

Обобщение предлагаемой процедуры на случай большего числа измерений достаточно очевидно, но его гораздо труднее изобразить на рисунке. В рассматриваемом случае двух измерений имеем

$$(\omega + \nu) \cdot a = 1,\tag{15.16}$$

$$(\omega + \nu) \cdot b = 1.\tag{15.17}$$

Таким образом, оба вектора a и b должны лежать на линии « $\omega + \nu$ ». Как показано на рис. 15.7, указанное требование полностью определяет « $\omega + \nu$ ».

Наше определение операции сложения ковекторов опиралось исключительно на понятия параллельности и пересечения прямых. Поэтому оно, очевидно, ковариантно относительно

Сложение

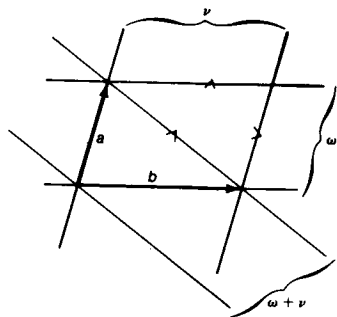


Рис. 15.7

Сложение ковекторов ω и ν , в результате которого получается новый ковектор $(\omega + \nu)$; при выполнении операции сложения вводятся два вспомогательных вектора a и b .

Ковариантность относительно линейных преобразований

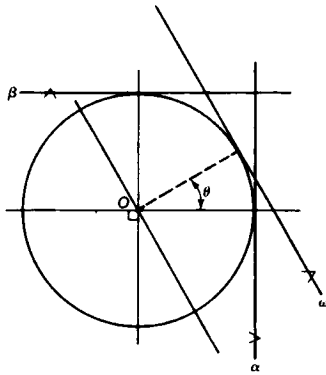


Рис. 15.8

произвольных линейных преобразований. Эта ковариантность будет играть чрезвычайно важную роль, когда мы перейдем к рассмотрению векторных и тензорных полей на многообразиях.

В пространстве двух измерений ковектор может быть представлен точкой на единичном контуре, которая ближе всего расположена к началу координат. Однако такое представление не ковариантно относительно линейных преобразований. Оно было использовано в разд. 1 при получении необычного правила сложения векторов в $\mathbb{R}^2$.

Мы видим, что векторы и ковекторы очень похожи. Соотношение между ними называется *дуальностью*. Мы говорим, что ковекторы дуальны векторам. Ковекторы являются линейными операторами в пространстве векторов; подобно этому векторы, очевидно, могут рассматриваться как линейные операторы в пространстве ковекторов. Соответствующее отображение осуществляется согласно правилу

$$a: V^* \rightarrow \mathbb{R}; \omega \mapsto \omega \cdot a.$$

Дуальность Таким образом, векторы также дуальны ковекторам. Действительно, любой линейный оператор, отображающий ковекторы в числа, является некоторым вектором. Справедливость этого утверждения не очевидна, однако, его нетрудно доказать. (Приведенное утверждение не справедливо для некоторых бесконечномерных функциональных пространств.)

ЗАДАЧИ

15.1. (18) На рис. 15.8 изображены окружность и три касательных к ней ковектора. Покажите, что

$$\omega = (\cos \theta)\alpha + (\sin \theta)\beta.$$

15.2. (13) Графическое правило сложения ковекторов неприменимо, когда ковекторы параллельны. Покажите, что ковектор α можно графически сложить с самим собой, если ввести произвольный ковектор β , который не параллелен α , и затем выполнить операцию сложения по правилу

$$\alpha + \alpha = [(\alpha + \beta) + \alpha] - \beta.$$

15.3. (16) Пусть V — пространство непрерывных и ограниченных на отрезке $0 \leq x \leq 1$ функций.

1. Убедитесь, что V — векторное пространство.

2. Покажите, что оператор

$$I_g: V \rightarrow \mathbb{R}; f(x) \rightarrow \int_0^1 f(x) g(x) dx$$

принадлежит V^* .

3. Найдите элемент пространства V^* , который не может быть представлен в таком виде.

15.4. (22) Пусть для некоторого неизвестного ковектора ω выполняется соотношение

$$\omega \cdot x = 0$$

для всех векторов x , таких, что

$$\nu \cdot x = 0,$$

где ν — некоторый заданный ковектор. Покажите графически, что в этом случае

$$\omega = \lambda \nu,$$

где λ — произвольный числовой множитель, который называется множителем Лагранжа.

15.5. (21) Решите задачу 15.4 для случая, когда вектор x удовлетворяет сразу нескольким условиям

$$\alpha \cdot x = 0,$$

$$\beta \cdot x = 0,$$

и т. д.

15.6. (16) Решите задачи 15.4 и 15.5 для дуального случая, когда неизвестный вектор x удовлетворяет условию

$$\omega \cdot x = 0$$

для всех ω , таких, что

$$\omega \cdot a = 0,$$

$$\omega \cdot b = 0$$

и т. д.

15.7. (26) В пространстве двух измерений ковектор может быть задан двумя числами l и m , которые представляют собой длины отрезков, отсекаемых на координатных осях прямоугольной системы координат (рис. 15.9).

а. Что изменится на рисунке, если ковектор умножить на число k ?

б. Покажите, что при таком представлении закон сложения ко-

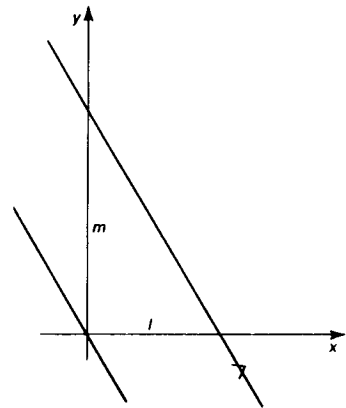


Рис. 15.9

векторов имеет вид

$$\frac{1}{l_3} = \frac{1}{l_1} + \frac{1}{l_2},$$

$$\frac{1}{m_3} = \frac{1}{m_1} + \frac{1}{m_2}.$$

(Указание: рассмотрите подобные треугольники.)

в. Что будет, если оси координат не ортогональны?

15.8. (24) Снова рассмотрите задачу 1.11. (Обратите внимание что оценка теперь другая!)

16. Касательные векторы и 1-формы

*Верх безумия учиться тому,
что потом должно быть забыто.*

Эразм

В предыдущем разделе мы ввели ковекторы как наиболее простые операторы, заданные на векторном пространстве, и показали, что имеются чисто геометрические основания для введения понятия ковектора. В настоящем разделе рассматривается одна из реализаций этого понятия. Конкретные физические приложения обсуждаются в двух последующих разделах. Следует отметить, что все стандартные курсы математической физики обладают общим недостатком — в них не делается различия между векторами и ковекторами. Трудно привыкнуть различать концепции, которые ранее были введены как одно понятие. Поэтому многие сторонники тензорного анализа не понимают полезности такого различия. Тем не менее физический мир полон ковекторами! Я надеюсь, что рассмотренные ниже примеры покажут, до какой степени понятие ковектора облегчает наше понимание окружающего мира.

Касательные векторы

Векторное пространство, в котором мы работаем, представляет собой пространство касательных векторов к всевозможным параметризованным кривым, проходящим через некоторую фиксированную точку P . Касательный вектор содержит в себе информацию о локальном поведении гладкой кривой, проходящей через точку P . Это отличает его от свободного вектора, который характеризует локальное поведение семейства гладких кривых в окрестности выделенной точки.

Параметризованные кривые

Напомним определение касательного вектора к параметризованной кривой, данное на стр. 102. Это определение было основано на операции предельного перехода, которая включала в себя растяжение окрестности точки P в некоторое число раз с

последующим устремлением степени растяжения к бесконечности. Переход к пределу позволяет получить линейную аппроксимацию параметризованных кривых в окрестности точки P .

Пример

На плоскости $\mathbb{R}^2$ рассмотрим изображенное на рис. 16.1. семейство параметризованных кривых

$$\gamma_a: \mathbb{R} \rightarrow \mathbb{R}^2; u \mapsto a(\cos \frac{u}{a^2}, \sin \frac{u}{a^2}). \quad (16.1)$$

Здесь различные значения параметра a соответствуют разным кривым, а различные значения u относятся к различным точкам на одной и той же кривой. Кривые представляют собой концентрические окружности с центром в начале координат; при этом смещение точки вдоль окружности, как функция u , уменьшается с увеличением радиуса. Используя описанную выше процедуру предельного перехода, рассмотрим поведение таких семейств кривых в окрестности точки $P = (1, 0)$. Будем увеличивать масштаб наших чертежей, используя для этой цели растянутые координаты

$$X \equiv \frac{x-1}{\varepsilon}. \quad (16.2)$$

$$Y \equiv \frac{y}{\varepsilon},$$

[Мы вычертили эти кривые на фоне координатной сетки, чтобы подчеркнуть, что наши чертежи верны в количественном отношении. Пожалуйста, обратите внимание на то, что мы могли бы выбрать любую другую линейную координатную сетку. Совершенно обязательно было выбрать прямоугольную систему координат. Это проявление хорошо знакомой нам ковариантности относительно линейных преобразований.]

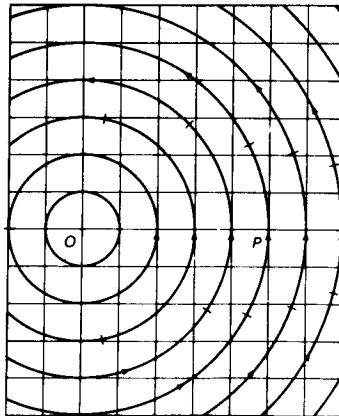


Рис. 16.1

Описанное в тексте семейство параметризованных кривых. Стрелки стоят в точках, соответствующих целым значениям параметра u , черточки — в точках, соответствующих полuceлому значению этого параметра.

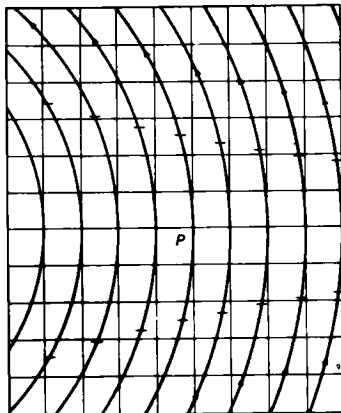


Рис. 16.2

Рис. 16.1 после двукратного увеличения ($\varepsilon = 1/2$). Число кривых увеличено в два раза; отмечены точки, соответствующие значениям параметра u , равным $1/2$ и $1/4$.

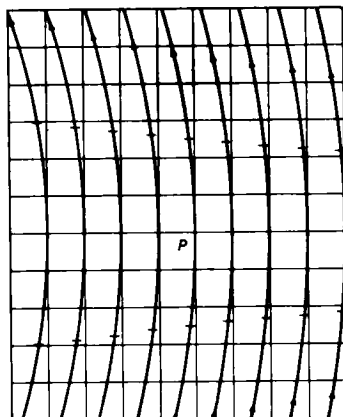


Рис. 16.3

Рис. 16.1 после пятикратного увеличения ($\varepsilon = 1/5$). Отмечены точки, для которых значение u равно $0,2$ и $0,1$.

[Когда мы говорим, что прямая является локальной линейной аппроксимацией некоторой кривой в точке P , то имеем в виду, что расхождение между кривой и аппроксимирующей прямой возрастает как квадрат расстояния от точки P .]

В результате такой замены переменных область, имевшая размер ε в старых координатах, теперь имеет размер порядка единицы. Так как мы рассматриваем только малую окрестность точки P , изменим масштаб параметра u и введем новый параметр

$$U \equiv \frac{u}{\varepsilon}. \quad (16.3)$$

Чтобы выделить кривые γ_a , проходящие вблизи точки $(1, 0)$, введем еще один новый параметр

$$A \equiv \frac{a-1}{\varepsilon}. \quad (16.4)$$

Заметим, что через точку $(1, 0)$ проходит только окружность γ_1 . В новых координатах наши кривые задаются отображением Γ_A :

$$\begin{aligned} \Gamma_A: U &\mapsto (X, Y); \\ X &= \frac{(1 + \varepsilon A)}{\varepsilon} \left[\cos \frac{\varepsilon U}{(1 + \varepsilon A)^2} \right] - \frac{1}{\varepsilon}; \\ Y &= \frac{(1 + \varepsilon A)}{\varepsilon} \sin \frac{\varepsilon U}{(1 + \varepsilon A)^2}. \end{aligned} \quad (16.5)$$

Эти кривые для $\varepsilon = 1/2$ и $\varepsilon = 1/5$ приведены на рис. 16.2 и 16.3. Поскольку рассматриваемое семейство состоит из гладких кривых, при достаточно малых ε их вид в (XY) -плоскости практически не будет меняться с уменьшением ε . Используя разложение в ряд Тейлора, можно перейти к пределу $\varepsilon \rightarrow 0$. Тогда получим

$$\begin{aligned} X &= A, \\ Y &= U. \end{aligned} \quad (16.6)$$

Формулы (16.6) описывают изображенное на рис. 16.4 семейство прямых, параллельных оси y .

После того как мы нашли линейную аппроксимацию рассматриваемого семейства кривых, можно вернуться к исходным переменным. Имеем

$$\begin{aligned} x &= a, \\ y &= u. \end{aligned} \quad (16.7)$$

Формулы (16.7) описывают семейство прямых линий, аппроксимирующих заданное семейство окружностей в окрестности точки $P = (1, 0)$.

Соответствующее поле свободных векторов изображено на рис. 16.5.

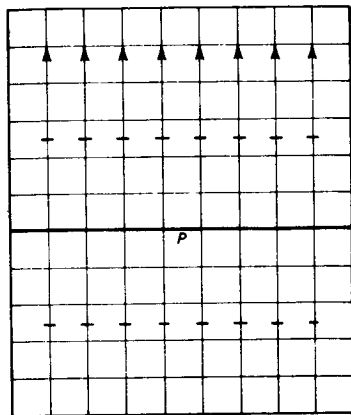


Рис. 16.4

Предел, к которому приводит описанная выше процедура при $\varepsilon \rightarrow 0$.

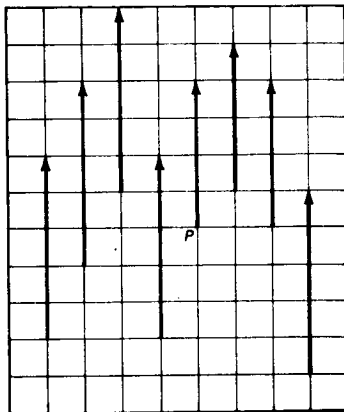


Рис. 16.5

Поле свободных векторов, соответствующее семейству параллельных прямых, которое изображено на рис. 16.4. Обратите внимание, что свободный вектор можно провести из любой точки, точно так же, как прямую на рис. 16.4 можно провести через любую точку.

Существует другое понятие, не менее важное, чем понятие кривой. Мы имеем в виду *функцию*. Функция ставит в соответствие каждому событию пространства-времени некоторое число. Таким образом, она представляет собой отображение пространства-времени в $\mathbb{R}$. Параметризованная кривая является дуальным понятием по отношению к функции: это отображение, которое ставит в соответствие каждой точке из $\mathbb{R}$ некоторое событие в пространстве-времени. Как мы знаем, поведение гладкой кривой в окрестности некоторой точки определяется касательным вектором. Точно так же поведение гладкой функции определяется ее градиентом. Чтобы найти локальную линейную аппроксимацию функции в окрестности некоторой точ-

Функции

Градиент

ки, можно воспользоваться описанной выше процедурой предельного перехода. Представим нашу функцию в виде набора линий уровня, затем растянем окрестность рассматриваемой точки в $1/\varepsilon$ раз и дорисуем новые линии уровня таким образом, чтобы переход с одной линии на соседнюю соответствовал изменению значения функции не на единицу, как до растяжения, а на величину ε . Если функция является гладкой, то этот процесс приводит к хорошо определенному пределу. Таким пределом будет семейство параллельных и эквидистантных линий. В градиенте нетрудно узнать нашего старого знакомого — свободный ковектор.

Градиент является представителем класса операторов, которые называются дифференциальными формами. Как первый член этого класса его называют также 1-формой. Термин «градиент» часто используется для обозначения вектора, поэтому во избежание недоразумений мы будем называть градиент 1-формой. Другие члены этого ряда, 2-формы и т. д., нам не требуются.

1-формы

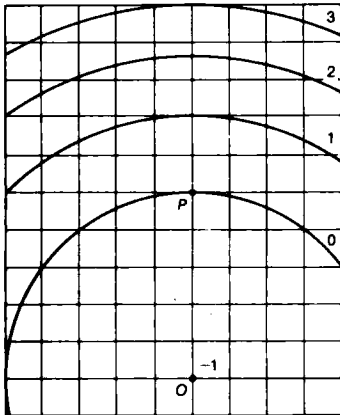


Рис. 16.6

Линии уровня функции $x^2 + y^2 - 1$. Точка P имеет координаты $(0,1)$.

Пример

Пусть на $\mathbb{R}^2$ задана функция

$$f(x,y) = x^2 + y^2 - 1; \quad (16.8)$$

найдем ее линейную аппроксимацию в окрестности точки $P = (0,1)$. Введем растянутые координаты

$$X \equiv \frac{x}{\varepsilon}, \quad (16.9)$$

$$Y \equiv \frac{y-1}{\varepsilon}$$

и рассмотрим новую функцию (точнее, исходную функцию в новых переменных)

$$F(X,Y) = (1 + \varepsilon Y)^2 + \varepsilon^2 X^2 - 1. \quad (16.10)$$

Линии уровня этой функции определяются из соотношения

$$(1 + \varepsilon Y)^2 + \varepsilon^2 X^2 - 1 = \varepsilon n, \quad (16.11)$$

где $n = 0, 1, 2, \dots$. Для $\varepsilon = 1$ и $\varepsilon = 1/4$ линии уровня изображены соответственно на рис. 16.6 и 16.7. В пределе $\varepsilon \rightarrow 0$ получим

$$2Y = n. \quad (16.12)$$

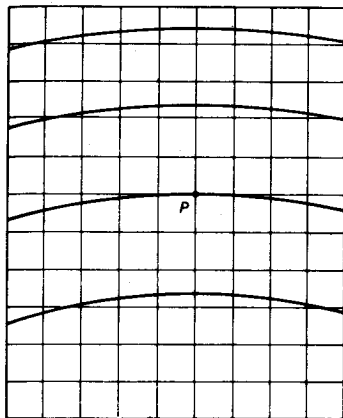


Рис. 16.7

Рис. 16.6 после увеличения в четыре раза.

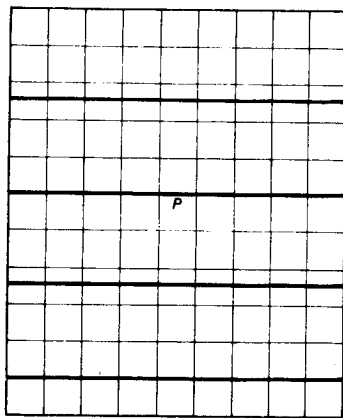


Рис. 16.8

Предел при бесконечном увеличении степени растяжения.

Эти параллельные и расположенные на равных расстояниях прямые линии показаны на рис. 16.8. Соответствующее поле свободных ковекторов изображено на рис. 16.9.

Теперь мы можем обратить процедуру предельного перехода. Возвращаясь к исходным координатам, получаем, что прямые

$$f \sim 2y - 2 \quad (16.13)$$

являются линейными функциями, аппроксимирующими функцию

$$f = x^2 + y^2 - 1 \quad (16.14)$$

в окрестности точки $(0, 1)$.

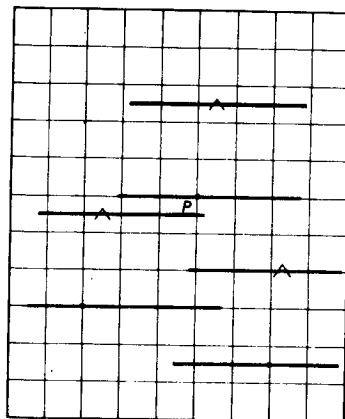


Рис. 16.9

Соответствующее рис. 16.8 поле свободных ковекторов.

Множество заданных в некоторой точке 1-форм представляет собой множество линейных функций, локально аппроксимирующих всевозможные функции более общего вида. Две такие линейные функции можно сложить, и любую линейную функцию можно умножить на число. Указанные операции превращают множество 1-форм в линейное векторное пространство. Аналогично касательные векторы в некоторой точке дают линейную аппроксимацию проходящих через эту точку параметризованных кривых. Касательные векторы можно умножать на число и складывать; они также образуют линейное векторное пространство.

Линейная структура пространства 1-форм

Линейная структура пространства касательных векторов

Пример

Линейной аппроксимацией кривой

$$u \mapsto (x, y) = (u, \sin u) \quad (16.15)$$

в точке $(0, 0)$ является прямая линия

$$u \mapsto (x, y) = (u, u). \quad (16.16)$$

Отображение

$$u \mapsto (2u, -u) \quad (16.17)$$

определяет другую прямую. Сумма этих двух прямых также представляет собой прямую линию

$$u \mapsto (3u, 0). \quad (16.18)$$

Более того, умножение второй прямой на число k

$$u \mapsto (2ku, -ku), \quad (16.19)$$

дает другую прямую линию.

Множество всех касательных векторов в некоторой точке называется касательным векторным пространством в этой точке. Множество всех заданных в точке 1-форм называется пространством, дуальным касательному пространству.

Линейные операторы

Ковекторы и 1-формы представляют собой линейные операторы, в результате действия которых на векторы получаются действительные числа. Какова геометрическая интерпретация этой операции? Предположим, что нам задана кривая γ :

$$\gamma: \mathbb{R} \rightarrow M; u \mapsto \gamma(u), \quad (16.20)$$

где мы по-прежнему под пространством-временем M понимаем пространство Минковского.

Предположим далее, что мы имеем функцию f :

$$f: M \rightarrow \mathbb{R}; p \mapsto f(p). \quad (16.21)$$

Если мы вычисляем значение функции f в точках, принадлежащих кривой γ , то мы получаем отображение, которое обозначается $f \circ \gamma$ и называется «композицией f и γ ». Это отображение имеет вид

$$(f \circ \gamma): \mathbb{R} \rightarrow \mathbb{R}; u \mapsto f[\gamma(u)] \quad (16.22)$$

(рис. 16.10). Отображение $f \circ \gamma$ представляет собой функцию в обычном смысле слова. В разд. 17 мы увидим, что вычисление результата действия ковектора градиента функции f на касательный вектор

[Ни одно из этих понятий не требует пересмотра, когда мы перейдем к обсуждению многообразий, поэтому уже сейчас вы можете под M понимать многообразие.]

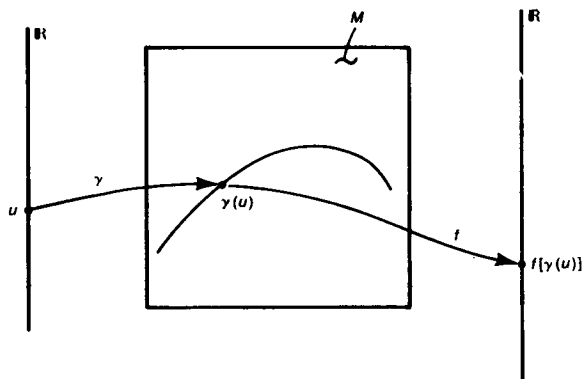


Рис. 16.10

касательный вектор к кривой γ сводится к вычислению производной функции $f \circ \gamma$, т.е. скорости изменения функции f вдоль кривой γ при изменении параметра u . Таким образом, понятие касательного вектора объединяет в себе две различные идеи. С одной стороны, это локальная линейная аппроксимация кривой или семейства кривых. С другой стороны — это оператор производной функции по направлению. В пространствах конечного числа измерений оба понятия представляются одним и тем же математическим объектом. То же самое справедливо и для большинства бесконечномерных пространств. Вектор v , понимаемый как производная по направлению, в обычных трехмерных обозначениях записывается в виде $\mathbf{v} \cdot \nabla$. Дальнейшее обсуждение оператора производной по направлению отложим до разд. 17.

Производная по направлению

В книгах по тензорному анализу касательные векторы и 1-формы называются соответственно контравариантными и ковариантными векторами. Однако мы надеемся, что рассмотрение обобщенной силы (1-форма) в разд. 18, а также групповой скорости (касательный вектор) и фазовой скорости (1-форма) в разд. 19 убедит читателя в том, что геометрическая картина, в которой проводится различие между векторами и 1-формами, вполне естественна, проста и интуитивно понятна.

ЗАДАЧИ

16.1. (11) Покажите, что 1-форма градиента функции

$$f(x, y) = x^2 + y^2$$

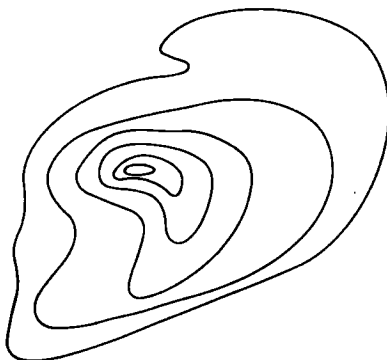


Рис. 16.11

обращается в нуль в начале координат. Сделайте это графически и аналитически.

16.2. (10) На рис. 16.11 изображен участок топографической карты. Для нескольких произвольно выбранных точек изображите 1-форму градиента.

16.3. (10) Пусть кривая γ представляет собой винтовую линию:

$$\gamma: s \mapsto (\cos s, \sin s, s),$$

а функция f имеет вид

$$f(x, y, z) = x^2 + y^2 + z^2.$$

Найдите функцию

$$(f \circ \gamma): \mathbb{R} \rightarrow \mathbb{R}.$$

17. Векторы координатного базиса

Хотя исключение координат из наших определений представляет собой простой и элегантный прием и хотя графические построения наглядны и интуитивно понятны, для эффективного использования тензоров в действительности необходимо производить аналитические вычисления. Чтобы перейти к числовому представлению, воспользуемся нашей системой координат и зададим специальный набор базисных векторов. Мы уже поступали так выше при рассмотрении канонических систем отсчета. Подчеркнем, что, хотя мы по-прежнему считаем пространство линейным векторным пространством, введенные здесь определения после небольших изменений будут соответство-

вать гораздо более общему случаю многообразия. Характерная особенность многообразий состоит в том, что для них в отличие от линейных векторных пространств необходимо использовать криволинейные системы координат.

В разд. 16 мы определили градиент функции как ее локальную линейную аппроксимацию. Поскольку сумма двух линейных функций представляет собой линейную функцию и поскольку в результате умножения линейной функции на константу мы снова получаем линейную функцию, множество линейных функций образует векторное пространство. Эффективный способ задания векторного пространства состоит в выборе базиса, т. е. некоторого стандартного набора векторов, и записи всех векторов линейного пространства в виде линейной комбинации базисных векторов. Теперь координаты сами становятся функциями, определенными на пространстве-времени, и мы можем вычислить их градиенты. Мы используем эти градиенты как базис в пространстве 1-форм.

Вернемся теперь к рассмотрению градиентов и для произвольной функции $f(x, y)$ повторим вычисления, проведенные в примере на стр. 129. Мы можем разложить функцию $f(x, y)$ в окрестности точки (x_0, y_0) в ряд Тейлора

$$f(x, y) \approx f(x_0, y_0) + \frac{\partial f}{\partial x}(x - x_0) + \frac{\partial f}{\partial y}(y - y_0) + \dots \quad (17.1)$$

Линейные члены в этом выражении мы называем градиентом функции f . Следуя общепринятой традиции, обозначим ко вектор градиента функции f символом df . Теперь это не бесконечно малая величина, а определенная выше 1-форма.

Что можно сказать о самой координате x ? Она тоже представляет собой функцию, и ее можно записать в виде

$$x = x_0 + (x - x_0). \quad (17.2)$$

В круглых скобках стоит градиент x ; следуя нашему соглашению, обозначим его dx . Линейные функции dx , dy и т. д. образуют базис в пространстве градиентов. Действительно, из приведенного выше разложения функции f в ряд Тейлора следует, что

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy. \quad (17.3)$$

Отсюда видно, в частности, почему мы обозначили 1-форму градиента символом df .

Базисные 1-формы

[Сначала мы рассматриваем пространство ко векторов, поскольку этот случай проще.]

Градиент

df

[Повсюду в этой книге символ d мы будем понимать именно в указанном смысле. Используемые физиками бесконечно малые приращения dx будут обозначаться символом Δx .]

dx

Базисные касательные векторы

Выбор системы координат эквивалентен заданию некоторого набора функций, или, что то же самое, некоторого множества кривых. Это линии, на которых все координаты, кроме одной, фиксированы. Прямая, локально аппроксимирующая заданную кривую линию, называется касательной. Касательные к координатным линиям будем обозначать символами $\partial/\partial x$, $\partial/\partial y$ и т.д.¹⁾ Мы скоро увидим, что такие своеобразные обозначения действительно имеют смысл. В точке (x_0, y_0) символ $\partial/\partial x$ соответствует прямой

$$u \mapsto (x_0 + u, y_0), \quad (17.4)$$

а $\partial/\partial y$ изображается прямой

$$u \mapsto (x_0, y_0 + u). \quad (17.5)$$

Кривая общего вида γ ,

$$\gamma: u \mapsto [X(u), Y(u)], \quad (17.6)$$

может быть локально аппроксимирована прямой линией, явный вид которой мы получим, если воспользуемся разложением в ряд Тейлора

$$u \mapsto (x_0 + \frac{dX}{du}u, y_0 + \frac{dY}{du}u) \quad (17.7)$$

[не нарушая общности, мы приняли, что $0 \mapsto (x_0, y_0)$]. На множестве таких прямых мы снова можем легко определить операции сложения и умножения на число. Тогда касательный вектор (прямую) к произвольной кривой γ , который мы будем обозначать $\dot{\gamma}$, можно записать в виде

$$\dot{\gamma} = \frac{dX}{du} \frac{\partial}{\partial x} + \frac{dY}{du} \frac{\partial}{\partial y}. \quad (17.8)$$

Мы будем использовать векторы $\partial/\partial x$, $\partial/\partial y$ и т. д. в качестве базиса в пространстве касательных векторов. Все сказанное выше нетрудно распространить на случай четырех и большего числа измерений.

Дуальность

Базисные векторы $\partial/\partial x$ и $\partial/\partial y$ и базисные ковекторы dx и dy дуальны друг другу. Это значит, что имеют место

¹⁾ Допуская некоторую вольность, автор не делает различия между векторами (операторами дифференцирования) и прямыми линиями в направлениях, задаваемых векторами. — Прим. ред.

соотношения

$$\begin{aligned}
 dx \cdot \frac{\partial}{\partial x} &= 1, \\
 dx \cdot \frac{\partial}{\partial y} &= 0, \\
 dy \cdot \frac{\partial}{\partial x} &= 0, \\
 dy \cdot \frac{\partial}{\partial y} &= 1.
 \end{aligned}
 \tag{17.9}$$

[Напоминаем, что точкой мы обозначаем операцию свертывания ковектора с вектором. Не следует путать эту операцию со скалярным произведением евклидовой геометрии или геометрии пространства Минковского.]

Проверим первое из этих равенств. Мы определили dx как градиент функции x , а $\partial/\partial x$ — как прямую линию

$$u \mapsto (x_0 + u, y_0). \tag{17.10}$$

Значением соответствующей сложной функции является x -координата точки, принадлежащей линии (17.10), при заданном значении параметра u . Это значение $(x_0 + u)$. Скорость изменения функции при изменении u , очевидно, равна единице. Аналогичным образом проверяются остальные равенства. На рис. 17.1, где изображен полученный результат, показаны базисные векторы и ковекторы в косоугольной системе координат. Результаты, очевидно, инвариантны относительно произвольных линейных преобразований.

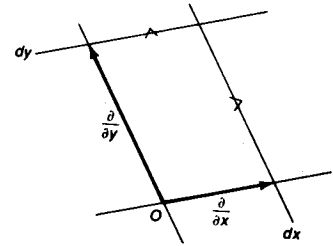


Рис. 17.1

Обозначение $\partial/\partial x$ для вектора координатного базиса подчеркивает его тесную связь с понятием производной по направлению.

Пример 1

В разд. 16 мы говорили, что для любой функции f величина $df \cdot \partial/\partial x$ представляет собой производную функцию f по направлению оси x . Действительно, полученные выше результаты позволяют написать

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \dots; \tag{17.11}$$

следовательно,

$$df \cdot \frac{\partial}{\partial x} = \frac{\partial f}{\partial x}. \tag{17.12} \quad \text{Производная по направлению}$$

Пример 2

Пусть функция f удовлетворяет дифференциальному уравнению

в частных производных

$$\frac{\partial f}{\partial x} + 2 \frac{\partial f}{\partial y} = 0 \quad (17.13)$$

с граничным условием

$$f(x, 0) = e^{-x^2}. \quad (17.14)$$

Эти требования определяют функцию f на всей плоскости.

Уравнение в частных производных (17.13) эквивалентно уравнению

$$df \cdot \left(\frac{\partial}{\partial x} + 2 \frac{\partial}{\partial y} \right) = 0, \quad (17.15)$$

которое означает, что функция f постоянна вдоль направления

$$\frac{\partial}{\partial x} + 2 \frac{\partial}{\partial y}. \quad (17.16)$$

Это позволяет найти функцию f с помощью простого геометрического построения, показанного на рис. 17.2. Получаем

$$f(x, y) = e^{-(2x - y)^2/4}. \quad (17.17)$$

[Мы продолжим рассмотрение этого примера в разд. 25.]

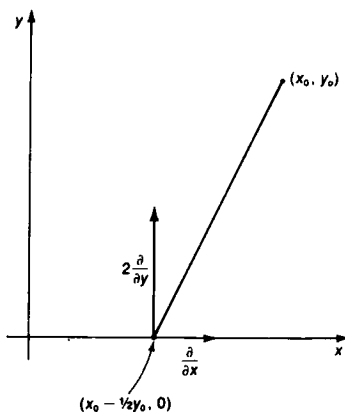


Рис. 17.2

ЗАДАЧИ

17.1. (08) Сложите графически

$$dx + 3 dy.$$

17.2. (12) Нарисуйте векторы

$$\xi = \frac{\partial}{\partial x} + \frac{\partial}{\partial t},$$

$$\eta = \frac{\partial}{\partial x} - \frac{\partial}{\partial t}.$$

Изобразите 1-формы, образующие дуальный им базис. Запишите их через dx и dt .

17.3. (13) Найдите касательный вектор к конической винтовой линии:

$$s \mapsto (s \cos s, s \sin s, s).$$

17.4. (13) Найдите градиент функции

$$f(x, y) = \sin(x + y).$$

Изобразите его на рисунке.

17.5. (16) Нарисуйте координатный базис, соответствующий полярной системе координат, и дуальный ему базис 1-форм.

17.6. (22) Выразите базисные ковекторы сферической системы координат через dx , dy и dz .

18. Пример: статическое равновесие

Мы могли бы продолжить изложение материала и перейти к рассмотрению тензоров, однако прежде следует до конца разобраться в том, что такое 1-формы и как они связаны с касательными векторами. Поэтому сделаем отступление и рассмотрим конкретный пример. Простым примером 1-формы в физике может служить обобщенная сила. Правда, это понятие не имеет прямого приложения в космологии, но оно хорошо известно всем, хотя не все знают, что обобщенная сила представляет собой 1-форму. Авторы стандартных учебников не рассматривают обобщенную силу с такой точки зрения, поэтому они вынуждены использовать абстрактные и трудные для понимания правила. Формулировки принципа виртуальных перемещений в той форме, как они приведены в большей части книг, выглядят достаточно запутанными. Как часто бывает, абстрактный язык и непонятные выражения указывают на отсутствие полной ясности.

В качестве примера рассмотрим состояние статического равновесия механической системы с конечным числом степеней свободы. Полученные результаты нетрудно распространить на случай систем, в которых действуют также электрические силы. Состояние рассматриваемой системы можно задать, указав конечное число параметров, которые называются обобщенными координатами. Определение состояния системы путем задания пространственного положения всех ее частей оказывается значительно менее эффективным. Пример возможной системы рассматриваемого типа изображен на рис. 18.1. Если мы пренебрегаем упругими волнами в пружинах, нитях, стержнях и массах, то состояние рассматриваемой системы определяется только двумя параметрами, x и θ . Его можно изобразить точкой на плоскости $\mathbb{R}^2$. Часть $\mathbb{R}^2$, состоящую из точек, соответствующих всем возможным состояниям системы, мы будем называть конфигурационным пространством. В конфигурационном пространстве понятие расстояния лишено естественного смысла. Иными словами, расстояние между точками конфигурационного пространства, изображающими два последовательных положения системы, не соответствует какому-то

Статика

Конфигурационное пространство

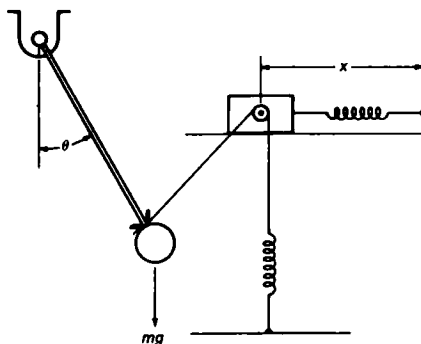


Рис. 18.1

реальному ее перемещению. Здесь евклидовость метрики $\mathbb{R}^2$ не существенна с точки зрения физики. В указанном смысле конфигурационное пространство может служить важным примером пространства без метрики.

Положение равновесия
Принцип виртуальных перемещений

Будучи предоставлена самой себе, механическая система в конце концов перейдет в положение равновесия. Эффективный метод отыскания таких положений дает принцип виртуальных перемещений (ПВП). Обычная формулировка принципа такова: «Данная механическая система будет находиться в состоянии равновесия тогда и только тогда, когда суммарная виртуальная работа всех приложенных сил равна нулю... Виртуальная работа сил реакции равна нулю для любого виртуального перемещения, совместимого с заданными кинематическими связями» [17]. Определяемую ПВП схему вычислений нетрудно реализовать. Не хватает только наглядной геометрической картины. Однако такую картину легко представить. ПВП требует

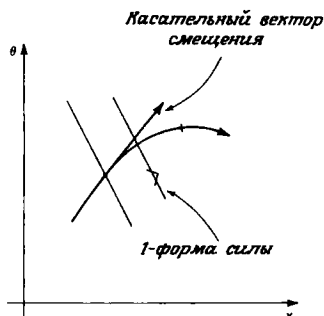


Рис. 18.2

1-форма обобщенной силы. На рисунке изображена ситуация, когда виртуальная работа равна двум единицам.

вычисления работы, которая совершается при «виртуальном перемещении» системы. Обобщенная сила связывает такое перемещение с «виртуальной работой». Следовательно, в геометрическом смысле обобщенная сила является оператором, преобразующим перемещение (касательный вектор) в работу (число). Поскольку перемещение должно быть бесконечно малым, вычисление виртуальной работы аналогично процедуре предельного перехода, которую мы ввели в предыдущем разделе. Такой предельный переход линеаризует оператор. Таким образом, обобщенная сила представляет собой линейный оператор, отображающий касательные векторы в числа, и, следовательно, является 1-формой (рис. 18.2).

Обобщенная сила

Понимание того, что сила является 1-формой, рассеивает ореол таинственности вокруг принципа виртуальных перемещений. Виртуальные перемещения оказываются необходимыми только для вычисления компонент 1-формы силы. Теперь мы можем сформулировать принцип виртуальных перемещений следующим образом: «Система находится в состоянии равновесия, если суммарная обобщенная сила, включая силы реакции связей, равна нулю».

Некоторые силы могут быть получены из потенциалов. Ясно, что такие силы представляют собой градиенты, т.е. 1-формы. Величину работы, совершаемой системой при ее движении по некоторому пути в конфигурационном пространстве, можно найти, подсчитав общее число пересекаемых линий уровня. Отсюда следует, что потенциальное поле не совершает работы при перемещении тела по замкнутому пути. Однако далеко не всякое поле является потенциальным. Примером может служить силовое поле, изображенное на рис. 18.3. Однако в каждой отдельной точке силу можно представить с помощью некоторой 1-формы.

Потенциалы

Как могло получиться, что мы до сих пор считали силу касательным вектором? Сила может быть представлена касательным вектором в том случае, если мы работаем в конфигурационном пространстве с естественной евклидовой геометрией, поскольку евклидовость метрики позволяет любой 1-форме поставить в соответствие касательный вектор. Мы покажем это, когда перейдем к рассмотрению метрических тензоров. Изучение же механики всегда начинают с механики точки, которая движется как раз в евклидовом пространстве.

Евклидово пространство

Наличие наглядного геометрического изображения сил дает нам преимущества, которые проявляются еще раз при рассмотрении систем со связями. Ограничимся случаем идеальных свя-

Связи

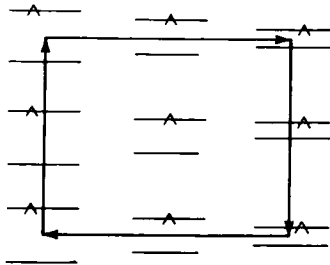


Рис. 18.3
Неконсервативное силовое поле. При обходе по изображенному замкнутому прямоугольному контуру совершается работа, равная примерно 8 единицам.

Множители Лагранжа

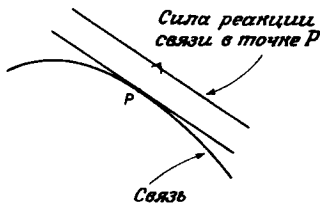


Рис. 18.4

зей, т. е. будем считать, что трение отсутствует. Обычно говорят, что сила реакции ортогональна поверхности связи. Однако если вы действительно поняли идею ковариантности относительно линейных преобразований, то вам должно быть ясно, что это утверждение бессмысленно. Действительно, перпендикулярность остается неопределенным понятием в любом конфигурационном пространстве, если только оно не обладает метрикой. Корректное с точки зрения геометрии представление связи дано на рис. 18.4. Мы видим, что сила, рассматриваемая как 1-форма, параллельна поверхности связи, а параллельность — ковариантное понятие. И в этом случае нековариантное описание возникло из-за неправомерного обобщения механики частицы в евклидовом пространстве.

Для нахождения положения равновесия системы со связями используется метод множителей Лагранжа. Ни в одном стандартном учебнике вы не найдете наглядного геометрического изображения этого метода. Однако такое изображение легко построить, поскольку мы знаем, что сила является 1-формой. По определению сила реакции связи может быть сколь угодно большой. Пусть связь задается условием

$$f(x, \theta) = 0, \quad (18.1)$$

тогда параллельная этой поверхности 1-форма, очевидно, совпадает с градиентом df . Само уравнение связи (18.1) определяет линию нулевого уровня функции f . Равновесие требует, чтобы суммарная сила обращалась в нуль. Если Q — сумма всех сил, за исключением силы реакции, и, следовательно, известна, то условие равновесия имеет вид

$$Q + \lambda df = 0, \quad (18.2)$$

где сила реакции связи записана в виде λdf . Необходимость введения неопределенного множителя λ продиктована тем, что мы знаем только направление действия силы реакции, но не знаем ее величину. Ковекторное уравнение (18.2) должно быть решено относительно n координат положения равновесия и множителя λ . Так как в уравнение (18.2) входят ковекторы, оно на самом деле представляет собой n уравнений, по одному для каждой компоненты. Еще одним уравнением служит само уравнение связи (18.1). Таким образом, мы имеем $(n + 1)$ уравнение для $(n + 1)$ неизвестных.

Пример

Рассмотрим маятник, изображенный на рис. 18.5. На рисунке показаны гравитационная сила и действующая со стороны стержня сила, величина которой не известна.

Если мы задаем состояние системы обобщенной координатой θ , то потенциальная энергия маятника в гравитационном поле дается выражением

$$V = mgl(1 - \cos \theta), \quad (18.3)$$

а соответствующая гравитационной силе 1-форма

$$Q = -dV \quad (18.4)$$

имеет вид

$$Q = -mgl \sin \theta d\theta. \quad (18.5)$$

Система находится в состоянии равновесия, когда сила Q обращается в нуль. Это имеет место, если

$$\sin \theta d\theta = 0, \quad (18.6)$$

т. е. при

$$\theta = 0, \pi. \quad (18.7)$$

В рассматриваемом случае конфигурационное пространство представляет собой окружность, и 1-форму силы легко изобразить. «Линии уровня» в действительности являются точками уровня; на рис. 18.6 мы указали их в виде отметок времени. На рис. 18.6 изображены также само конфигурационное пространство и 1-форма силы.

С другой стороны, мы можем описывать состояние рассматриваемой системы с помощью обобщенных координат (x, y) при условии, что теперь мы будем учитывать возникающую в стержне упругую силу (в θ -пространстве упругую силу можно было не учитывать, так как ее проекция на касательную к окружности равнялась нулю). В этих переменных гравитационная сила γ дается выражением

$$\gamma = -mg dy, \quad (18.8)$$

а действующая со стороны стержня сила упругости имеет вид

$$p = \lambda(\cos \theta dy - \sin \theta dx), \quad (18.9)$$

где нам известно только направление действия силы p , а величину λ необходимо определить. Согласно условию равновесия,

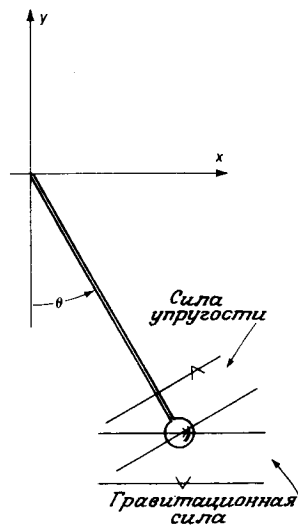


Рис. 18.5

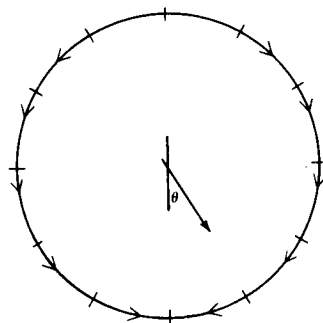


Рис. 18.6

Одномерное конфигурационное пространство, соответствующее маятнику. Изображена 1-форма гравитационной силы. Равновесие имеет место в точках $\theta = 0$ и $\theta = \pi$, где сила обращается в нуль.

[См. задачу 15.1.]

суммарная сила должна обращаться в нуль,

$$\gamma + p = 0, \quad (18.10)$$

т. е. должно выполняться равенство

$$(\lambda \cos \theta - mg) dy - \lambda \sin \theta dx = 0. \quad (18.11)$$

Это ковекторное уравнение, поэтому каждая компонента должна независимо обращаться в нуль. Последнее требование определяет как положение равновесия

$$\sin \theta = 0, \quad (18.12)$$

так и величину возникающей в стержне силы упругости

$$\lambda \cos \theta = mg. \quad (18.13)$$

В результате получим

$$\theta = 0, \quad \lambda = mg, \quad (18.14)$$

$$\theta = \pi, \quad \lambda = -mg. \quad (18.15)$$

Эту задачу можно также рассматривать как задачу со связью и решать с помощью введения множителей Лагранжа. Как показано на рис. 18.7, равновесие имеет место в точках,

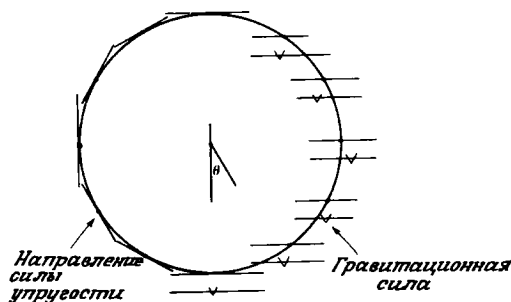


Рис. 18.7

Маятник, рассматриваемый как двумерная система со связью. Справа изображена 1-форма гравитационной силы, слева указано направление силы упругости, действующей на чечевицу маятника. Величина силы, действующей со стороны связи, меняется, но если сила трения отсутствует, направление должно быть таким, как изображено. Эти две силы могут взаимно компенсироваться, только если их направления совпадают, что имеет место в верхней и нижней точках.

где силы параллельны связям. Детали вычислений мы оставим до задач 18.1 и 18.2.

Ценность любого нового формализма должна быть доказана путем применения его к конкретным случаям. Надеюсь, что рассмотренные в книге примеры и задачи убедят читателя, сколь рационален язык 1-форм в применении к силам и смещениям. Когда мы перейдем к рассмотрению движения частиц, мы увидим, что импульс также является 1-формой. Таким образом, сила и скорость изменения импульса имеют одну и ту же геометрическую природу.

ЗАДАЧИ

18.1. (20) Рассмотрите пример на стр. 145, используя множители Лагранжа (см. задачи 15.4 и 15.5).

18.2. (08) Предположим, что на маятник вместо гравитационной силы действует сила

$$Q = 2x dx + y dy.$$

Найдите соответствующий этой силе потенциал.

18.3. (12) Работая непосредственно в θ -пространстве, найдите положения равновесия для системы, рассмотренной в задаче 18.2. Сделайте соответствующее геометрическое построение.

18.4. (15) Для задачи 18.2 найдите положения равновесия, рассматривая ее как задачу со связью. Сделайте соответствующее геометрическое построение и сравните с результатом задачи 18.3.

18.5. (15) В задаче 18.2 найдите силу, действующую на маятник со стороны стержня.

18.6. (20) Предположим, что на стержень действует гравитационная сила плюс сила, явный вид которой приведен в задаче 18.2. Любым способом найдите положения равновесия. Сделайте соответствующее геометрическое построение.

18.7. (20) Пусть на маятник действует сила, потенциал которой имеет вид

$$V = x^4 + y^4.$$

Найдите положения равновесия. Сначала проведите графическое построение, затем вычислите положения равновесия, используя множители Лагранжа.

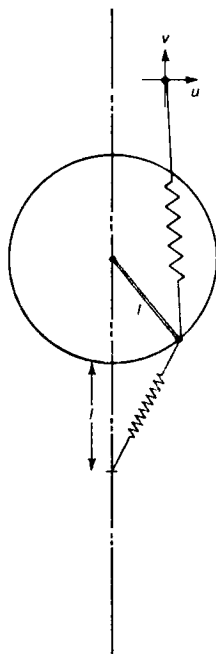


Рис. 18.8

18.8. (24) Повторите задачу 18.7 для потенциала

$$V = x^3 + y^3.$$

18.9. (20) Повторите задачу 18.7 для потенциала

$$V = |x|^3 + |y|^3.$$

18.10. (35) Теория катастроф изучает изменения положений равновесия, вызванные изменениями параметров задачи (см., например, Scientific American, April 1976, а также книгу [31]). Типичный пример такой системы изображен на рис. 18.8. Предположим, что потенциальная энергия растяжения пружин имеет вид

$$V = \frac{k}{2} (L - L_0)^2,$$

где L_0 — некоторая постоянная, а L — расстояние между концами пружины. Исследуйте зависимость положения равновесия этой системы от положения точки, где закреплена верхняя пружина. Считайте, что $l = k = L_0 = 1$.

19. Пример: диспергирующие волны

В качестве еще одного примера физической задачи, при решении которой мы естественным образом приходим к понятию 1-формы, рассмотрим распространение пакета диспергирующих волн. Такими волнами могут быть волны на воде, упругие волны или волны, которые соответствуют квантовомеханическим частицам. Этот пример заслуживает внимательного рассмотрения не только потому, что он дает возможность проследить взаимосвязь между касательными векторами и 1-формами, но и потому, что в последующих разделах книги движение волновых пакетов будет использоваться в качестве модели движения частиц в общей теории относительности. Мы воспользуемся этой моделью при рассмотрении красного смещения, симметрий, тахионов и черных дыр. Движение волн на воде обладает поразительной симметрией и может рассматриваться как еще одна физическая реализация математической модели, используемой в специальной теории относительности. Поэтому мы посвятим целый раздел изучению этого вопроса. Геометрический подход к описанию движения волновых пакетов может оказаться полезным для физиков, которые занимаются физикой твердого тела и называют такие пакеты квазичастицами.

[Простым и достаточно хорошим введением в теорию волн может служить книга [27].]

Движение волновых пакетов можно обсуждать, вообще не упоминая о метрической структуре пространства-времени. Наоборот, чуть позже мы покажем, что метрическую структуру можно изучать, исследуя движение волновых пакетов. Такой подход позволяет рассматривать лоренцеву метрику специальной теории относительности как проявление особенно простого закона движения волн.

Структура пространства-времени проявляется в нетривиальных физических закономерностях. Поскольку мы не рассматриваем метрику, нужно рассчитывать на что-то другое. Оказывается, что интересную информацию о структуре пространства-времени дает анализ линейных уравнений в частных производных. Решения этих уравнений представляет собой волны. Прежде всего мы рассмотрим решения, имеющие характер плоских волн. Потом сформируем из плоских волн волновые пакеты. Амплитуда волнового пакета уже не является постоянной величиной. Мы ограничимся рассмотрением случая высокочастотных волн, для которых амплитуда пакета мало меняется на расстояниях порядка нескольких длин волн.

С точки зрения крупномасштабного усреднения волновой пакет выглядел бы просто сгустком энергии, в котором нельзя различить отдельные волны. Именно такую модель частицы мы будем использовать, а движение пакетов рассматривать как модель движения частиц. Мы увидим, что существуют некоторый касательный вектор и некоторая 1-форма, которые естественным образом сопоставляются волновому пакету. Касательный вектор указывает направление движения волнового пакета, 1-форма описывает расположение в пространстве-времени гребней волны, касательный вектор называется групповой скоростью, а 1-форма — фазовой скоростью¹⁾. При традиционном подходе с фазовой скоростью обращаются как с касательным вектором. Однако это возможно только в случае двух измерений. В двумерном пространстве 1-форма описывается набором параллельных линий, и ее ориентацию можно рассматривать как «нечто вроде» вектора. Сложности, вызванные тем, что это только «нечто вроде», исчезают, если понять, что фазовая скорость в действительности представляет собой 1-форму. Наш подход к проблеме, и в особенности разработанный в разд. 27 метод волновых диаграмм, является новым. Изложе-

Волновые пакеты

Уравнения в частных производных

[Для наших целей вполне достаточно иметь общее представление об уравнениях в частных производных.]

Полуклассический подход

[Эта процедура перехода к пределу проиллюстрирована на рис. 19.6 — 19.8.]

Групповая скорость и фазовая «скорость»

[Чтобы облегчить изучение материала, мы будем возвращаться к рассмотрению отдельных примеров на протяжении всего обсуждения. Соответствующие места в тексте, как обычно, будут выделяться вертикальной линией.]

¹⁾ Вводимые автором здесь групповая и фазовая скорости в отличие от общепринятого определения являются геометрическими объектами в пространстве-времени. — *Прим. ред.*

ние традиционного подхода к вопросу о групповой скорости можно найти во многих книгах; лучше всего, по нашему мнению, он освещен в книге [42].

Начнем с рассмотрения линейного уравнения в частных производных второго порядка.

Пример 1

Квантовомеханическое движение бесспиновых частиц описывается уравнением Клейна—Гордона

$$\frac{\partial^2 \Psi}{\partial t^2} - \frac{\partial^2 \Psi}{\partial x^2} + \Psi = 0. \quad (19.1)$$

[Не обязательно знать, каким образом уравнение получено. Примем его как данное. Уравнение (19.1) описывает также распространение упругих волн по балке.]

Плоские волны

Прежде всего исследуем решения нашего уравнения в частных производных, имеющие вид плоских волн, т. е. решения вида

$$\Psi = A \cos \theta, \quad (19.2)$$

где A — амплитуда волны, которую мы пока считаем постоянной, а θ — фаза волны. Положение гребней волны дается уравнением

$$\theta = 2\pi n. \quad (19.3)$$

где n — целое число. Как показано на рис. 19.1, покоящийся наблюдатель видит сечение волны вертикальной мировой линией W . Скорость, с которой гребни волны проходят мимо наблюдателя, называется частотой. Мы будем пользоваться угловой частотой, которая определяется соотношением

$$\omega \equiv \frac{\partial \theta}{\partial t}. \quad (19.4)$$

В любой фиксированный момент времени пространственное расположение гребней волн дается горизонтальным сечением соответствующей волне пространственно-временной диаграммы (рис. 19.2). Скорость появления гребней волн при изменении пространственной координаты пропорциональна волновому числу, которое определяется следующим образом:

$$k \equiv \frac{\partial \theta}{\partial x}. \quad (19.5)$$

Функция

$$\Psi = A \cos (\sqrt{2}t - x) \quad (19.6)$$

является частным решением уравнения (19.1), имеющим вид

[Множитель 2π в выражении для фазы будет возникать постоянно. Он связан с соответствующим множителем в формуле для преобразования Фурье, и, так же как и там, его нельзя исключить.]

[В случае трех пространственных измерений у k , конечно, было бы три компоненты.]

[Продолжение примера 1.]

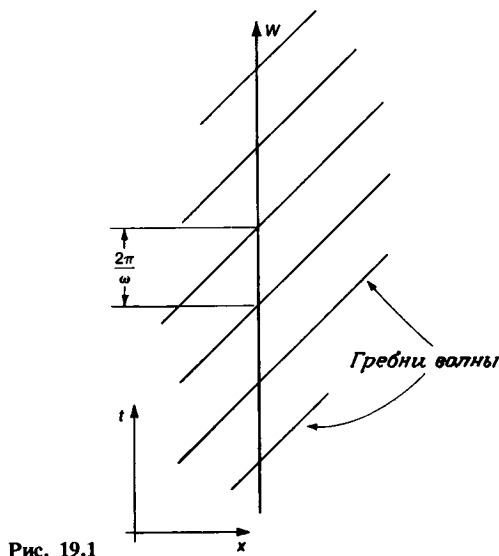


Рис. 19.1

плоской волны. Соответствующая этому решению фаза θ имеет вид

$$\theta = \sqrt{2}t - x. \quad (19.7)$$

Конфигурация гребней волн изображена на рис. 19.3. Частота ω и волновое число k равны соответственно

$$\omega = \sqrt{2}, \quad (19.8)$$

$$k = -1. \quad (19.9)$$

Частота и волновое число, очевидно, являются компонентами 1-формы градиента фазы. Действительно,

$$d\theta = \frac{\partial \theta}{\partial x} dx + \frac{\partial \theta}{\partial t} dt, \quad (19.10)$$

и, следовательно,

$$d\theta = k dx + \omega dt. \quad (19.11)$$

Градиент фазы описывает кинематику гребней волн, их локальное расположение в пространстве-времени. Строго говоря, конфигурация гребней волн несет в себе меньше информации, чем 1-форма градиента фазы. Действительно, одно и то же расположение гребней волн соответствует как самой 1-форме, так и 1-форме, умноженной на минус единицу. С другой стороны,

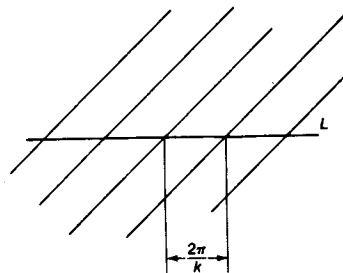


Рис. 19.2

1-форма градиента фазы

конфигурация гребней дает более подробное описание волны, так как градиент определяет фазу только с точностью до постоянного фазового сдвига. Выбор того или другого описания определяется характером рассматриваемой задачи.

[Продолжение примера 1.]

Для рассматриваемой в примере плоской волны градиент фазы имеет вид

$$d\theta = \sqrt{2} dt - dx. \quad (19.12)$$

Как видно из рис. 19.3, гребни волн перемещаются вправо со скоростью

$$\frac{\omega}{k} = \sqrt{2}. \quad (19.13)$$

Дисперсионное уравнение

При каких условиях плоская волна, имеющая вид (19.2), будет решением рассматриваемого уравнения в частных производных? Чтобы ответить на этот вопрос, сведем уравнение в частных производных к некоторому алгебраическому соотношению, связывающему компоненты градиента фазы.

[Продолжение примера 1.]

Для волны

$$\Psi = A \cos \theta, \quad (19.14)$$

если k и ω постоянны, как это имеет место в плоской волне, получаем

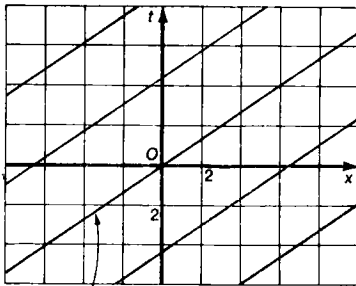
$$\frac{\partial^2 \Psi}{\partial t^2} = -A\omega^2 \cos \theta, \quad (19.15)$$

$$\frac{\partial^2 \Psi}{\partial x^2} = -Ak^2 \cos \theta. \quad (19.16)$$

Таким образом, должно выполняться соотношение

$$-\omega^2 + k^2 + 1 = 0. \quad (19.17)$$

Это соотношение называется дисперсионным уравнением. Оно полностью эквивалентно исходному уравнению в частных производных. Нетрудно понять, каким образом можно восстановить дифференциальное уравнение по дисперсионному уравнению. Просто нужно каждой степени ω поставить в соответствие производную по времени, а каждой степени k — производную по координате. Действительно, распространение волн часто описывают, задавая именно дисперсионное уравнение, а не соответствующее уравнение в частных производных. При экспериментальном исследовании волновых систем дисперсионное уравнение часто получается непосредственно в результате измерений.



Гребень волны

Рис. 19.3

[Если вы изучали квантовую механику, то эти правила должны быть вам знакомы.]

Несколько распространенных дисперсионных уравнений: обычное волновое уравнение

$$\omega^2 = k^2; \quad (19.18)$$

уравнение Клейна—Гордона

$$\omega^2 = k^2 + 1; \quad (19.19)$$

упругие волны в напряженном жестком стержне

$$\omega^2 = k^2 - k^4; \quad (19.20)$$

волны на глубокой воде

$$\omega^4 = k^2. \quad (19.21)$$

Когда существенно то обстоятельство, что градиент фазы представляет собой 1-форму, а не касательный вектор? Рассмотрим градиент фазы в случае трех измерений (рис. 19.4). Для изображенной на рисунке ситуации мы не можем указать направление распространения гребней волн в пространстве. Однако мы можем определить скорость изменения фазы волны с точки зрения любого заданного наблюдателя. Если λ — касательный вектор к мировой линии наблюдателя, то скорость изменения фазы дается выражением $d\theta \cdot \lambda$, но это не что иное, как результат действия 1-формы на касательный вектор.

Теперь рассмотрим решения, для которых k и ω не строго постоянны. Такие волны уже не являются плоскими. Однако если k и ω мало меняются на расстояниях порядка длины волны и за время, равное периоду волны, то такие решения мало отличаются от плоских волн и для их описания можно приближенно использовать плосковолновой формализм. Пример почти плоской волны приведен на рис. 19.5. При рассмотрении почти плоских волн считают, что величины k и ω зависят не от x и t , а от медленно меняющихся переменных ϵx и ϵt , где $\epsilon \rightarrow 0$. В этом пределе частота и волновое число меняются медленно по сравнению с самой фазой. Тогда для решения вида

$$\Psi = A \cos \theta \quad (19.22)$$

получаем

$$\frac{\partial \Psi}{\partial t} = -A\omega \sin \theta, \quad (19.23)$$

Распространенные дисперсионные уравнения

[Чтобы упростить соответствующие выражения, мы исключили из них все физические константы. Это эквивалентно выбору канонической системы координат для каждой задачи.]

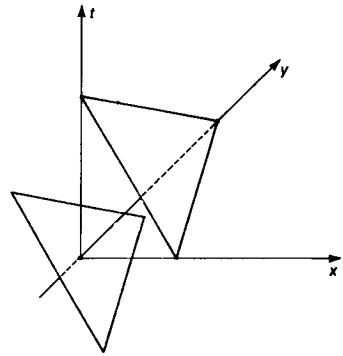


Рис. 19.4

Гребни в случае двух пространственных и одного временного измерений. Изображены два гребня, один из которых проходит через начало координат.

Почти плоские волны

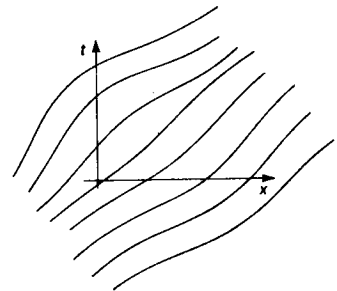


Рис. 19.5

$$\frac{\partial^2 \Psi}{\partial t^2} = -A\omega^2 \cos \theta - \varepsilon A \frac{\partial \omega}{\partial t} \sin \theta. \quad (19.24)$$

Мы видим, что наибольшую величину при $\varepsilon \rightarrow 0$ имеют члены, которые входят в дисперсионное уравнение для плоских волн. Таким образом, частота и волновое число изменяются по величине, но в любой точке должны по-прежнему удовлетворять дисперсионному уравнению.

[Продолжение примера 1.]

В рассматриваемом случае в качестве решения, мало отличающегося от плоской волны, можно было бы взять решение вида

$$\Psi = A \cos \left\{ \frac{1}{\varepsilon} [1 + 2\sqrt{2}\varepsilon t - 2\varepsilon x - (\varepsilon t)^2 - (\varepsilon x)^2]^{1/2} \right\}. \quad (19.25)$$

Это решение пригодно вплоть до x порядка $1/\varepsilon$, поэтому члены εx и $\varepsilon^2 x^2$ в (19.25), вообще говоря, могут не быть малыми и ими нельзя пренебречь. Вблизи начала системы отсчета получаем

$$\Psi \sim A \cos (\sqrt{2} t - x + \text{const}). \quad (19.26)$$

Соответственно градиент фазы имеет вид

$$d\theta = \sqrt{2} dt - dx. \quad (19.27)$$

При работе с волновыми пакетами абсолютное значение фазы обычно не представляет интереса, поэтому постоянной в выражении (19.26) можно пренебречь.

При больших x и (или) t членами типа εx пренебречь нельзя. В этом случае частота дается выражением

$$\omega = \frac{\sqrt{2} + (\varepsilon t)}{[1 + 2\sqrt{2}\varepsilon t - 2\varepsilon x + (\varepsilon t)^2 - (\varepsilon x)^2]^{1/2}}. \quad (19.28)$$

Из (19.28) видно, что в окрестности любой точки частота действительно меняется медленно по сравнению с самой фазой. То же справедливо и для волнового числа.

[Совсем не очевидно, каким образом можно находить решения такого типа. Соответствующие методы можно найти в превосходной книге [42].]

Групповая скорость

Раз мы перешли от плоской волны к волне, частота и волновое число которой могут изменяться, появляется возможность каким-то способом пометить волновой фронт и следить за его перемещением в пространстве-времени. В качестве метки может служить, например, максимум частотного распределе-

ния. Вектор, указывающий направление перемещения волнового пакета в пространстве-времени, называется *вектором групповой скорости*. Позже мы увидим, что энергия волны и максимумы амплитуды, равно как и информация о частоте, перемещаются в том же самом пространственно-временном направлении.

Уравнение переноса для частоты можно получить из дисперсионного уравнения следующим образом. Мы знаем, что

Частота

$$\frac{\partial \omega}{\partial x} = \frac{\partial^2 \theta}{\partial x \partial t} = \frac{\partial^2 \theta}{\partial t \partial x} = \frac{\partial k}{\partial t}. \quad (19.29)$$

Если мы решим дисперсионное уравнение относительно волнового числа

$$k = K(\omega) \quad (19.30)$$

и затем используем равенство (19.29), то получим уравнение, описывающее перенос информации о частоте.

Пример 2

Выше мы приводили дисперсионное уравнение для волн на глубокой воде

$$\omega^4 = k^2. \quad (19.31)$$

Одно из решений этого уравнения имеет вид

$$k = -\omega^2. \quad (19.32)$$

Используя уравнение (19.29), получаем уравнение

$$\frac{\partial \omega}{\partial x} + \frac{\partial}{\partial t} (\omega^2) = 0 \quad (19.33)$$

или, иначе

$$\frac{\partial \omega}{\partial x} + 2\omega \frac{\partial \omega}{\partial t} = 0. \quad (19.34)$$

Из последнего уравнения следует, что частота не меняется при перемещении в пространственно-временном направлении, которое задается вектором

$$v = \frac{\partial}{\partial x} + 2\omega \frac{\partial}{\partial t}. \quad (19.35)$$

Этот вектор совпадает по направлению с вектором групповой скорости для волн на глубокой воде.

[Обратите внимание, насколько естественно наше обозначение для касательного вектора. Еще раз просмотрите обсуждение операции взятия производной по направлению в разд. 17.]

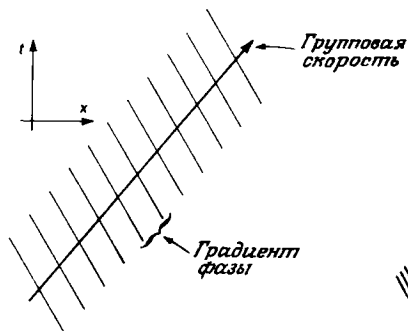


Рис. 19.6

Пространственно-временная диаграмма для волнового пакета. Конечную протяженность пакета мы учли, изобразив гребни волн короткими линиями. Для наглядности представлен случай относительно большой длины волны. Рис. 19.7 в большей степени соответствует высокочастотному волновому пакету.

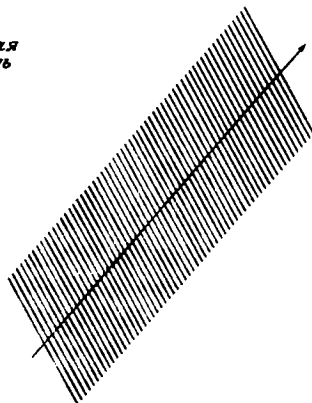


Рис. 19.7

Волновой пакет, частота которого в пять раз больше частоты пакета, изображенного на рис. 19.6.

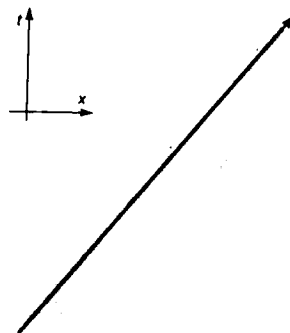


Рис. 19.8

Предел.

Амплитуда

Ниже мы увидим, что движение максимума волнового пакета происходит в том же направлении. Мы представляем себе волновой пакет в виде решетки, изображенной на рис. 19.6. На рис. 19.7 и 19.8 изображены волновые пакеты, частота которых больше. Заметим, что фактическая протяженность волнового пакета для электрона порядка $8 \cdot 10^{-21}$ с. Мы рассматриваем волны, амплитуды которых имеют максимум в некоторой точке пространства и плавно спадают по сторонам от волнового пакета. Положение гребней волн определяется градиентом фазы, компоненты которого удовлетворяют дисперсионному уравнению. Волновой пакет перемещается в пространстве-времени в направлении вектора групповой скорости, который находится дифференцированием дисперсионного уравнения. В последующих разделах мы распространим полученные результаты на случай уравнений в частных производных с медленно меняющимися коэффициентами. Такие уравнения будут использоваться при обсуждении движения волновых пакетов во внешних полях.

ЗАДАЧИ

19.1. (10) Вычислите k для плоской волны, рассмотренной в примере на стр. 154.

19.2. (12) Нарисуйте на пространственно-временной диаграмме волновой пакет для волн, распространяющихся на воде. Укажите на диаграмме k , ω и компоненты групповой скорости. Выполните рисунок корректно в количественном отношении.

19.3. (21) Найдите значение групповой скорости для упругих волн с дисперсионным уравнением (19.20).

19.4. (20) Покажите, что для волн на глубокой воде скорость переноса энергии в два раза меньше скорости гребней волны. Используя пространственно-временную диаграмму, объясните, каким образом это можно было бы наблюдать.

19.5. (28) Покажите, что в примере на стр. 154 все векторы групповых скоростей выходят из одной и той же точки в прошлом.

20. Отображения

Отображение множеств — фундаментальное математическое понятие. В тех случаях, когда множества обладают структурой, естественно поставить вопрос, в какой степени эта структура переносится при прямом и обратном отображении. В этом разделе мы будем рассматривать только множества, которые являются линейными векторными пространствами. Однако полученные результаты справедливы и в более общем случае многообразий, обсуждаемом в разд. 26. Мы увидим, что существует тонкое свойство дуальности в соотношениях между векторами и ковекторами, которое теряется при традиционном координатном подходе.

Начнем с отображения ψ между двумя векторными пространствами A и B . При этом не обязательно, чтобы пространства имели одинаковую размерность. Если $A = \mathbb{R}$, то при отображении мы получаем параметризованную кривую; если $B = \mathbb{R}$, то имеем дело с обычной функцией. В основном нас будет интересовать случай, когда размерность B больше размерности A , поэтому все наши рисунки будут относиться к такому случаю. Однако полученные результаты остаются в силе и при снятии указанного ограничения. Если размерность A меньше размерности B , то можно сказать, что отображение задает в B параметризованное подмножество; это понятие яв-

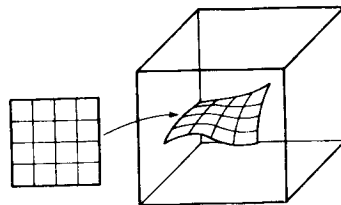


Рис. 20.1
Отображение, выделяющее подпространство в пространстве большего числа измерений.

Отображение

ляется обобщением введенного ранее понятия параметризованной кривой (рис. 20.1).

Функция

Предположим далее, что на множестве B задана некоторая функция, т. е. отображение

$$f: B \rightarrow \mathbb{R} \quad (20.1)$$

(рис. 20.2). Тогда отображение ψ естественным образом определяет функцию f^* на множестве A :

$$f^*: A \rightarrow \mathbb{R}; a \mapsto f[\psi(a)]. \quad (20.2)$$

Ограничение на подпространство

[Мы используем общепринятое обозначение с помощью звездочки. Оно не имеет отношения к операции комплексного сопряжения.]

Это также изображено на рис. 20.2. Чтобы вычислить значение f^* в точке a , принадлежащей множеству A , нужно сначала с помощью ψ отобразить a в множество B , а затем взять значение функции f в точке $\psi(a)$. Функцию f^* будем называть ограничением (pullback) функции f на подпространство, задаваемое отображением ψ .

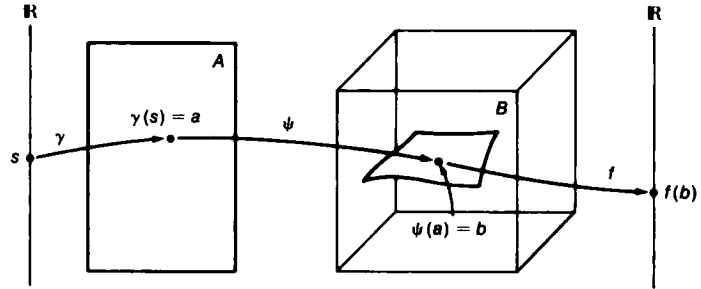


Рис. 20.2

Параметризованная кривая в двумерном пространстве A , отображение из A в трехмерное пространство B и функция на B .

Параметризованная кривая

Существует также дуальная операция, ассоциируемая с параметризованной кривой в A :

$$\gamma: \mathbb{R} \rightarrow A; \quad (20.3)$$

она схематически изображена на рис. 20.2. Теперь она осуществляется с помощью отображения ψ в том же направлении. Результатом является кривая γ_* :

Продолжение из подпространства

$$\gamma_*: \mathbb{R} \rightarrow B; s \mapsto \psi[\gamma(s)], \quad (20.4)$$

которую будем называть продолжением (push-forward) γ с помощью отображения ψ .

Касательные векторы и 1-формы являются локальными линейными аппроксимациями кривых и функций. Очевидно, что несущественно, когда проводится аппроксимация — до или после отображения ψ . Это требование определяет гладкое отображение. Всюду ниже мы будем предполагать, что рассматриваются гладкие отображения.

Гладкие отображения позволяют определить продолжение касательных векторов пространства A в пространство B и ограничения на 1-формы. При этом мы будем использовать введенное выше обозначение с помощью звездочки. Заметим, что касательный вектор, определенный в B , не может быть ограничен на подпространство A . Действительно, представим себе A как подпространство пространства B . Как показано на рис. 20.3, не существует однозначного способа спроецировать вектор на подпространство. Аналогичным образом, хотя 1-форма, определенная на B , естественным образом определяет 1-форму на подпространстве (рис. 20.4), обратное не верно.

Ограничение формы на подпространство осуществляется с помощью цепного правила. Пусть $(u, v, \dots)$ — координаты в A , а $(x, y, \dots)$ — координаты в B . Координаты не обязательно должны быть линейными. Отображение ψ задается наб-

[Для наших целей вполне достаточно общего представления о том, что такое гладкость. Различия между C^r , C^∞ и C^ω не имеют глубокого физического содержания.]

Касательные векторы и 1-формы

Вычисление ограничения на подпространство

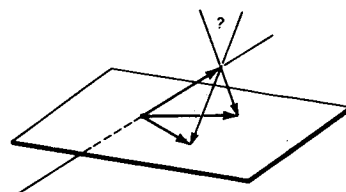


Рис. 20.3

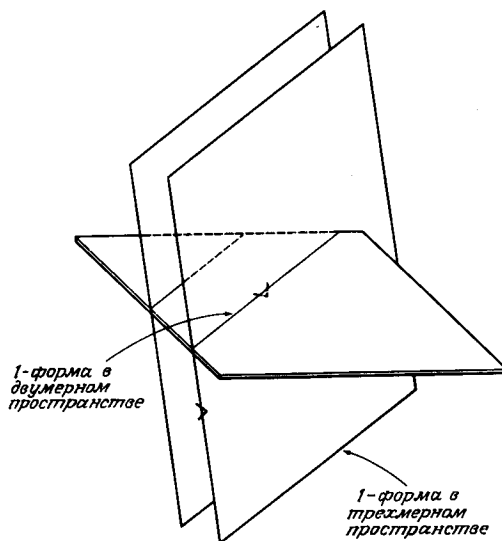


Рис. 20.4

ром функций

$$\begin{aligned}x &= X(u, v, \dots), \\y &= Y(u, v, \dots), \\&\text{и т. д.}\end{aligned}\tag{20.5}$$

Простейшая функция, определенная на B , имеет вид $f = x$. Ее градиент dx представляет собой базисную 1-форму. Ограничением функции x на A является функция

$$x^* = X(u, v, \dots),\tag{20.6}$$

градиент которой имеет вид

$$dx^* = \frac{\partial X}{\partial u} du + \frac{\partial X}{\partial v} dv + \dots\tag{20.7}$$

Поскольку любая базисная 1-форма может быть ограничена на A , а оператор градиента d линеен и коммутирует с отображением ψ , любую 1-форму можно представить в виде линейной комбинации базисных 1-форм и, следовательно, ограничить на A .

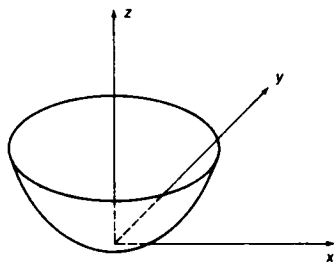


Рис. 20.5

Пример

Пусть $A = \mathbb{R}^2$, $B = \mathbb{R}^3$, и пусть отображение ψ имеет вид (рис. 20.5)

$$\psi: (u, v) \mapsto (x, y, z) = (u, v, u^2 + v^2).\tag{20.8}$$

Тогда имеем

$$\begin{aligned}x^* &= u, \\y^* &= v, \\z^* &= u^2 + v^2;\end{aligned}\tag{20.9}$$

следовательно,

$$\begin{aligned}dx^* &= du, \\dy^* &= dv, \\dz^* &= 2u du + 2v dv.\end{aligned}\tag{20.10}$$

Как мы видим, $dz^* = 0$ в точке $(0,0)$, что вполне естественно, поскольку в этой точке параболоид параллелен 1-форме dz .

Продолжение касательных векторов осуществляется аналогично предыдущему. Кривая в A

$$\gamma: s \mapsto (s, 0, \dots) \quad (20.11)$$

имеет касательный вектор, который мы обозначаем $\partial/\partial u$. Касательный вектор $(\partial/\partial u)_*$ к кривой γ_*

$$\gamma_*: \mathbb{R} \rightarrow B; s \mapsto [X(s, 0, \dots), Y(s, 0, \dots), \dots]. \quad (20.12)$$

можно найти при помощи цепного правила

$$\left(\frac{\partial}{\partial u}\right)_* = \frac{\partial X}{\partial u} \frac{\partial}{\partial x} + \frac{\partial Y}{\partial u} \frac{\partial}{\partial y} + \dots \quad (20.13)$$

В случае когда пространства A и B имеют одинаковую размерность, отображение ψ можно обратить; тогда векторы и 1-формы могут отображаться в обе стороны. С этой ситуацией мы часто сталкиваемся при переходе от одной системы координат к другой.

Вычисление продолжения из подпространства

[Еще раз обратите внимание на то, насколько естественно обозначение для касательного вектора!]

[Первый раз мы воспользуемся операцией ограничения на подпространство в разд. 29 при рассмотрении движения волновых пакетов. Она будет использована также в разд. 39 для нахождения метрики 3-сферы.]

21. Тензоры

*Напряженной, — сказал Тензор,
Напряженной, — сказал Тензор,
и напряжения, опасения,
разногласия начались.*

Альфред Бестер

Теперь мы достаточно подготовлены, чтобы перейти к изучению линейных операторов более общего вида, чем ковекторы. Ковекторами мы называли отображения вида $V \rightarrow \mathbb{R}$, а множество всех заданных на линейном пространстве V ковекторов обозначали символом V^* . Что можно сказать об операторах, осуществляющих отображение $V^* \rightarrow \mathbb{R}$? Ничего нового. Таким оператором может быть любой вектор v из линейного пространства V . Действительно, любой такой вектор определяет отображение

$$v: \omega \mapsto \omega \cdot v, \quad (21.1)$$

более того, любой оператор такого вида эквивалентен некоторому вектору.

Теперь рассмотрим несколько более сложные линейные операторы вида $V \rightarrow V^*$. Такие операторы мы будем называть

Тензоры

[Для обозначения тензоров мы по возможности будем использовать рукописный шрифт. Здесь v и w — векторы, а k — действительное число.]

тензорами. Некоторые авторы называют их тензорами второго ранга и применяют термин «тензор» для любого линейного оператора. В таких более общих случаях мы будем говорить, что имеем дело с тензором общего вида. В основе тензорной алгебры лежит понятие линейности. Какие бы операторы ни действовали между пространствами E и F , имеющими структуру линейного вектора пространства, из их множества всегда можно выделить особый класс линейных операторов. Как V , так и V^* являются векторными пространствами. Оператор $\mathcal{B}$, определяемый следующим образом:

$$\mathcal{B}: V \rightarrow V^*, \quad (21.2)$$

представляет собой линейный оператор, если справедливы равенства

$$\mathcal{B}(kv) = k \mathcal{B}(v), \quad (21.3)$$

$$\mathcal{B}(v + w) = \mathcal{B}(v) + \mathcal{B}(w). \quad (21.4)$$

В левой части равенства (21.4) складываются векторы, в правой части — ковекторы.

Линейность

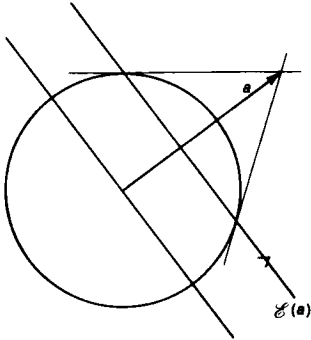


Рис. 21.1

Пример

Все метрические соотношения в евклидовой геометрии определяются заданием некоторого тензора. В разд. 3 было показано, что при введении канонической системы координат этот тензор задает окружность. Окружность можно рассматривать как оператор (мы обозначим его $\mathcal{E}$), который отображает векторы в ковекторы. Способ, позволяющий осуществить такое отображение, показан на рис. 21.1. Возьмем произвольный вектор, конец которого лежит вне единичной окружности. Из конца вектора (который мы обозначим через a) проведем касательные к окружности. Через точки касания проведем прямую линию. отождествим эту линию с линией уровня нашего ко-вектора, соответствующей его значению, равному единице.

Является ли этот оператор тензором? Да, если он линеен, а он действительно линеен. Чтобы доказать это, мы должны сначала убедиться, что

$$\mathcal{E}(ka) = k\mathcal{E}(a), \quad (21.5)$$

а затем показать справедливость соотношения

$$\mathcal{E}(a + b) = \mathcal{E}(a) + \mathcal{E}(b). \quad (21.6)$$

Первое из указанных равенств следует из рассмотрения подоб-

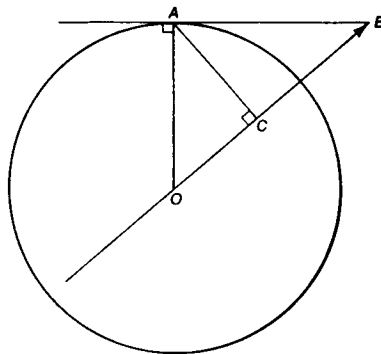


Рис. 21.2

Треугольники OAC и OBA подобны, следовательно, произведение $OC \times OB$ постоянно.

ных треугольников на рис. 21.2. Используя правило умножения на число, можно определить результат действия оператора $\mathcal{E}$ на векторы, которые лежат внутри окружности: сначала умножим вектор на число так, чтобы его конец оказался вне окружности, найдем ковектор, соответствующий новому вектору, затем полученный ковектор умножим на обратное число.

Условие (21.6) довольно трудно доказать, не выходя за рамки обычной геометрии, однако изложенная в следующем разделе техника расчета позволит нам это сделать. По-видимому, суть этого уравнения понятна и без строгого доказательства, которое не слишком поучительно. На рис. 21.3 мы изобразили операции, которые выполняются в левой части уравнения (21.6), а на рис. 21.4—операции, которые выполняются в правой части. Рисунки показывают, что уравнение (21.6) отнюдь не тривиально.

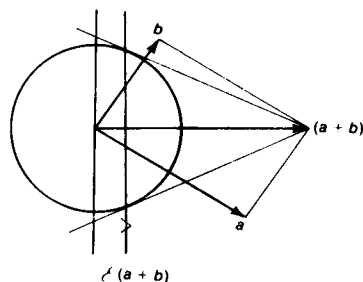


Рис. 21.3

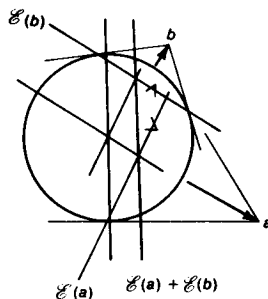


Рис. 21.4

Предположим, что нам задан некоторый тензор $\mathcal{B}$. С ним связан ряд других операторов. С помощью $\mathcal{B}$ мы всегда можем построить оператор, который действует в пространстве пар векторов.

Декартово произведение

Декартово произведение: пусть E и F — два векторных пространства; тогда произведением $E \times F$ назовем векторное пространство, элементами которого являются пары векторов, один вектор из E , другой из F . Операции умножения вектора на число и сложения векторов определены соотношениями

$$k(e, f) = (ke, kf), \quad (21.7)$$

$$(e_1, f_1) + (e_2, f_2) = (e_1 + e_2, f_1 + f_2). \quad (21.8)$$

[Здесь e — вектор из E , f — вектор из F , а k — действительное число.]

Плоскость $\mathbb{R}^2$ является декартовым произведением $\mathbb{R} \times \mathbb{R}$.

Ассоциированные тензоры

Обозначим этот оператор символом $\bar{\mathcal{B}}$:

$$\bar{\mathcal{B}}: V \times V \rightarrow \mathbb{R} \quad (21.9)$$

и определим его соотношением

$$\bar{\mathcal{B}}(a, b) \mapsto \mathcal{B}(a) \cdot b. \quad (21.10)$$

Здесь $\mathcal{B}(a)$ — ковектор, который действует на вектор b , в результате чего получается число. Если, наоборот, нам задан оператор типа $\bar{\mathcal{B}}$, то можно обратить процедуру и построить оператор типа $\mathcal{B}$. Определим оператор $\mathcal{B}$ следующим образом:

$$\mathcal{B}: V \rightarrow V^*; a \mapsto \bar{\mathcal{B}}(a, _). \quad (21.11)$$

В (21.11) символом $\bar{\mathcal{B}}(a, _)$ обозначен частично свернутый оператор. Величина $\bar{\mathcal{B}}(a, _)$ представляет собой ковектор! Почему? Потому что, если им подействовать на вектор (вставьте вектор на пустующее место), то мы получим число. Этот оператор линеен, если линеен оператор $\bar{\mathcal{B}}$.

Полилинейность

Линейность операторов, которые действуют в пространстве декартовых произведений, определяется просто. Потребуем, чтобы имела место линейность по каждому аргументу, т.е.

чтобы выполнялись равенства

$$\overline{\mathcal{B}}(a + b, c) = \overline{\mathcal{B}}(a, c) + \overline{\mathcal{B}}(b, c), \quad (21.12)$$

$$\overline{\mathcal{B}}(a, b + c) = \overline{\mathcal{B}}(a, b) + \overline{\mathcal{B}}(a, c). \quad (21.13)$$

Для данного оператора $\mathcal{B}: V \rightarrow V^*$ мы можем определить обратный оператор, т. е. линейный оператор $\mathcal{B}^{-1}: V^* \rightarrow V$, такой, что для всех векторов a справедливо равенство

$$\mathcal{B}^{-1}[\mathcal{B}(a)] = a. \quad (21.14)$$

Не у каждого тензора существует обратный ему тензор.

Мы будем называть линейные операторы типа $V \rightarrow V^*$ и все связанные с ними операторы, осуществляющие отображения $V \times V \rightarrow \mathbb{R}$, $V^* \rightarrow V$, $V^* \times V^* \rightarrow \mathbb{R}$, $V \rightarrow V$ и $V^* \rightarrow V^*$, тензорами. В большинстве учебников эти операторы называются тензорами второго ранга, векторы и ковекторы называются тензорами первого ранга, а числа — скалярами. Можно определить тензоры еще более общего вида, но лишь немногие из них нам понадобятся. При обсуждении кривизны мы упомянем мимоходом только один тензор высшего ранга — тензор Римана, который осуществляет отображение $V \times V \times V \rightarrow V$. Именно поэтому здесь для простоты используется термин «тензор» для обозначения тензора второго ранга.

Тензору $\mathcal{E}$, который мы ввели при рассмотрении евклидовой геометрии, ставится в соответствие тензор $\overline{\mathcal{E}}: V \times V \rightarrow \mathbb{R}$. Отличительная особенность этого тензора состоит в том, что для него справедливо равенство

$$\overline{\mathcal{E}}(a, b) = \overline{\mathcal{E}}(b, a). \quad (21.15)$$

Такие тензоры называются симметричными. Привычные для традиционной геометрии понятия расстояний и углов можно определить в произвольном векторном пространстве, если ввести в нем симметричный невырожденный тензор. Невырожденность означает, что существует обратный тензор.

В случае симметричного тензора нет необходимости указывать, по какому из аргументов производится частичное свертывание, поэтому в таких случаях мы будем использовать сокращенные обозначения:

$$\mathcal{E} \cdot a \equiv \overline{\mathcal{E}}(a, _). \quad (21.16)$$

Кроме описания метрических свойств пространства симметричные тензоры имеют множество других приложений, прежде

Тензоры второго ранга

Евклидова геометрия

[Это равенство легко доказать, воспользовавшись уравнениями (21.5), (21.6) и простыми геометрическими соображениями.]

Симметричные тензоры

Анизотропные среды

[К нашему сожалению, анизотропные среды не рассматриваются в обычных курсах, очевидно, именно потому, что для их описания необходимо вводить тензоры.]

Закон Гука

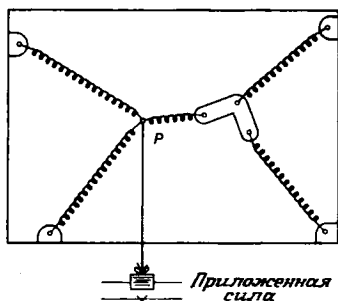


Рис. 21.5

Пример системы, для которой жесткость зависит от направления.

Пример 1

Рассмотрим силу F , приложенную в некоторой точке к достаточно сложной механической системе. Если реакция на силу не зависит от направления последней, то при малых деформациях связь между величиной приложенной силы и смещением d дается законом Гука

$$F = kd. \quad (21.17)$$

В более общем случае естественно ожидать, что упругость системы будет различной в различных направлениях. Можно, например, представить себе систему типа изображенной на рис. 21.5. Вспомним также, что сила является 1-формой. Сочетание этих двух фактов естественным образом приводит нас к понятию тензора упругости $\mathcal{K}$. Здесь мы даем только графическое представление тензора $\mathcal{K}$. В следующем разделе будет рассмотрено аналитическое и численное представление тензоров.

Тензор упругости мы будем изображать в виде эллипса; при этом 1-форма силы и вектор смещения оказываются связанными посредством процедуры, представляющей собой прямое обобщение правила, сформулированного для метрических тензоров (рис. 21.6). Совсем не очевидно, что $\mathcal{K}$ — действительно линейный оператор. Легче всего это показать, если воспользоваться аналитическим методом, описанным в разд. 22.

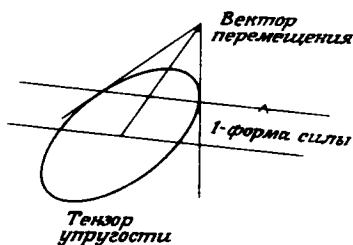


Рис. 21.6

Применение эллипса, который представляет тензор упругости. Для малых сил такое построение неприменимо. В этом случае следует воспользоваться свойством линейности: умножить вектор на число и перейти к рассмотренному случаю, а затем умножить вектор смещения на обратное число.

Пример 2

В качестве еще одного примера задачи, которая естественным образом приводит нас к понятию тензора, рассмотрим распространение тока (вектор) в анизотропном проводнике, например в кристалле, под действием градиента электростатического потенциала (1-форма). Теперь вас уже не удивит, что электрические свойства кристалла описываются тензором проводимости $\mathcal{R}$. Соответствующее правило, которое можно понять, обратившись к рис. 21.7, аналогично правилу, сформулированному при рассмотрении предыдущего примера. Так как мы работаем в евклидовом пространстве, длина вектора тока непосредственно дает плотность тока в амперах на квадратный метр, а направление вектора совпадает с направлением тока.

Тензор проводимости

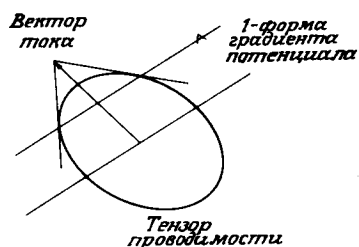


Рис. 21.7

Кратко охарактеризуем некоторые типичные ситуации, при описании которых вводится понятие тензора проводимости. Можно выделить два простейших случая. В первом случае рассматривается тонкая кристаллическая пластинка, на обе поверхности которой наложены проводящие электроды (рис. 21.8). Если проводимость электродов много больше проводимости кристалла, то эквипотенциальные поверхности представляют собой плоскости, параллельные плоскости электродов. Способ нахождения тока при заданном тензоре проводимости изображен на рис. 21.9. Ток имеет две компоненты. Символом

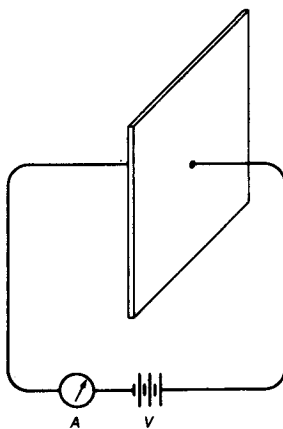


Рис. 21.8
Измерение проводимости тонкой пластинки.

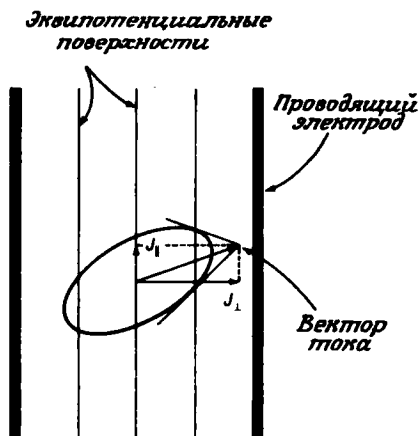


Рис. 21.9

Распределение потенциала и тока в проводящей пластинке. Как показано на рисунке, разность потенциалов между электродами равна 4. Отношение поперечной плотности тока к продольной равно примерно $2\frac{3}{4}$.

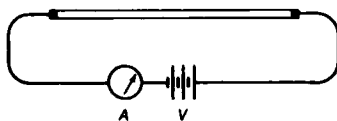


Рис. 21.10

Измерение проводимости длинного тонкого стержня.

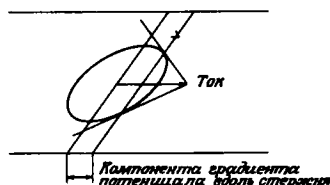


Рис. 21.11

Распределение потенциала и тока в стержне, изображенном на рис. 21.10. Падение напряжения на представленном участке проводника равно примерно 10.

$J_{\perp}$ на рисунке обозначена компонента, перпендикулярная пластинкам. Этот ток измеряется амперметром, изображенным на рис. 21.8. Компонента тока $J_{\parallel}$ соответствует переносу заряда вдоль пластин. Ее явный вид зависит от граничных условий. Величина тока $J_{\parallel}$ пропорциональна не площади поверхности образца, а произведению площади на толщину пластины. В случае тонкой пластины током $J_{\parallel}$ можно пренебречь.

В противоположном случае мы имеем дело с тонким длинным стержнем, изображенным на рис. 21.10. Теперь нам известен ток, но мы ничего не знаем о том, как выглядят эквипотенциальные поверхности. На рис. 21.11 изображена ситуация, когда стержень вырезан из того же материала, что и пластинка, в направлении, перпендикулярном плоскости пластины. Для заданного тока 1-форма градиента потенциала должна иметь изображенный на рисунке вид. И опять граничные условия на концах образца зависят от того, каким способом подсоединены электроды. Для длинного тонкого образца влиянием условий на концах можно пренебречь. Градиент потенциала создает разность потенциалов на концах образца, которая зависит от длины образца и величины компоненты градиента вдоль образца.

Две рассмотренные экспериментальные ситуации взаимно дополняют друг друга. Первый случай соответствует материа-

лам с низкой проводимостью, второй — хорошим проводником электричества.

ЗАДАЧИ

21.1. (14) Дайте геометрическое доказательство справедливости равенства (21.5).

21.2. (26) На рис. 21.12 тензор $\mathcal{E}$ действует на вектор, который оканчивается в точке с координатами (x, y) ; в результате мы получаем 1-форму, которая задается двумя числами l и m . Покажите, что

$$lx = 1,$$

$$my = 1.$$

21.3. (13) Используя результаты задач 15.7 и 21.2, докажите, что оператор $\mathcal{E}$ линеен.

21.4. (34) Повторите описанную в настоящем разделе процедуру, используя вместо окружности $\mathcal{E}$ гиперболу. Покажите, что в таком случае мы также имеем дело с тензором. Удобно использовать гиперболу

$$xy = 1.$$

Еще раз внимательно просмотрите задачи 21.1—21.3.

21.5. (18) Многие кривые можно использовать как операторы, которые превращают векторы в ковекторы, осуществляя для этого описанную выше процедуру. Покажите, что оператор, определяемый параболой

$$y = x^2,$$

не является тензором.

21.6. (20) Рассмотрите отображение векторов в ковекторы

$$\mathcal{D}: v \mapsto \mathcal{D}(v),$$

которое изображено на рис. 21.13. Здесь $\mathcal{D}(v)$ представляет собой набор трех параллельных линий, а величины a , b и c удовлетворяют соотношению

$$ab = c^2.$$

Покажите, что это отображение представляет собой тензор. Является ли этот тензор симметричным?

21.7. (34) Сравните проводимость, т.е. отношение полного то-

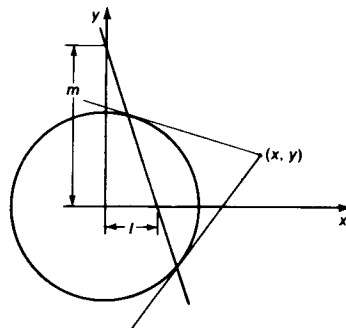


Рис. 21.12

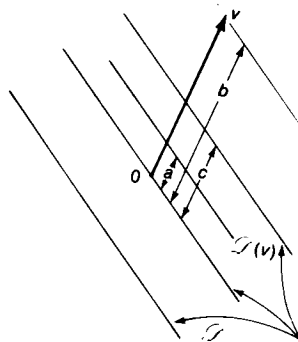


Рис. 21.13

$\mathcal{D}$ представляется тремя заданными параллельными линиями.

[Эта задача очень важна. Отнеситесь к ней внимательно.]

ка к разности потенциалов, для двух пластинок, если одна из них вырезана горизонтально, а другая вертикально из материала, тензору проводимости которого соответствует эллипс на рис. 21.7.

21.8. (17) Как выглядит кривая, описываемая тензором проводимости для системы проводящих стержней, пространство между которыми заполнено диэлектриком?

21.9. (11) Почему отображение

$$V \rightarrow V; a \mapsto a + b,$$

где b — заданный вектор, не является тензором?

22. Базисные векторы и тензорное произведение

*Хорошенькое дельце, — говаривал он,
— и такое же занимательное, как алгебра.*

Н. Дуглас

Тензорная алгебра основана на использовании свойства линейности. Линейность пространства касательных векторов следует из определения касательного вектора посредством процедуры предельного перехода. Тензор представляет собой линейный оператор, что должно быть ясно из его определения и графических примеров, рассмотренных в предыдущем разделе. Здесь, как и при рассмотрении ковекторов, нам необходимо, с одной стороны, развивать интуицию, а с другой — научиться практически работать с тензорами. Для работы с тензорами необходимо использовать их численное представление. Тензоры образуют линейное векторное пространство, т. е. их можно складывать и умножать на числа. Как и в любом другом векторном пространстве, для перехода к численному представлению нужно выбрать совокупность базисных элементов и записать все тензоры в виде линейной комбинации базисных тензоров. Мы построим базис нашего пространства тензоров с помощью базисных векторов исходного векторного пространства. При этом оказывается, что базис исходного линейного пространства не только естественным образом определяет базис в пространстве ковекторов, но и позволяет построить базис для различных тензорных пространств.

Для построения базисных тензоров введем новое правило умножения, так называемое тензорное произведение. Эта опе-

[Вас может смутить используемый здесь язык. К сожалению, это общепринятая терминология. Не только векторы являются тензорами, но и, как вы скоро увидите, тензоры являются векторами.]

Базисные тензоры

Тензорное произведение

рация позволит нам, например, взять два ковектора ω и ν и «перемножить» их таким образом, чтобы в результате получился тензор $\omega \otimes \nu$. Как оператор произведение $\omega \otimes \nu$, представляющее собой тензор, определяется следующим образом:

$$\omega \otimes \nu: V \times V \rightarrow \mathbb{R}; (a, b) \mapsto (\omega \cdot a)(\nu \cdot b). \quad (22.1)$$

Здесь a и b — векторы, а $(\omega \cdot a)$ и $(\nu \cdot b)$ — числа, перемножение которых дает окончательный результат. Оператор (22.1) представляет собой тензор, если он линейен. Последнее легко доказать. Действительно, ω и ν — ковекторы и, следовательно, линейны, поэтому их тензорное произведение $\omega \otimes \nu$ также будет линейным оператором.

[Выражение $\omega \otimes \nu$ следует читать как «тензорное произведение омега на ню».]

[Точкой обозначена операция свертывания ковектора с вектором, определенная в разд. 16. Крестиком обозначено декартово произведение, которое мы ввели в предыдущем разделе.]

Пример 1

Предположим, что

$$\begin{aligned} \omega &= dx + dy, \\ \nu &= dx - dy, \\ a &= \frac{\partial}{\partial x}, \\ b &= \frac{\partial}{\partial y}; \end{aligned} \quad (22.2)$$

[Эти базисные векторы и ковекторы были определены в разд. 18.]

тогда

$$\omega \otimes \nu \cdot (a, b) = (\omega \cdot a)(\nu \cdot b) = -1. \quad (22.3)$$

Тензорное произведение можно подвергнуть частичному свертыванию. В результате мы найдем, что величина $\omega \otimes \nu(a, -)$ является ковектором; точно так же ковектор получается в результате свертывания $\omega \otimes \nu(-, a)$. Заметим, что эти два свертывания не обязательно дают одинаковый результат:

[Здесь мы, следуя общепринятой традиции, обозначаем точкой операцию свертывания тензора с элементом векторного пространства, на котором он определен.]

$$\begin{aligned} \omega \otimes \nu \cdot (a, _) &= dx - dy, \\ \omega \otimes \nu \cdot (b, _) &= dx - dy, \\ \omega \otimes \nu \cdot (_, a) &= dx + dy, \\ \omega \otimes \nu \cdot (_, b) &= -dx - dy. \end{aligned} \quad (22.4) \quad \text{[Продолжение примера 1.]}$$

Другие тензоры могут быть построены путем перемножения двух векторов или вектора и ковектора. Соответствующие правила достаточно очевидны.

Пример

$$\frac{\partial}{\partial x} \otimes dx \cdot \left(dx + dy, 3 \frac{\partial}{\partial x} - \frac{\partial}{\partial y} \right) = 3, \quad (22.5)$$

$$\frac{\partial}{\partial x} \otimes dy \cdot (3 dx - dy, _) = 3 dy, \quad (22.6)$$

$$\frac{\partial}{\partial y} \otimes dx \cdot (3 dx - dy, _) = -dx, \quad (22.7)$$

$$\frac{\partial}{\partial y} \otimes dx \cdot (_, dy) = \text{TILT!} \quad (22.8)$$

Последняя операция не имеет смысла. Мы не можем определить операцию действия ковектора на ковектор, пока не введем специальный тензорный оператор.

Пространство векторов типа $dx \otimes dx$ обозначается символом $V^* \otimes V^*$. В рассмотренном выше примере мы имели дело с тензорами, которые являются элементами пространства $V \otimes V^*$.

Пространство $V \otimes V^*$ знакомо физикам. Элементами такого пространства являются произведения векторов-столбцов на векторы-строки, т. е. матрицы, например

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} (1 \ 1 \ 0) = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (22.9)$$

а также операторы, представляющие собой произведения векторов состояний квантовой системы вида

$$|\psi\rangle \langle \psi'|.$$

Линейная структура пространства тензоров

Тензоры, полученные с помощью тензорного произведения, порождают новое векторное пространство. Это не значит, что они сами по себе образуют векторное пространство. Не каждый вектор, принадлежащий, например, пространству $V^* \otimes V^*$, можно записать как произведение двух ковекторов. Это очевидно для матриц, так как не каждую матрицу можно представить как тензорное произведение двух векторов; существует n^2 различных матриц $n \times n$ и только $2n$ различных пар векторов. Однако любой тензор можно записать в виде суммы таких тензорных произведений; именно это мы имели в виду, когда говорили, что тензорное произведение порождает

новое векторное пространство. Правила умножения на число и сложения для тензорных произведений отличаются от соответствующих правил для декартова произведения, приведенных в предыдущем разделе. Умножение на число и сложение удовлетворяют следующим правилам:

$$k(\omega \otimes \nu) = (k\omega) \otimes \nu = \omega \otimes (k\nu), \quad (22.10)$$

$$\omega \otimes (\alpha + \beta) = \omega \otimes \alpha + \omega \otimes \beta. \quad (22.11)$$

Они вытекают из требования линейности и введенных выше определений.

Пример

Чтобы вывести закон умножения на число, заметим, что в соответствии с условием линейности должно выполняться равенство

$$[k(\omega \otimes \nu) \cdot (a, b) = k[\omega \otimes \nu \cdot (a, b)]], \quad (22.12)$$

откуда, согласно нашему определению, имеем

$$\begin{aligned} &= k[(\omega \cdot a)(\nu \cdot b)] = \\ &= (k\omega \cdot a)(\nu \cdot b) = \\ &= (\omega \cdot a)(k\nu \cdot b). \end{aligned} \quad (22.13)$$

Эти правила делают ненужным употребление скобок во многих приведенных выше выражениях.

Теперь мы в состоянии построить базис для любого тензорного пространства. В случае двух измерений, когда базис пространства ковекторов состоит из двух элементов dx и dy , базис пространства $V^* \otimes V^*$ образуется четырьмя тензорами:

$$\begin{aligned} &dx \otimes dx, \\ &dx \otimes dy, \\ &dy \otimes dx, \\ &dy \otimes dy. \end{aligned} \quad (22.14)$$

Пространство $V \otimes V^*$ имеет следующий базис:

$$\begin{aligned} &\frac{\partial}{\partial x} \otimes dx, \\ &\frac{\partial}{\partial x} \otimes dy, \\ &\frac{\partial}{\partial y} \otimes dx, \\ &\frac{\partial}{\partial y} \otimes dy \end{aligned} \quad (22.15)$$

Координатный базис

[Симметричными называются тензоры, для которых справедливо равенство (21.15).]

и т. д. Базис подпространства симметричных тензоров, принадлежащих пространству $V^* \otimes V^*$, состоит из трех элементов

$$\begin{aligned} dx \otimes dx, \\ dx \otimes dy + dy \otimes dx, \\ dy \otimes dy. \end{aligned} \quad (22.16)$$

Подпространство антисимметричных тензоров пространства $V^* \otimes V^*$ одномерно; соответствующий базисный вектор имеет вид

$$dx \otimes dy - dy \otimes dx. \quad (22.17)$$

[Этот пример показывает, каким образом графические построения можно описать алгебраически.]

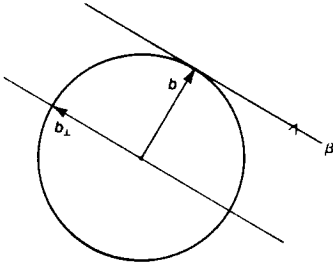


Рис. 22.1

[По поводу симметричности тензора $\mathcal{E}$ в форме (22.18.) см. задачу 22.10.]

Пример

Убедимся, что метрический тензор евклидова пространства можно записать в рассматриваемом базисе в виде

$$\mathcal{E} = dx \otimes dx + dy \otimes dy. \quad (22.18)$$

По построению это, очевидно, линейный оператор. Мы должны показать, что он эквивалентен графическому представлению, которое приводилось выше.

Прежде всего заметим, что вектор

$$b = \cos \theta \frac{\partial}{\partial x} + \sin \theta \frac{\partial}{\partial y} \quad (22.19)$$

лежит на единичной окружности для любых значений параметра θ . Заметим также, что вектор

$$b_{\perp} \equiv -\sin \theta \frac{\partial}{\partial x} + \cos \theta \frac{\partial}{\partial y} \quad (22.20)$$

перпендикулярен вектору b . Линия уровня 1-формы β совпадает с касательной к окружности в точке b , если выполняются равенства (рис. 22.1).

$$\beta \cdot b = 1, \quad (22.21)$$

$$\beta \cdot b_{\perp} = 0. \quad (22.22)$$

Эти требования определяют β однозначно:

$$\beta = \cos \theta dx + \sin \theta dy. \quad (22.23)$$

Теперь возьмем произвольный вектор a , лежащий вдоль оси x . Благодаря симметрии задачи это возможно всегда. Пока-

жем, что описанный в разд. 21 графический способ построения 1-формы по заданному вектору a согласуется с результатом выполнения операции $\mathcal{E} \cdot a$.

Пусть

$$a = k \frac{\partial}{\partial x}, \quad (22.24)$$

[Сокращенное обозначение для операции частичного свертывания было введено на стр. 165.]

где $k \geq 1$.

При графическом построении мы должны провести из конца вектора a касательные к единичной окружности. Эти касательные можно рассматривать как линии единичного уровня для двух 1-форм $\beta_{\pm}$, имеющих вид (22.23), для которых справедливо равенство

$$\beta_{\pm} \cdot a = 1. \quad (22.25)$$

Если мы выберем θ таким образом, что

$$k \cos \theta = 1, \quad (22.26)$$

то интересующие нас 1-формы примут вид

$$\beta_{\pm} = \cos \theta dx \pm \sin \theta dy. \quad (22.27)$$

Они касаются единичной окружности в точках с радиус-векторами

$$b_{\pm} = \cos \theta \frac{\partial}{\partial x} \pm \sin \theta \frac{\partial}{\partial y}, \quad (22.28)$$

как изображено на рис. 22.2. Наконец, построим 1-форму α , та-

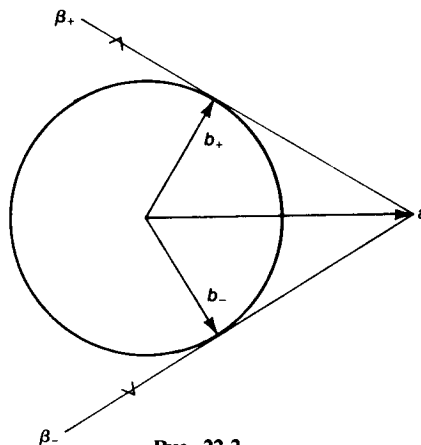


Рис. 22.2

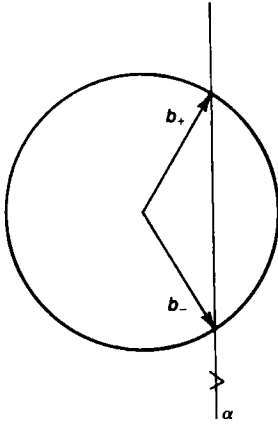


Рис. 22.3

кую, что

$$\alpha \cdot b_{\pm} = 1. \quad (22.29)$$

Сделаем это так, как показано на рис. 22.3. Если взять α в виде

$$\alpha = k dx, \quad (22.30)$$

то мы, очевидно, получим

$$\alpha \cdot b_{\pm} = k \cos \theta = 1. \quad (22.31)$$

Таким образом, наше графическое построение осуществляет отображение

$$k \frac{\partial}{\partial x} \mapsto k dx. \quad (22.32)$$

Алгебраический расчет дает

$$\begin{aligned} \mathcal{E} \cdot a &= (dx \otimes dx + dy \otimes dy) \cdot k \frac{\partial}{\partial x} = \\ &= dx \otimes dx \cdot k \frac{\partial}{\partial x} + dy \otimes dy \cdot k \frac{\partial}{\partial x} = \\ &= k \left(dx \cdot \frac{\partial}{\partial x} \right) dx + k \left(dy \cdot \frac{\partial}{\partial x} \right) dy = \\ &= k dx. \end{aligned} \quad (22.33)$$

Полученный результат совпадает с (22.32).

Часто используют сокращенную форму записи для метрических тензоров, т. е. симметричных тензоров, принадлежащих пространству $V^* \otimes V^*$, при которой знак $\otimes$ опускается. Метрику евклидова пространства записывают в виде

$$\mathcal{E} = dx^2 + dy^2, \quad (22.34)$$

понимая под этим, что

$$\mathcal{E} = dx \otimes dx + dy \otimes dy. \quad (22.35)$$

Определенные в настоящем разделе тензоры могут действовать не только на пары векторов, но и на тензоры. Это станет ясно, когда мы рассмотрим следующий пример.

Пример

$$\mathcal{E} \cdot \left(2 \frac{\partial}{\partial x} \otimes \frac{\partial}{\partial x} \right) = dx \otimes dx \cdot \left(2 \frac{\partial}{\partial x} \otimes \frac{\partial}{\partial x} \right) + 0 =$$

$$\begin{aligned}
 &= 2 \left(dx \cdot \frac{\partial}{\partial x} \right) \left(dx \cdot \frac{\partial}{\partial x} \right) = \\
 &= 2.
 \end{aligned}
 \tag{22.36}$$

ЗАДАЧИ

22.1. (7) Изобразите графически тензор

$$dx \otimes dx + 2 dy \otimes dy.$$

22.2. (10) Изобразите графически тензор

$$dx \otimes dy + dy \otimes dx.$$

22.3. (13) Найдите результат свертывания $\mathcal{F} \cdot (\partial/\partial x)$ и $\mathcal{A}(\partial/\partial y)$, если тензор $\mathcal{F}$ имеет вид

$$\mathcal{F} = dx \otimes dx + dx \otimes dy + dy \otimes dx.$$

22.4. (14) Найдите тензор, обратный приведенному выше тензору.

22.5. (7) Строго покажите, что тензор

$$dx \otimes dy + dy \otimes dx$$

симметричен.

22.6. (14) Найдите тензор, обратный метрическому тензору евклидова пространства.

22.7. (39) Придумайте графическое представление для антисимметричных тензоров в пространстве двух, трех и четырех измерений.

22.8. (13) Дайте численное представление тензора проводимости, изображенного на рис. 21.7.

22.9. (10) Используя (22.18), покажите, что тензор $\mathcal{E}$ симметричен.

22.10. (16) Покажите, что вид тензора $\mathcal{E}$ не меняется при поворотах базисных векторов

$$\begin{aligned}
 dx' &= \cos \phi \, dx + \sin \phi \, dy, \\
 dy' &= -\sin \phi \, dx + \cos \phi \, dy.
 \end{aligned}$$

23. Пространство Минковского

Теперь мы можем воспользоваться преимуществами, которые нам дает тензорная алгебра. Наконец, можно раскрыть мате-

Метрический тензор специальной теории относительности

[Мы не будем заменять символ $\mathcal{E}$, введенный для хроноструктуры события, общепринятым обозначением ds^2 . Символ d имеет совсем другой смысл, чем тот, который ему приписывается в выражении ds^2 .]

Метрические фигуры

[Выражение (23.6) показывает, что вектор v является касательным к линии $x = x_0 s$, $t = t_0 s$, которая при $s = 1$ проходит через точку с координатами (t_0, x_0) .]

матический смысл графических построений, которые мы проводили при изучении специальной теории относительности. Он оказывается простым. Математическая структура специальной теории относительности определяется метрическим тензором $\mathcal{N}$, имеющим вид

$$\mathcal{N} = -dt^2 + dx^2 = -dt \otimes dt + dx \otimes dx \quad (23.1)$$

или в случае четырех измерений

$$\mathcal{N} = -dt^2 + dx^2 + dy^2 + dz^2. \quad (23.2)$$

Такой подход позволяет легко обобщить специальную теорию относительности, чтобы включить в рассмотрение гравитацию. Сейчас мы покажем, каким образом этот метрический тензор связан с нашей гиперболой, что позволит нам дать простую интерпретацию рассмотренной в разд. 13 формулы для доплеровского сдвига.

В разд. 22 было показано, что метрический тензор евклидова пространства может быть изображен единичной окружностью, т. е. множеством векторов v , которые удовлетворяют соотношению

$$\mathcal{E}(v, v) = 1. \quad (23.3)$$

Аналогичное представление возможно для любого симметричного тензора. В случае метрического тензора $\mathcal{N}$ мы имеем два различных семейства единичных векторов

$$\mathcal{N} \cdot (v, v) = +1, \quad (23.4)$$

$$\mathcal{N} \cdot (v, v) = -1. \quad (23.5)$$

Отрицательная ветвь изображается знакомой нам гиперболой. Действительно, вектор, выходящий из начала координат и оканчивающийся в точке с координатами (t_0, x_0) , имеет вид

$$v = t_0 \frac{\partial}{\partial t} + x_0 \frac{\partial}{\partial x}, \quad (23.6)$$

и мы получаем

$$\mathcal{N} \cdot (v, v) = -t_0^2 + x_0^2. \quad (23.7)$$

Поэтому множество векторов, удовлетворяющих уравнению (23.5), которое задает времениподобную гиперболу, определяется уравнением

$$t_0^2 - x_0^2 = 1. \quad (23.8)$$

Это и есть знакомая нам гипербола.

Метрический тензор $\mathcal{N}$ осуществляет также отображение

векторов в ковекторы. Определим ковектор α соотношением

$$\alpha \equiv \mathcal{N} \cdot a. \quad (23.9)$$

Тогда, если вектор a удовлетворяет уравнению (23.5), получаем

$$\alpha \cdot a = -1. \quad (23.10)$$

Рассмотрим вектор $a + \varepsilon$, где ε — малый вектор, квадратом которого в пределе $\varepsilon \rightarrow 0$ можно пренебречь. Если вектор $a + \varepsilon$ также удовлетворяет уравнению (23.5), т. е.

$$\mathcal{N} \cdot (a + \varepsilon, a + \varepsilon) = -1, \quad (23.11)$$

то, используя свойство линейности, получаем

$$-1 + \mathcal{N} \cdot (a, \varepsilon) + \mathcal{N} \cdot (\varepsilon, a) = -1. \quad (23.12)$$

Тензор $\mathcal{N}$ симметричен, т. е.

$$\mathcal{N} \cdot (a, \varepsilon) = \mathcal{N} \cdot (\varepsilon, a); \quad (23.13)$$

следовательно, имеем

$$\mathcal{N} \cdot (a, \varepsilon) = 0, \quad (23.14)$$

т. е.

$$\alpha \cdot \varepsilon = 0. \quad (23.15)$$

По построению вектор ε является касательным вектором к гиперболе, и из уравнения (23.15) следует, что линии уровня ковектора α параллельны ε . Более того, уравнение (23.10) показывает, что линия уровня ковектора α , соответствующая его значению, равному -1 , проходит через конец вектора a . Обратите внимание на знак минус! Отображение вектора a в ковектор, осуществляемое тензором $\mathcal{N}$, показано на рис. 23.1. Результат действия тензора $\mathcal{N}$ на вектор, имеющий то же направление, но другую длину, можно найти с помощью операции умножения на число.

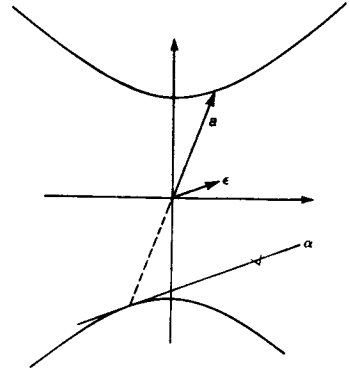


Рис. 23.1

Пример

$$\mathcal{N} \cdot \frac{\partial}{\partial t} = -dt. \quad (23.16)$$

Пространственноподобные векторы никогда не пересекают времениподобную гиперболу. Они должны пересекать пространственноподобную гиперболу

$$x_0^2 - t_0^2 = 1. \quad (23.17)$$

Процедура отображения векторов в ковекторы с помощью пространственноподобной гиперболы аналогична описанной выше

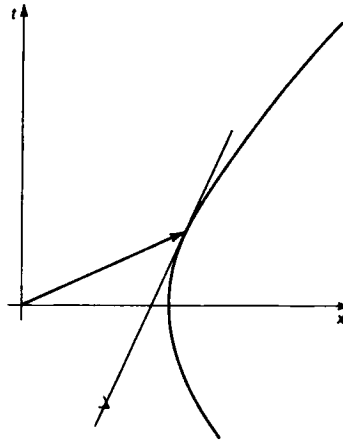


Рис. 23.2

(рис. 23.2). Обратите внимание на другой знак.

Пример

$$\mathcal{N} \cdot \frac{\partial}{\partial x} = dx. \quad (23.18)$$

Четырехмерное скалярное произведение нетрудно записать через метрический тензор

$$\mathcal{N} \cdot (a, b) = a \cdot b. \quad (23.19)$$

[Обратите внимание на то, что жирная точка в правой части равенства обозначает четырехмерное скалярное произведение.]

Доказательство

Возьмем два произвольных вектора

$$a = a^t \frac{\partial}{\partial t} + a^x \frac{\partial}{\partial x} + \dots, \quad (23.20)$$

$$b = b^t \frac{\partial}{\partial t} + \dots; \quad (23.21)$$

тогда

$$\begin{aligned} \mathcal{N} \cdot (a, b) &= (-dt \circ dt + \dots) \cdot (a^t \frac{\partial}{\partial t} + \dots, b^t \frac{\partial}{\partial t} + \dots) = \\ &= -a^t b^t + a^x b^x + \dots \end{aligned} \quad (23.22)$$

[Обратите внимание на символ «...», который обозначает «каждому ясно, что следует написать дальше».]

Геометрия пространства-времени специальной теории относительности тесно связана со структурой волнового уравнения Волновое уравнение

$$\frac{\partial^2 \psi}{\partial t^2} = \frac{\partial^2 \psi}{\partial x^2}. \quad (23.23)$$

Введение метрического тензора позволяет записать соответствующее этому уравнению дисперсионное уравнение

$$\omega^2 = k^2. \quad (23.24)$$

Для метрического тензора $\mathcal{N}$ существует обратный тензор $\mathcal{N}^{-1}$, который имеет вид

$$\mathcal{N}^{-1} = -\frac{\partial}{\partial t} \otimes \frac{\partial}{\partial t} + \frac{\partial}{\partial x} \otimes \frac{\partial}{\partial x}. \quad (23.25)$$

Доказательство

Возьмем произвольный вектор

$$a = u \frac{\partial}{\partial t} + v \frac{\partial}{\partial x}. \quad (23.26)$$

Тогда можно написать

$$\mathcal{N} \cdot \left(u \frac{\partial}{\partial t} + v \frac{\partial}{\partial x} \right) = -u dt + v dx, \quad (23.27)$$

следовательно,

$$\begin{aligned} \mathcal{N}^{-1} \cdot (-u dt + v dx) &= \\ &= -\frac{\partial}{\partial t} \otimes \frac{\partial}{\partial t} \cdot (-u dt) + \frac{\partial}{\partial x} \otimes \frac{\partial}{\partial x} \cdot (v dx) = \\ &= u \frac{\partial}{\partial t} + v \frac{\partial}{\partial x}. \end{aligned} \quad (23.28)$$

Таким образом, тензор $\mathcal{N}^{-1}$ действительно представляет собой тензор, обратный тензору $\mathcal{N}$.

Используя тензор $\mathcal{N}^{-1}$, мы можем записать дисперсионное уравнение (23.24) в виде

$$\mathcal{N}^{-1} \cdot (d\theta, d\theta) = 0. \quad (23.29) \quad \text{Дисперсионное уравнение}$$

Уравнение (23.29) устанавливает связь между метрическим тензором $\mathcal{N}$ специальной теории относительности и волновым уравнением. Можно также показать, что вектор групповой ско-

[Отложим доказательство этого равенства до разд. 27.]

рости σ и градиент фазы $d\theta$ связаны соотношением

$$\sigma = \mathcal{N}^{-1} \cdot d\theta. \quad (23.30)$$

Для случаев двух и трех пространственных измерений это соотношение графически представлено соответственно на рис. 23.3 и 23.4.

Электромагнетизм

Волновое уравнение, используемое в электродинамике, представляет собой обобщение приведенного выше скалярного уравнения на случай векторного поля. Соответствующее ему дисперсионное уравнение имеет тот же самый вид. Поэтому все сказанное выше применимо и к световым волнам. В однородном пространстве-времени специальной теории относительности распространение волны не сопровождается изменением ее частоты ω и волнового числа k . Однако, исходя из вышеизложенного, нетрудно понять, как возникает эффект доплеровского сдвига (см. разд. 13).

Доплеровское смещение

[См. рис. 13.5.]

Конфигурация гребней волн определяется 1-формой градиента фазы $d\theta$. Наблюдаемая частота — это скорость, с которой гребни волн проходят мимо данного наблюдателя (рис. 23.5). Для наблюдателя с 4-скоростью λ эта частота дается выражением $(d\theta \cdot \lambda)/2\pi$, а отношение частот для двух наблюдателей, движущихся с различными скоростями, имеет вид

$$\frac{d\theta \cdot \lambda}{d\theta \cdot \lambda'}. \quad (23.31)$$

Используя равенство

$$d\theta = \mathcal{N} \cdot \sigma, \quad (23.32)$$

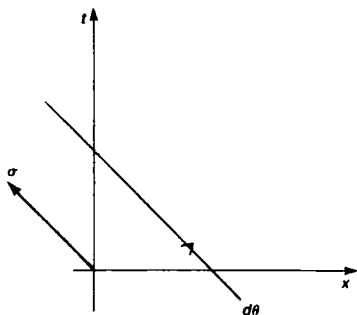


Рис. 23.3

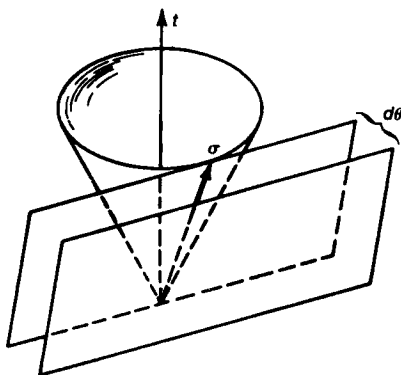


Рис. 23.4

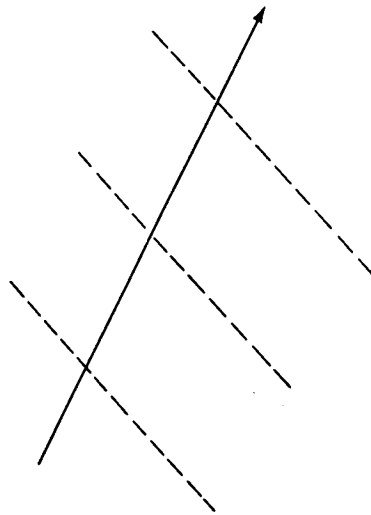


Рис. 23.5

обратное равенству (23.30), для отношения частот получаем

$$\frac{\mathcal{N} \cdot (\sigma, \lambda)}{\mathcal{N} \cdot (\sigma, \lambda')} = \frac{\sigma \cdot \lambda}{\sigma \cdot \lambda'}. \quad (23.33)$$

Таким образом, мы пришли к нашему старому выражению.

Мы осуществили нашу программу и разработали математический аппарат, который описывает различные хроноструктуры. В общем случае они описываются дисперсионными уравнениями. В более простых случаях можно ввести тензоры и построить более изящные теории. Именно так мы действовали при рассмотрении специальной теории относительности и соответствующего волнового уравнения. Четырехмерное векторное пространство, наделенное метрикой $\mathcal{N}$, называется пространством Минковского.

[Обратите внимание, что в правой части этого равенства точка обозначает четырехмерное скалярное произведение.]

ЗАДАЧА

23.1. (31) Один из возможных способов построения гиперболы изображен на рис. 23.6. Проведем окружность и возьмем некоторую точку P вне нее. Проведем прямую из точки P до пересечения с окружностью, а через точку пересечения проведем прямую L , перпендикулярную исходной прямой. Как показано

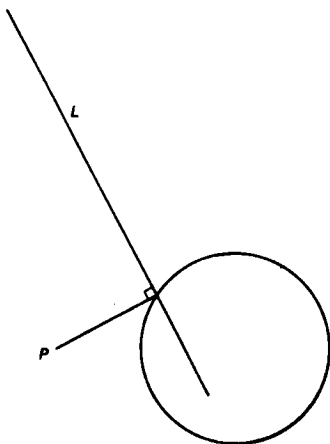


Рис. 23.6

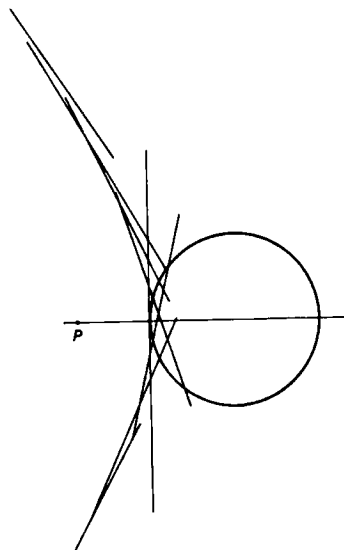


Рис. 23.7

[Как ни удивительно, с построением огибающих мы сталкиваемся чрезвычайно часто.]

на рис. 23.7, все такие линии L будут касательными к гиперболе, т. е. гипербола является огибающей нашего семейства прямых. Будем рассматривать эту гиперболу как изображающую метрический тензор $\mathcal{N}$ пространства Минковского, а линии L — как изображения 1-форм. Покажите, что описанная процедура действительно приводит к нужному результату.

24. Индексные обозначения

Такие люди, как Фарадей, интуитивно чувствовали, что такое вектор. Они могли ничего не знать о правилах действия с ними, но если они думали о векторах, они представляли их себе как векторы, т. е. как направленные величины. Ни один такой человек не мог, да и не стал бы представлять себе три компоненты вектора по отдельности и как не связанные друг с другом.

О. Хевисайд

До сих пор мы поступали в духе приведенного в эпиграфе высказывания и всячески подчеркивали внутренние геометрические свойства тензоров. Определение тензоров как линейных

операторов и полученные нами графические представления не зависели от выбора системы координат. Однако при выполнении конкретных расчетов и доказательств часто более эффективными оказываются индексные обозначения, основанные на координатном представлении. Мы начали с изучения внутренних свойств тензоров с тем, чтобы впоследствии не путать свойства представления, такие, как ковариантность, со свойствами самих тензоров. В старых учебниках по тензорному анализу основное внимание уделялось трансформационным свойствам тензоров. Эти свойства будут играть определенную роль при обсуждении свойств симметрии волн на воде в разд. 28, но для нас они имеют второстепенное значение.

Произвольный тензор можно выразить через построенные в разд. 23 базисные тензоры двумя различными способами. Один из способов заключается в точном перечне всех ненулевых компонент тензора. При таком подходе метрический тензор пространства-времени $\mathcal{N}$ имел бы следующий вид:

$$\mathcal{N} = -dt \otimes dt + dx \otimes dx. \quad (24.1)$$

При индексных обозначениях для представления тензора используется все множество его компонент. Если в выражении

$$\mathcal{N} = \eta_{tt} dt \otimes dt + \eta_{tx} dt \otimes dx + \eta_{xt} dx \otimes dt + \eta_{xx} dx \otimes dx, \quad (24.2)$$

символами η_{tt} , η_{tx} , η_{xt} и η_{xx} обозначить компоненты тензора $\mathcal{N}$, то тензор $\mathcal{N}$ мы можем записать в виде матрицы

$$\begin{pmatrix} \eta_{tt} & \eta_{tx} \\ \eta_{xt} & \eta_{xx} \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (24.3)$$

Способ, основанный на перечислении всех ненулевых компонент, более удобен при работе с векторами и тензорами, у которых мало отличных от нуля компонент, и им мы будем преимущественно пользоваться. Можно было бы ограничиться этим способом, однако индексные обозначения широко распространены и в некоторых случаях более эффективны, поэтому студенты, изучающие теорию относительности, должны их знать тоже.

В индексных обозначениях вектор или тензор представляется в виде разложения по координатному базису, где верхние и нижние индексы используются для обозначения операций сум-

[Традиционное изложение тензорного анализа можно найти в превосходной книге [37].]

Представления

Явное выражение

Таблица

Верхние и нижние индексы

мирования. Положение индекса указывает тип вектора или тензора. Так, касательные векторы и их тензорные произведения представляются компонентами с верхними индексами. Компоненты ковекторов и их тензорных произведений имеют нижние индексы.

Пример

Разложение касательного вектора λ будет записываться в виде

$$\lambda = \lambda^t \frac{\partial}{\partial t} + \lambda^x \frac{\partial}{\partial x} + \dots, \quad (24.4)$$

а разложение ковектора α — в виде

$$\alpha = \alpha_t dt + \alpha_x dx + \dots \quad (24.5)$$

Немые индексы

При записи таких сумм будем пользоваться *немymi индексами*. Пространственно-временные индексы будут обозначаться греческими буквами, причем часто без точного указания значений, которые пробегает индекс. Допуская некоторую вольность, будем впредь обозначать базисные ковекторы следующим образом:

$$dx^\mu, \quad \mu = \{x, y, z, t\}.$$

Базисные векторы и ковекторы

Аналогично запишем произвольный базисный вектор в виде $\partial/\partial x^\mu$. Тогда наши разложения приобретут следующий вид:

$$\lambda = \sum_\mu \lambda^\mu \frac{\partial}{\partial x^\mu} \quad (24.6)$$

для касательного вектора и

$$\alpha = \sum_\mu \alpha_\mu dx^\mu \quad (24.7)$$

для 1-формы. Авторы некоторых книг проявляют определенный педантизм, требуя, чтобы координаты обозначались как x^1, x^2 и т. д. и суммирование осуществлялось по числовому индексу. Это не очень хороший способ, и мы предпочитаем следовать соглашению, что суммирование по немому индексу осуществляется по всему набору его значений.

Свободные индексы

Если в уравнении имеется свободный индекс, то оно должно быть справедливо при любом значении индекса. Такой индекс должен стоять у каждого члена уравнения, причем в одной и той же позиции, наверху или внизу; в противном случае нарушилось бы условие линейности для тензоров.

Пример 1

Уравнение

$$a^\mu b^\nu = h^{\mu\nu} \quad (24.8)$$

означает, что

$$a^t b^t = h^{tt}, \quad (24.9)$$

$$a^t b^x = h^{tx} \quad (24.10)$$

и т. д. для всех шестнадцати уравнений в пространстве четырех измерений. Обратите внимание, насколько компактны такие обозначения.

Пример 2

Что означает выражение

$$dx^\mu \cdot \frac{\partial}{\partial x^\nu}?$$

Это результат действия ковектора на вектор, следовательно, оно должно быть числом, которое зависит от того, каковы значения μ и ν . Как легко проверить,

$$\partial x^\mu \cdot \frac{\partial}{\partial x^\nu} = \begin{cases} 1, & \text{если } \mu = \nu, \\ 0, & \text{если } \mu \neq \nu. \end{cases} \quad (24.11)$$

так как dx^μ и $(\partial/\partial x^\nu)$ дуальны. Например,

$$dx \cdot \frac{\partial}{\partial x} = 1, \quad (24.12)$$

$$dx \cdot \frac{\partial}{\partial y} = 0. \quad (24.13)$$

[Чтобы понять смысл тензорного уравнения, полезно выписать несколько его компонент.]

Наконец, в тензорных уравнениях суммирование всегда проводится по двум повторяющимся немой индексам, один из которых верхний, другой нижний. Существенное упрощение достигается, если принять, что по двум повторяющимся индексам всегда подразумевается суммирование. Уравнение, в котором немой индекс появляется в каждом члене только один раз, должно быть справедливым при любом допустимом значении индекса. Если в каждом члене уравнения немой индекс появляется дважды, то по такому индексу должно быть проведено суммирование. (Если индекс появляется три или большее число раз, то это значит, что допущена ошибка.) Так как в качестве немой индекса может быть выбран любой символ, его можно изменять по нашему усмотрению при условии, что он меняется одновременно во всех членах уравнения. При вычислениях

Правило суммирования

иногда в уравнении появляются четыре одинаковых немых индекса. Такая распространенная ошибка возникает, когда одно суммирование подставляют в другое, не изменив надлежащим образом обозначения одного из индексов суммирования. Первое время при необходимости мы будем указывать на возможность совершения такой ошибки.

Свертывание

Индексные обозначения позволяют записать в компактной форме все изученные выше операции над тензорами. Свертывание 1-формы с вектором определяется следующим образом:

$$\omega \cdot v \equiv (\omega_\mu dx^\mu) \cdot \left(v^\alpha \frac{\partial}{\partial x^\alpha} \right). \quad (24.14)$$

В силу свойства линейности можно поменять местами операции суммирования и свертывания; в результате получаем

$$\omega \cdot v = \omega_\mu v^\alpha \left[dx^\mu \cdot \frac{\partial}{\partial x^\alpha} \right]. \quad (24.15)$$

В рассмотренном выше примере 2 было показано, что величина в квадратных скобках обращается в нуль, если только индексы μ и α не совпадают. Это сводит двойную сумму к одной сумме «по диагонали»

$$\begin{aligned} \omega \cdot v &= \omega_t v^t + \omega_t \psi^x + \dots + \\ &+ \omega_x \psi^t + \omega_x v^x + \dots + \\ &+ \dots, \end{aligned} \quad (24.16)$$

Полученную диагональную сумму можно записать в виде

$$\omega \cdot v = \omega_\mu v^\mu. \quad (24.17)$$

Члены типа $dx^\mu \cdot (\partial/\partial x^\nu)$ появляются довольно часто, поэтому для простоты мы определим особый набор чисел δ_ν^μ , который аналогичен единичной матрице и задается следующим образом:

$$\delta_\nu^\mu \equiv \begin{cases} 1, & \text{если } \mu = \nu, \\ 0, & \text{если } \mu \neq \nu. \end{cases} \quad (24.18)$$

Величина δ_ν^μ называется дельта-символом Кронекера.

Пример

Справедливо соотношение

$$\omega_\mu \delta_\nu^\mu = \omega_\nu, \quad (24.19)$$

в чем легко убедиться, расписав компоненту $\nu = x$ этого равенства:

$$\omega_x \delta_x^x + \omega_t \delta_x^t + \dots = \omega_x. \quad (24.20)$$

Тензор $\mathcal{G}$, принадлежащий пространству $V^* \otimes V^*$, можно представить в виде

$$\mathcal{G} = g_{\mu\nu} dx^\mu \otimes dx^\nu. \quad (24.21)$$

Тогда отображение $\mathcal{G}: V \times V \rightarrow \mathbb{R}$ с использованием линейности можно осуществить следующим образом:

$$\begin{aligned} \mathcal{G} \cdot (a, b) &= (g_{\mu\nu} dx^\mu \otimes dx^\nu) \cdot \left(a^\alpha \frac{\partial}{\partial x^\alpha}, b^\beta \frac{\partial}{\partial x^\beta} \right) = \\ &= g_{\mu\nu} a^\alpha b^\beta \left(dx^\mu \cdot \frac{\partial}{\partial x^\alpha} \right) \left(dx^\nu \cdot \frac{\partial}{\partial x^\beta} \right) = \\ &= g_{\mu\nu} a^\alpha b^\beta \delta_\alpha^\mu \delta_\beta^\nu = \\ &= g_{\mu\nu} a^\mu b^\nu. \end{aligned} \quad (24.22)$$

Как легко показать, операцию частичного свертывания можно представить в виде

$$\mathcal{G} \cdot a = g_{\mu\nu} a^\mu dx^\nu. \quad (24.23)$$

Компоненты тензорного произведения равны произведению компонент сомножителей

$$a \otimes b = a^\mu b^\nu \frac{\partial}{\partial x^\mu} \otimes \frac{\partial}{\partial x^\nu}. \quad (24.24)$$

Симметричность тензора $\mathcal{G} = g_{\mu\nu} dx^\mu \otimes dx^\nu$ находит отражение в симметричности его компонент

$$g_{\mu\nu} = g_{\nu\mu}. \quad (24.25)$$

Что можно сказать о тензоре, обратном симметричному тензору $\mathcal{G}$? Запишем обратный тензор в виде

$$\mathcal{G}^{-1} \equiv g^{\mu\nu} \frac{\partial}{\partial x^\mu} \otimes \frac{\partial}{\partial x^\nu}, \quad (24.26)$$

тогда тождество

$$\mathcal{G}^{-1}(\mathcal{G} \cdot a) = a \quad (24.27)$$

можно преобразовать к виду

$$g^{\mu\alpha} g_{\alpha\beta} a^\beta = a^\mu = \delta_\beta^\mu a^\beta. \quad (24.28)$$

[Не обязательно использовать одну и ту же букву для обозначения тензора и его компонент, хотя обычно это и удобно.]

Метрический тензор

[Вот пример ситуации, когда возможно совершить указанную выше ошибку. Немые индексы у a и b должны быть различными и отличаться от μ и ν .]

Частичное свертывание

[Напомним, что $\mathcal{G} \cdot a$ — это сокращенное обозначение для выражения $\mathcal{G} \cdot (a, -)$.]

Тензорное произведение

Обратный тензор

Так как это справедливо для любых a^β , должно иметь место равенство

$$g^{\mu\alpha}g_{\alpha\beta} = \delta_\beta^\mu. \quad (24.29)$$

В случае четырех измерений уравнение (24.29) представляет собой 16 алгебраических уравнений для 16 компонент $g^{\mu\nu}$ тензора $\mathcal{S}^1$.

Тензоры: операторы или наборы чисел

Физики-релятивисты придают слишком большое значение индексным обозначениям. Вместо того чтобы дать формальное определение компонент тензора, они рассматривают снабженные индексами символы как другое определение тензора. Действительно, индексные обозначения особенно удобны, так как число и расположение индексов говорит нам, с каким типом тензора или вектора мы имеем дело. Если проводится операция свертывания, то появляются повторяющиеся индексы. Поэтому мы обычно называем $g_{\mu\nu}$ тензором, а не компонентами тензора. Недостаток классического тензорного анализа не в том, что в нем используются компоненты, а в том, что в нем не проводится различия между v^μ и v_μ . Это не приводит к ошибкам, если задана метрика, но даже в таком случае отождествление векторов и ковекторов не слишком удобно.

Пример

Если рассматривать тензоры как набор чисел, то появляется искушение написать

$$\delta_\nu^\mu = \delta_\mu^\nu. \quad (24.30)$$

Почему это уравнение бессмысленно? Предположим, что мы хотим восстановить исходное тензорное уравнение, свертывая обе части уравнения с соответствующими базисными элементами. При суммировании по индексу μ мы получим либо

$$\delta_\nu^\mu \frac{\partial}{\partial x^\mu} = \delta_\mu^\nu dx^\mu, \quad (24.31)$$

либо

$$\delta_\nu^\mu dx^\mu = \delta_\mu^\nu dx^\mu. \quad (24.32)$$

На нашем языке оба равенства не имеют смысла.

ЗАДАЧА

24.1. (15) Рассмотрите тождественное отображение

$$I: V^* \rightarrow V^*; \omega \mapsto \omega.$$

Покажите, что это тензор и найдите его компоненты. (Используйте символ δ^μ_ν .)

25. Векторные поля

До сих пор мы занимались построением алгебры векторов и тензоров в некоторой фиксированной точке. Сделаем теперь следующий шаг и перейдем к рассмотрению векторных и тензорных полей. Говорят, что задано векторное поле, если каждой точке пространства-времени ставится в соответствие некоторый вектор. Аналогичным образом определяется тензорное поле. Анализ векторных и тензорных полей в пространствах общего вида значительно более сложен, чем в пространствах-времени специальной теории относительности. К счастью, нам не потребуется чересчур углубляться в эти вопросы. Нам достаточно лишь немного видоизменить интуитивное понятие векторного поля, которое вводится в курсах электромагнетизма или механики жидкости. Новое, с чем мы здесь столкнемся, заключается в установлении тесной связи между векторными полями, системами обыкновенных дифференциальных уравнений и преобразованиями пространства.

Выбор системы координат определяет в каждой точке пространства-времени набор базисных векторов. При этом гладкое векторное поле представляется набором компонент вектора, которые являются гладкими функциями координат точки.

Пример 1

Примером гладкого векторного поля на плоскости может служить поле

$$v = \frac{x}{2} \frac{\partial}{\partial y} - \frac{y}{2} \frac{\partial}{\partial x}; \quad (25.1)$$

оно изображено на рис. 25.1.

Если задано некоторое векторное поле, то можно попытаться отыскать семейство кривых, для которых векторное поле в каждой точке дает локальную линейную аппроксимацию. Отдельная параметризованная кривая, касательный вектор к которой в каждой точке совпадает с вектором, принадлежащим векторному полю, называется интегральной кривой векторного поля.

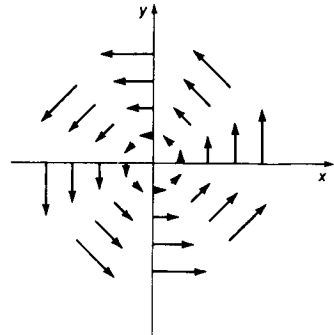


Рис. 25.1

Векторное поле

[Мы откладываем подробное обсуждение криволинейных координат до разд. 26.]

Интегральные кривые

[Продолжение примера 1. Здесь a — параметр, различающий кривые. Мы выбираем $a > 0$.]

[По поводу этого равенства см. разд. 17 и стр. 138.]

Семейство кривых

$$\gamma_a: \mathbb{R} \rightarrow \mathbb{R}^2; \theta \mapsto (a \cos \frac{\theta}{2}, a \sin \frac{\theta}{2}) \quad (25.2)$$

состоит из интегральных кривых определенного выше векторного поля v . Покажем это. Касательный вектор $\dot{\gamma}_a$ к кривой γ_a в точке θ имеет вид

$$\dot{\gamma}_a = -\frac{a}{2} \sin \frac{\theta}{2} \frac{\partial}{\partial x} + \frac{a}{2} \cos \frac{\theta}{2} \frac{\partial}{\partial y}. \quad (25.3)$$

Если точка (x, y) лежит на кривой γ_a , то

$$\begin{aligned} x^2 + y^2 &= a^2, \\ \cos \theta &= \frac{x}{a}, \\ \sin \theta &= \frac{y}{a}, \end{aligned} \quad (25.4)$$

поэтому

$$\dot{\gamma}_a = -\frac{y}{2} \frac{\partial}{\partial x} + \frac{x}{2} \frac{\partial}{\partial y}. \quad (25.5)$$

Полученный результат показывает, что кривые γ_a представляют собой интегральные кривые векторного поля v . Они изображены на рис. 25.2.

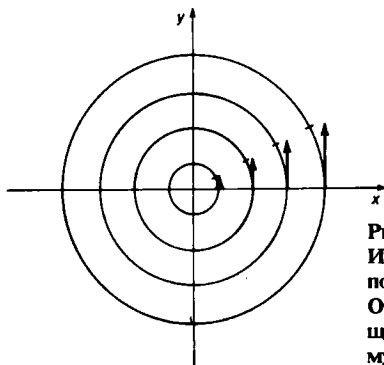


Рис. 25.2

Интегральные кривые векторного поля, изображенного на рис. 25.1. Отмечены точки, соответствующие значению параметра, равному единице.

Как найти интегральные кривые? Возьмем произвольную кривую

$$\gamma: s \mapsto [X(s), Y(s), \dots]. \quad (25.6)$$

Соответствующий касательный вектор имеет вид

$$\dot{\gamma} = \frac{dX}{ds} \frac{\partial}{\partial x} + \frac{dY}{ds} \frac{\partial}{\partial y} + \dots \quad (25.7)$$

Чтобы вектор (25.7) соответствовал заданному векторному полю

$$v = v^x \frac{\partial}{\partial x} + v^y \frac{\partial}{\partial y} + \dots, \quad (25.8)$$

должны выполняться равенства

$$\begin{aligned} \frac{dX}{ds} &= v^x(x, y, \dots), \\ \frac{dY}{ds} &= v^y(x, y, \dots), \\ &\dots, \end{aligned} \quad (25.9)$$

Система обыкновенных дифференциальных уравнений

где каждая из компонент $v^x, v^y, \dots$ является гладкой функцией координат. Мы видим, что нахождение интегральных кривых векторного поля эквивалентно нахождению решения системы обыкновенных дифференциальных уравнений. Именно поэтому их называют интегральными кривыми. С таким же успехом можно свести рассмотрение системы обыкновенных дифференциальных уравнений к исследованию соответствующего векторного поля, что оказалось весьма продуктивным подходом к решению некоторых задач.

Пример 2

Простой гармонический осциллятор описывается уравнением Гармонический осциллятор

$$\ddot{x} + \omega^2 x = 0. \quad (25.10)$$

Если рассматривать скорость как независимую переменную

$$y \equiv \dot{x}, \quad (25.11)$$

то на плоскости (x, y) , которая называется фазовой плоскостью, движение системы описывается системой уравнений

$$\begin{aligned} \dot{x} &= y, \\ \dot{y} &= -\omega^2 x. \end{aligned} \quad (25.12)$$

Решение этой системы эквивалентно рассмотрению на фазовой

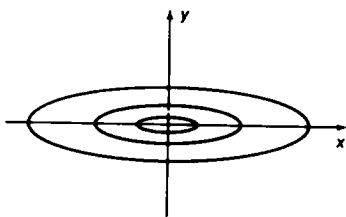


Рис. 25.3

Интегральные кривые, соответствующие значению $\omega = \sin 15^\circ = 0,26$.

Нелинейный осциллятор

плоскости векторного поля

$$v = y \frac{\partial}{\partial x} - \omega^2 x \frac{\partial}{\partial y}. \quad (25.13)$$

Интегральные кривые поля v имеют вид

$$\gamma_a: t \mapsto (a \cos \omega t, -a\omega \sin \omega t). \quad (25.14)$$

Эти кривые изображены на рис. 25.3.

Пример 3

К числу чаще всего рассматриваемых нелинейных осцилляторов относится осциллятор Ван-дер-Поля, который описывается уравнением

$$\ddot{y} + ky - \mu \dot{y} \left[1 - \left(\frac{y}{d} \right)^2 \right] = 0. \quad (25.15)$$

Это уравнение задает векторное поле

$$v = \{ky - \mu x \left[1 - \left(\frac{y}{d} \right)^2 \right]\} \frac{\partial}{\partial x} - x \frac{\partial}{\partial y}. \quad (25.16)$$

Векторное поле (25.16) изображено на рис. 25.4; несколько интегральных кривых, соответствующих значениям параметров $k = 0,231$, $d = 0,465$, $\mu = 0,769$, приведено на рис. 25.5. Чтобы не загромождать рисунок, все векторы уменьшены в пять раз.

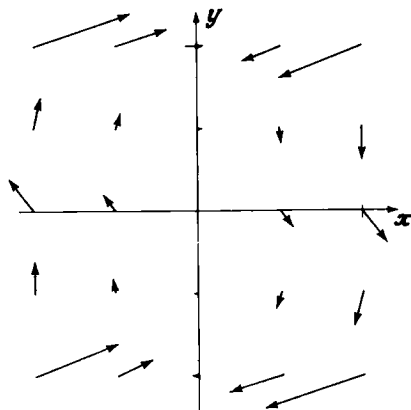


Рис. 25.4

Векторное поле Ван-дер-Поля.

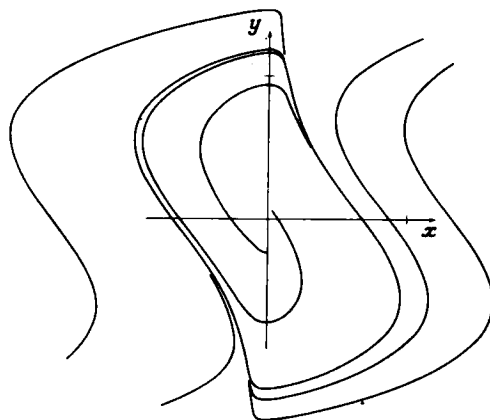


Рис. 25.5

Интегральные кривые поля Ван-дер-Поля.

Пример 4

В механике Гамильтона рассматривается движение системы в $2n$ -мерном фазовом пространстве с координатами

$$(q^1, q^2, \dots, q^n, p_1, p_2, \dots, p_n) = (q, p). \quad (25.17)$$

Любая гамильтонова система описывается функцией $H(q, p)$, заданной в фазовом пространстве. Динамика системы определяется векторным полем

$$v = \frac{\partial H}{\partial p_\mu} \frac{\partial}{\partial q^\mu} - \frac{\partial H}{\partial q^\mu} \frac{\partial}{\partial p_\mu}, \quad (25.18)$$

что эквивалентно системе обычных дифференциальных уравнений

$$\begin{aligned} \frac{dq^\mu}{dt} &= \frac{\partial H}{\partial p_\mu}, \\ \frac{dp_\mu}{dt} &= -\frac{\partial H}{\partial q^\mu}. \end{aligned} \quad (25.19)$$

Фазовое пространство (q, p) лишено метрики. Оно обладает своей собственной специфической геометрической структурой, которая порождается тензорным полем

$$\Omega = dp_\mu \otimes dq^\mu - dq^\mu \otimes dp_\mu. \quad (25.20)$$

Механика Гамильтона

[Здесь мы не придерживаемся общепринятой практики и не пишем координатные индексы сверху, а используем достаточно понятную сокращенную форму записи.]

[Правило суммирования!]

[Как dp_μ , так и dq^μ представляют собой ковекторы. Мы не придерживаемся традиционного расположения индексов.]

Пример 5

Не вдаваясь в детали, отметим как интересный факт, что уравнения Максвелла имеют ту же самую форму

$$\begin{aligned} \frac{\partial \mathbf{B}}{\partial t} &= -\nabla \times \mathbf{E}, \\ \frac{\partial \mathbf{E}}{\partial t} &= \nabla \times \mathbf{B} - 4\pi \mathbf{J}. \end{aligned} \quad (25.21)$$

Это уравнения в частных производных, и соответствующее векторное поле задано в бесконечномерном пространстве.

Электромагнетизм

Предположим, что мы нашли полное семейство интегральных кривых, причем через каждую точку проходит только одна кривая. Используя эти кривые, можно задать однопараметрическое семейство преобразований пространства Φ_w . Будем действовать следующим образом. Возьмем точку p и найдем интегральную кривую γ , проходящую через p . Пусть

Преобразования

[Преобразования были рассмотрены в разд. 4 и 6.]

u — значение параметра вдоль кривой, которое соответствует точке p :

$$\gamma(u) = p. \quad (25.22)$$

Тогда определим преобразование следующим образом:

$$\Phi_w: p \mapsto \gamma(u + w). \quad (25.23)$$

При таком преобразовании каждая точка p перемещается вдоль соответствующей ей интегральной кривой и занимает положение, соответствующее новому значению параметра, изменившемуся на величину w . Это семейство преобразований называется потоком векторного поля.

Инфинитезимальные преобразования

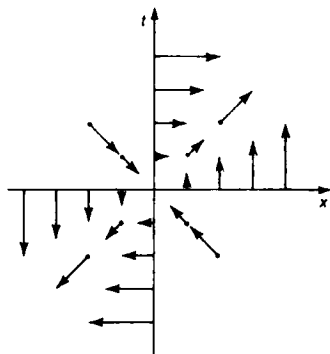


Рис. 25.6

Чтобы рисунок не выглядел слишком сложным, все векторы уменьшены в два раза.

Пример

Для векторного поля

$$v = x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \quad (25.24)$$

поток Φ_w соответствует повороту на угол w в плоскости (x, y) .

Векторное поле определяет инфинитезимальное преобразование.

Пример

Векторное поле

$$v = x \frac{\partial}{\partial t} + t \frac{\partial}{\partial x} \quad (25.25)$$

определяет инфинитезимальное преобразование Лоренца

$$(x, t) \mapsto (x \operatorname{ch} \psi + t \operatorname{sh} \psi, x \operatorname{sh} \psi + t \operatorname{ch} \psi). \quad (25.26)$$

(рис. 25.6 и 25.7).

Симметрия

[Подробнее о симметрии см. книгу [39].]

Важное свойство пространства-времени — его симметрия. Мы описываем симметрию, задавая преобразования, которые оставляют пространство-время неизменным. Удобнее и проще задавать бесконечно малые преобразования, т. е. векторные поля.

Гладкое ковекторное поле определяется аналогичным образом; его можно использовать, например, для описания фазы волнового пакета, распространяющегося в пространстве-

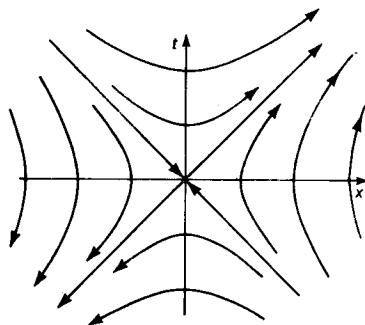


Рис. 25.7

времени. Геометрия пространства-времени описывается гладким тензорным полем, задающим метрику.

Связь между касательными векторами и обыкновенными дифференциальными уравнениями определяется смыслом касательных векторов как локальной линейной аппроксимации кривых. Касательный вектор является также оператором, который действует на функции; в этом качестве касательные векторы имеют отношение к уравнениям в частных производных. Все сказанное оказывается полезным при изучении распространения диспергирующих волн.

Дифференциальные операторы

Пример

Предположим, что функция $f(x, y)$ удовлетворяет уравнению

[Это продолжение примера 2 из разд. 17.]

$$2y \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} = 0 \quad (25.27)$$

с граничным условием

$$f(x, 0) = e^{-x^2}. \quad (25.28)$$

Достаточно ли этого для определения функции f ? Каким образом мы вычисляем ее?

С геометрической точки зрения уравнение (25.27) означает, что функция f постоянна вдоль направления

$$2y \frac{\partial}{\partial x} + \frac{\partial}{\partial y}.$$

Следовательно, она постоянна на интегральных кривых этого векторного поля. Такие интегральные кривые называются ха-

Характеристики

рактеристиками. В нашем случае интегральные кривые представляют собой параболы

$$\gamma_a: u \mapsto (x, y) = (u^2 + a, u), \quad (25.29)$$

где a — параметр, которым обозначаются кривые.

Значение функции f в произвольной точке (x_0, y_0) находим непосредственно (рис. 25.8). Разрешим (25.29) относительно u и a :

$$u = y_0, \quad (25.30)$$

$$a = x_0 - y_0^2. \quad (25.31)$$

Кривая γ_a пересекает ось x в точке $x = a$. Значение функции f в этой точке равно ее значению в точке (x_0, y_0) , поэтому

$$f(x_0, y_0) = e^{-(x_0 - y_0^2)^2}. \quad (25.32)$$

ЗАДАЧИ

25.1. (16) Опишите преобразование, которое порождается векторным полем

$$y \frac{\partial}{\partial x}.$$

25.2. (16) Найдите векторное поле, которое задает однородное растяжение.

25.3. (14) Найдите окружности, которые, если их рассматривать как множество точек, инвариантны относительно инфинитезимальных поворотов, задаваемых векторным полем

$$(x - 1) \frac{\partial}{\partial y} - y \frac{\partial}{\partial x}.$$

25.4. (19) Покажите, что гипербола

$$t^2 - x^2 = 1$$

инвариантна относительно преобразований Лоренца

$$x \frac{\partial}{\partial t} + t \frac{\partial}{\partial x}.$$

25.5. (38) Если в вашем распоряжении имеется программируемый микрокалькулятор, найдите несколько численных решений уравнения Ван-дер-Поля:

25.6. (33) Какая специфическая проблема возникает при реше-

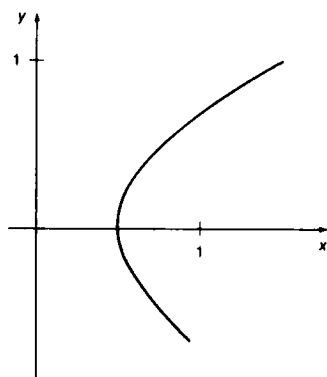


Рис. 25.8
Кривая, на которой функция f постоянна.

нии уравнения в частных производных

$$\frac{\partial f}{\partial x} + 2x \frac{\partial f}{\partial y} = 0$$

с граничным условием, заданным на оси x ?

26. Многообразия

Он купил большую карту,
на которой было изображено море
без каких-либо признаков берегов,
и моряки ликовали, что все хорошо ее понимают.
«Что хорошего в картах Меркатора
с их Северным полюсом и экватором,
тропиками, зонами и меридианами,
— восклицала команда, —
Это всего лишь условные знаки!
На других картах только мешают
все эти острова и мысы!
Мы должны торжественно поблагодарить
нашего храброго капитана — он купил нам
самую лучшую, совершенную,
не ограниченную, абсолютную карту!»

Льюис Кэрролл

До сих пор мы относились к пространству-времени как к линейному пространству. Однако наши определения допускали использование криволинейных координат, например полярных, и предположение, что в основе находится векторное пространство, на самом деле не использовалось. В отличие от специальной теории относительности пространства эйнштейновской общей теории относительности не являются векторными пространствами. Перейдем теперь к рассмотрению пространств более общего вида, которые выглядят как векторное пространство только в малой пространственно-временной области и совсем не похожи на векторное пространство в больших масштабах. Такие пространства, соответствующим образом определенные, называются многообразиями.

Пример 1

Самым известным примером многообразия служит поверхность сферы. В окрестности любой точки ее можно аппроксимировать плоскостью, но в целом она существенно отличается от плоскости. Следует почаще обращаться к этому примеру при рассмотрении конкретных свойств многообразий.

[Бегло просмотрев разд. 38, вы увидите, почему нам необходимы эти понятия и как мы собираемся их использовать.]

Многообразия

Многообразие отличается от векторного пространства во многих отношениях. Оно не обладает линейной структурой векторного пространства. В нем нет нулевого элемента, не определены также операции сложения и умножения на число. Не существует способа определения понятия параллельности на расстоянии. Отсутствует представление о постоянном векторном поле. Кроме того, не всегда возможно ввести единую систему координат, гладко покрывающую все многообразие. Тем не менее все развитые здесь геометрические идеи без всякого изменения применимы к многообразиям. Последний раз мы использовали предположение о векторной структуре пространства-времени в разд. 12 при определении ускорения. Именно с расчетом на будущее мы так тщательно определили там касательный вектор.

Многообразие представляет собой подходящую модель для общерелятивистского описания Вселенной. Почему в таком случае авторы очень многих книг поверхностно комментируют это понятие, давая вид, что оно совершенно очевидно и не нуждается в обсуждении? Причина состоит в том, что большинство вводных курсов носит чисто описательный характер, и авторы считают возможным ограничиться достаточно беглым рассмотрением понятия многообразия (как впрочем и любого другого). Мы же собираемся проводить конкретные вычисления и не можем довольствоваться неопределенными и непроверенными представлениями об объекте исследования. Существует и другая причина того, почему даже во вводных математических курсах не рассматривается понятие многообразия. Она связана с тем, что последовательное изложение соответствующего раздела математики носит достаточно абстрактный и специальный характер и вместе с тем мало что дает с точки зрения умения проводить конкретные вычисления на многообразиях. В этой книге мы избираем неформальный подход. Мы исходим из допущения, что многообразия необходимы здесь и во многих других областях физики. Каждый серьезный физик должен иметь о них представление. Однако мы не будем пытаться дать строгое определение многообразия. Такой подход мало пригоден на первом этапе изучения многообразий; не слишком много пользы принесут читателю и замысловатые определения. Будем вместо этого опираться на замечательную способность человеческого сознания улавливать характерное при рассмотрении нескольких примеров. Усвоив материал этой книги, можно приступить к тщательному изучению основных идей исчисления на многообразиях. Приведен-

ные в книге примеры послужат основой для усвоения абстрактных понятий этой области математики.

Для читателей, обладающих достаточной математической подготовкой, которые не могут обойтись без формальных определений, скажем, что все многообразия, играющие важную роль в космологии, можно рассматривать как подпространства некоторых векторных пространств. В таких случаях касательные пространства в действительности представляют собой касательные гиперплоскости, и все определения, сформулированные для случая линейного пространства, могут использоваться на законных основаниях. Мы действительно будем использовать такое погружение для определения некоторых из наших многообразий.

При переходе от евклидова метрического пространства к векторному пространству теряется нечто в геометрической структуре пространства: понятия длины, угла и перпендикулярности перестают иметь смысл. Это было отражено в том представлении, которое мы использовали для таких пространств. Мы перешли от представлений, ковариантных относительно ортогональных преобразований, к представлениям, ковариантным относительно произвольных линейных преобразований.

При переходе к многообразиям теряется линейная структура пространства. Теперь наши определения должны быть ковариантны относительно произвольных криволинейных преобразований координат. Действительно, на протяжении всего обсуждения геометрии мы придавали нашим понятиям форму, которая обладала такой ковариантностью. Благодаря этому теперь мы можем описывать многообразия, не производя больших изменений в формализме, а только учитывая перечисленные выше различия. На многообразиях больше не существует выделенного класса линейных координат, но мы и не пользовались такими координатами где бы то ни было при построении геометрии.

Возникают два вопроса: «Как точно определить многообразие?» и «Как мы описываем многообразие?». На оба этих вопроса можно дать такой ответ. *Многообразия* представляют собой множества, которые могут быть заданы следующим образом. Части многообразий могут быть покрыты картами. *Картой* называется взаимнооднозначное отображение некоторой части многообразия в $\mathbb{R}^n$. Точнее говоря, карта отображает подмножество многообразия на открытое множество в $\mathbb{R}^n$ (рис. 26.1). Многообразие может быть представлено набором

Утраченная структура

[Здесь вы можете вернуться назад и просмотреть обсуждение проблемы ковариантности в разд. 3 и 4.]

Представление и определение

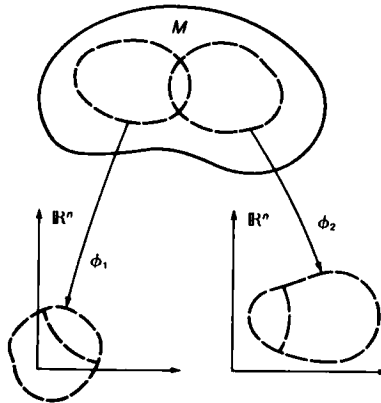


Рис. 26.1

Две перекрывающиеся карты, заданные на многообразии M .

Карты

[Для наших целей достаточно обычного определения открытого и замкнутого множеств. Можно было бы воспользоваться топологическими определениями открытого множества, но соответствующие достаточно экзотические ситуации редко встречаются.]

Размерность

таких карт. Как правило, одна карта не может покрыть всего многообразия. Каждая карта наделяет соответствующую часть многообразия координатами. Там, где карты перекрываются, преобразование от одной системы координат к другой должно быть гладким и иметь гладкое обратное преобразование. Каждая точка многообразия должна отображаться хотя бы одной картой. Все карты осуществляют отображение в пространство $\mathbb{R}^n$ одной и той же размерности n ; число n называется размерностью многообразия. Указанные ограничения приводят к тому, что в окрестности любой точки многообразие ведет себя как обычное пространство $\mathbb{R}^n$.

Окружность

Пример 2

Окружность, т. е. множество принадлежащих $\mathbb{R}^2$ точек, таких, что

$$x^2 + y^2 = 1, \quad (26.1)$$

представляет собой одномерное многообразие. Для окружности картами должны быть отображения в $\mathbb{R}$. В качестве карт можно выбрать отображения

$$(x, y) \mapsto u = 2\sqrt{\frac{1-y}{1+y}} \quad (26.2)$$

для всех точек окружности, за исключением точки с координатами

тами $(0, -1)$ и

$$(x, y) \mapsto v = 2\sqrt{\frac{1+y}{1-y}} \quad (26.3)$$

для всех точек, кроме точки $(0, 1)$ (рис. 26.2 и 26.3). И в том, и в другом случае отображение осуществляется на всю действительную прямую, которая является открытым множеством. При $x \neq 0$ карты перекрываются, причем преобразование от координаты v к координате u осуществляется посредством гладкой функции

$$u = \frac{4}{v}. \quad (26.4)$$

Таким образом, каждая точка окружности отображается какой-нибудь картой; окружность представляет собой многообразие.

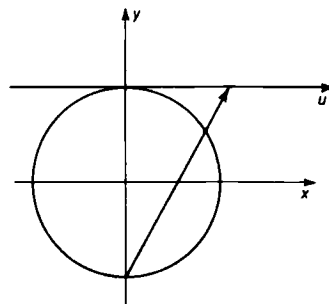


Рис. 26.2
Одна карта для окружности.

Пример 3

Почему обычный полярный угол θ не определяет карту многообразия в строгом смысле? Чтобы каждая точка покрывалась картой не более одного раза, мы должны ограничить область изменения угла θ интервалом $0 \leq \theta \leq 2\pi$. Но это замкнутое множество. Если ограничиться открытым интервалом $0 < \theta < 2\pi$, то одна точка окажется непокрытой. Таким образом, сформулированные выше требования не выполняются.

Такие неувязки возникают в тех случаях, когда мы не следуем сформулированным правилам. В нашем случае функция $f = \theta$ не является однозначной функцией на окружности, хотя является таковой на θ -карте.

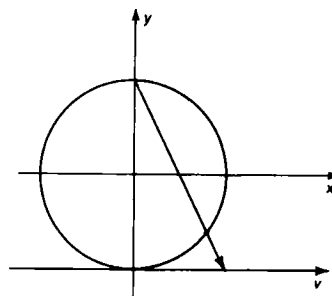


Рис. 26.3
Другая карта для окружности.

Неполные карты типа рассмотренной в примере 3 бесполезны, если мы хотим доказать, что данное множество представляет собой многообразие. Однако если такая карта определена, то она служит в высшей степени полезным представлением многообразия; необходимо лишь помнить, что часть точек не отображена должным образом. Многообразия, которые используются в космологии и в общей теории относительности, почти всегда представляются такими неполными картами.

Неполные карты

[Использование таких неполных карт освещено в разд. 38.]

Пример 4

Множество точек $\mathbb{R}^2$, которое определяется соотношением

$$xy = 0 \quad (26.5)$$

Контрпример

[Эти замечания можно пропустить, если вы не знакомы с топологическим представлением об открытых множествах, которые можно задавать произвольно в рамках определенных постулатов.]

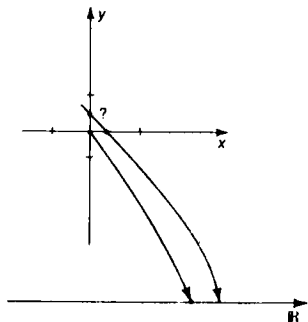


Рис. 26.4

Пример множества, которое не является многообразием. Попытка найти взаимнооднозначное отображение терпит неудачу.

[Вспомните обсуждение операции сложения касательных кривых, данное на стр. 134.]

Касательные векторы

(рис. 26.4), не является многообразием. В этом виновата точка $(0,0)$. Невозможно построить гладкое взаимнооднозначное отображение открытой окрестности этой точки на открытое множество точек, принадлежащих действительной прямой. Открытое множество, содержащее точку $(0,0)$, дается соотношениями

$$\begin{aligned} a < x < b, & \quad a < 0 < b, \\ c < y < d, & \quad c < 0 < d. \end{aligned} \quad (26.6)$$

Если мы честно отобразим точки оси x , то не останется места для точек, лежащих на оси y , и наоборот.

С другой стороны, нельзя считать открытыми множества

$$\begin{aligned} a < x < b, \\ y = 0 \end{aligned} \quad (26.7)$$

и

$$\begin{aligned} c < y < d, \\ x = 0, \end{aligned} \quad (26.8)$$

так как пересечение открытых множеств, в нашем случае точка $(0,0)$, также должно быть открытым множеством. Но если назвать точку $(0,0)$ открытым множеством, то мы не сможем взаимнооднозначно отобразить ее на открытое множество в $\mathbb{R}$, где отдельная точка является замкнутым множеством.

Важнейшие понятия кривых, функций, касательных векторов и 1-форм были определены таким образом, что они полностью сохраняют свой смысл и на многообразиях.

Пример 5

Предположим, что мы пытаемся проводить вычисления на множестве, которое определено в примере 4. Для линий, проходящих через точку $(0,0)$, мы можем найти локальную линейную аппроксимацию. Однако соответствующие касательные векторы нельзя складывать. Они не образуют векторное пространство.

Наше определение касательного вектора как локальной линейной аппроксимации кривой было сформулировано безотносительно к какой-либо линейной системе координат. Линейность пространства касательных векторов была обусловлена процедурой предельного перехода, используемой для их определения. В бесконечно малой области любое гладкое преобра-

зование является линейным. Поэтому в различных системах координат мы получим различные представления для любого касательного вектора, но все они будут представлениями одного и того же геометрического объекта.

Мы можем убедиться путем вычислений в явном виде, что линейная структура касательного пространства сохраняется при произвольных криволинейных преобразованиях координат. Предположим, что в некоторой системе координат $(x^1, x^2, \dots, x^n)$, которую мы будем сокращено обозначать x , имеется кривая γ . Кривая задается отображением γ_x :

$$\gamma_x: s \mapsto X(s). \quad (26.9)$$

Локальное рассмотрение линейной структуры

[Здесь под X понимается n функций одной переменной.]

Касательный вектор этой кривой в выбранной системе координат имеет вид

$$\dot{\gamma}_x = \frac{dX^\mu}{ds} \frac{\partial}{\partial x^\mu}. \quad (26.10)$$

Рассмотрим вторую систему координат y , которые связаны с координатами x с помощью n заданных функций n переменных

$$y = Y(x). \quad (26.11)$$

Кривая γ в новой системе координат задается отображением

$$\gamma_y: s \mapsto Y[X(s)]; \quad (26.12)$$

соответствующий касательный вектор может быть представлен в виде

$$\dot{\gamma}_y = \frac{d}{ds} \left\{ Y^\mu[X(s)] \right\} \frac{\partial}{\partial y^\mu}. \quad (26.13)$$

Это выражение можно упростить, если воспользоваться цепным правилом для частных производных; тогда получаем

$$\dot{\gamma}_y = \frac{\partial Y^\mu}{\partial x^\nu} \frac{dX^\nu}{ds} \frac{\partial}{\partial y^\mu}. \quad (26.14)$$

Результат показывает, что компоненты касательного вектора в одной системе координат представляют собой линейные функции компонент в любой другой системе. Числа $\partial Y^\mu / \partial x^\nu$ образуют матрицу $n \times n$. Нелинейное преобразование $Y(x)$ линеаризуется дифференцированием.

Так как компоненты связаны линейными уравнениями, линейная структура пространства касательных векторов полно-

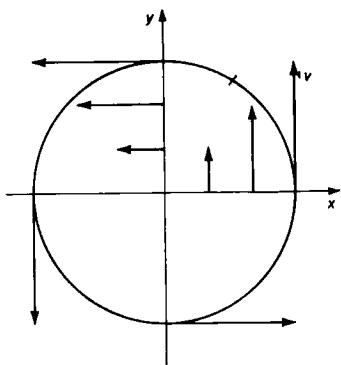


Рис. 26.5

Векторное поле и соответствующие интегральные кривые. Отмечена точка, соответствующая значениям параметров, равным единице.

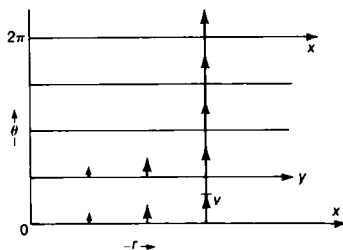


Рис. 26.6

Те же самые векторы, интегральные кривые и оси, что и на рис. 26.5. Почему в этом случае вдоль кривой укладывается пять векторов?

стью ковариантна относительно этих преобразований. Две кривые, которые имеют одинаковые касательные в одной системе координат, имеют одинаковые касательные в любой другой системе. Результаты операций сложения и умножения на число, выполненных в каждой из двух систем координат, согласуются.

Можно найти определения касательных векторов, которые явным образом ковариантны. В таких определениях вообще не используется понятие координат. При этом исключается необходимость проверки должной ковариантности определения. Такие безкоординатные определения выглядят несколько абстрактными и недостаточно обоснованными с точки зрения физиков. В одном из них вектор определяется как класс эквивалентности параметризованных кривых, соприкасающихся в заданной точке. В другом вектор определяется как линейный дифференциальный оператор, который действует на функции. Несмотря на то что такие определения весьма изящны и эффективны, мне кажется, что начинать с таких определений — не самый лучший способ. Действительно, все определения эквивалентны, поэтому не имеет значения, что мы выбрали не самый изящный способ.

Пример 6

Векторное поле

$$v = x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \quad (26.15)$$

имеет интегральные кривые

$$\gamma_x: s \mapsto (a \cos s, a \sin s), \quad (26.16)$$

которые изображены на рис. 26.5. Их можно записать также в полярных координатах; тогда интегральные кривые имеют вид

$$\gamma_y: s \mapsto [R(s), \Theta(s)] = (a, s), \quad (26.17)$$

а векторное поле дается выражением

$$v = \frac{\partial}{\partial \theta}. \quad (26.18)$$

Оно изображено на рис. 26.6. Мы используем полярные координаты, без всякого предубеждения изображая их под прямым углом, как любые другие координаты. Помните о ковариант-

ности многообразий! Мы не должны выделять какую-либо систему координат.

Ковариантность нашего представления касательного пространства проявляется и в наших графических построениях. Мы изображали касательные пространства непосредственно на карте многообразия. Это не совсем верно. Касательные векторы получаются в результате предельного перехода и лежат в своем собственном векторном пространстве. Изображение их на карте многообразия действительно дает нам правильное представление об их линейной структуре. Однако совпадение конца вектора с какой-нибудь точкой многообразия не имеет никакого значения. Такое совпадение важно только в том случае, если мы умножим вектор на ε , а затем устремим ε к нулю.

Ковариантность

Пример 7

Посмотрите снова на рис. 26.5 и 26.6. На первом рисунке конец вектора a совпадает с точкой $(x, y) = (1, 1)$, а второй — с совершенно другой точкой $(r, \theta) = (1, 1)$.

Представление касательного пространства непосредственно на карте многообразия — весьма полезный метод, и необходимо научиться им пользоваться. Идея метода поясняется на рис. 26.7. Вложенные поверхности могут служить примерами многообразий, для которых касательные пространства на самом деле являются касательными плоскостями. При использовании описанного выше представления мы проецируем касательные плоскости обратно на многообразие. Понятно, что эта операция неоднозначна.

Наиболее важное понятие, которое мы теряем при переходе от линейного пространства к многообразию, — это понятие параллельности на расстоянии. Даже в случае многообразия с метрикой нельзя сравнивать направления векторов в двух различных точках. Например, на рис. 26.6 векторное поле выглядит так, как будто оно постоянно по направлению, тогда как про рис. 26.5 этого сказать нельзя. Таким образом, в случае многообразия не существует понятия постоянного векторного поля.

Векторы в различных точках

Другое важное свойство многообразия состоит в том, что мы, вообще говоря, не можем покрыть все многообразие одной координатной картой. Здесь могут возникнуть трудности, аналогичные тем, с которыми мы встретились в рассмотрен-

Может потребоваться несколько карт

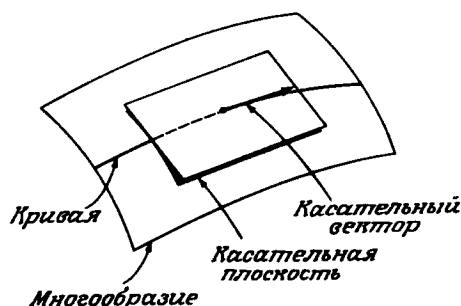


Рис. 26.7

[Это обстоятельство будет обсуждаться далее и использоваться в разд. 38, когда мы перейдем к рассмотрению топологии 3-сферы.]

ном примере с полярными координатами. Когда мы имеем дело с многообразиями, такие недостатки, как правило, устранить нельзя. Но они вносят не больше осложнений, чем полярные координаты. Математики формально справляются с названными трудностями, используя несколько различных карт, заданных таким образом, чтобы каждая точка многообразия была внутренней точкой по крайней мере одной карты. Мы же сможем легко обходить эти трудности, руководствуясь здравым смыслом.

Пространство поворотов

[Важно понять, что повороты являются линейными преобразованиями и, следовательно, могут быть представлены матрицами 3×3 . У любой такой матрицы имеется по крайней мере один действительный собственный вектор, который определяет ось вращения.]

Пример 8

Множество всех возможных поворотов твердого тела в трехмерном пространстве может служить важным примером многообразия. Карты для этого многообразия можно отыскивать следующим образом. Любой поворот оставляет фиксированной ось вращения. С точки зрения геометрии это утверждение не очевидно, но его можно доказать алгебраически. Проведем такое доказательство. Чтобы задать поворот, прежде всего найдем соответствующую ось. Затем найдем величину поворота вокруг оси. Для этого нужно договориться о том, какой знак приписывать каждому повороту. Будем пользоваться правилом правой руки, которое поясняется на рис. 26.8. Поворот представляется вектором, направленным вдоль оси поворота; длина вектора равна углу поворота ($< \pi$), а направление поворота определяется с помощью правила правой руки. Наглядный пример поворота вместе с соответствующим вектором поворота изображен на рис. 26.9. Векторы поворота являются элементами $\mathbb{R}^3$ и, следовательно, могут использоваться для построения карт.



Рис. 26.8

Пример, показывающий, каким образом поворот может быть представлен с помощью вектора.

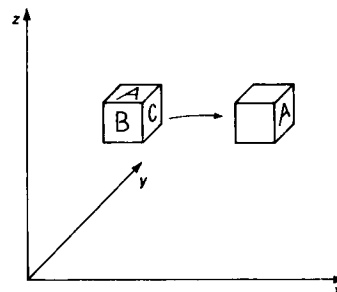


Рис. 26.9

Поворот, который представляется вектором $(2\pi/3\sqrt{3})[(\partial/\partial x) + (\partial/\partial y) + (\partial/\partial z)]$.

Первая карта отображает все повороты на углы меньше π в шар

$$x^2 + y^2 + z^2 < \pi^2. \quad (26.19)$$

Эта карта содержит все повороты, кроме поворотов, которые представляются векторами с длиной π . Назовем эти повороты *переворотами*. Поведение многообразия вблизи точек переворотов несколько более сложно. Различные точки на поверхности $x^2 + y^2 + z^2 = \pi^2$ не обязательно соответствуют разным поворотам. Диаметально противоположные точки представляют один и тот же поворот. Чтобы найти карту, изображающую некоторый переворот F , мы должны взять некоторую открытую окрестность точки F (рис. 26.10). Точки, принадлежащие этой окрестности и лежащие вне поверхности (26.19), соответствуют поворотам на углы $\pi + \alpha$. Такие повороты эквивалентны поворотам на углы $2\pi - \alpha$ в противоположном направлении. Повороты на углы $2\pi - \alpha$ изображаются точками, которые лежат внутри поверхности (26.19) с противоположной стороны. Таким образом, наша окрестность представляет собой открытое множество, каждая точка внутри нее представляет только один поворот, и ни один поворот не появляется в ней больше одного раза. Следовательно, это карта.

Такую карту можно построить для любого переворота. Следовательно, множество всех поворотов в трехмерном пространстве представляет собой трехмерное многообразие. При работе с ним удобнее использовать только первую карту, рас-

Перевороты

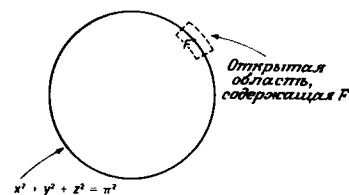


Рис. 26.10.

Пространство векторов поворотов и открытая карта, включающая переворот F . Представлен случай двух измерений.

ширив ее за поверхность

$$x^2 + y^2 + z^2 = \pi^2, \quad (26.20)$$

чтобы покрыть точки переверотов. При этом необходимо помнить, что некоторые повороты будут представлены двукратно.

Многообразия в общей теории относительности

[Только после того как вы немного поработаете с касательными пространствами, они станут для вас привычными и естественными. Разд. 27 и 28 дают вам такую возможность.]

В качестве модели пространства-времени в общей теории относительности используется многообразие, оснащенное метрическим тензором. Как мы скоро увидим, принять такую точку зрения нас заставляет наблюдаемое нарушение постулата свободных частиц. При наличии гравитационных полей инерциальные системы отсчета существуют только в бесконечно малых областях пространства-времени. Вот почему мы вынуждены использовать весь аппарат многообразий!

ЗАДАЧИ

26.1. (20) Цилиндр можно представить в виде полосы $0 \leq x \leq 1$ в плоскости (x, y) при условии, что для всех y точки $(0, y)$ и $(1, y)$ считаются одной и той же точкой. Какие из приведенных ниже векторных полей непрерывны в цилиндре:

а) $\frac{\partial}{\partial x}$,

б) $\frac{\partial}{\partial y}$,

в) $\sin 2\pi x \frac{\partial}{\partial y}$,

г) $x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x}$?

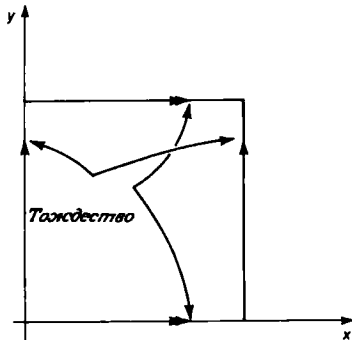


Рис. 26.11

26.2. (12) Тор можно представить в виде части плоскости $0 \leq x, y \leq 1$ при условии, что точки, лежащие на противоположных сторонах этого квадрата, отождествляются (рис. 26.11). Непрерывно ли на торе векторное поле

$$y \frac{\partial}{\partial x} ?$$

26.3. (20) Сколько координатных карт потребуется, чтобы покрыть тор таким образом, чтобы каждая точка тора являлась внутренней точкой хотя бы одной координатной карты?

26.4. (28) Покажите, что в случае векторного пространства определения касательного вектора, основанные на представлении о многообразии (разд. 17) и на представлении о векторном пространстве (разд. 12), согласуются.

26.5. (30) Рассмотрите множество всех окружностей максимально большого радиуса на сфере. Покажите, что это множество является многообразием (оно называется действительным проективным двумерным пространством, RP^2), и найдите для него удобное представление.

26.6. (28) Покажите, что в пространстве поворотов (см. последний пример этого раздела) любая плоскость, проходящая через начало координат, вырезает подмножество, которое представляет рассмотренное выше многообразие.

26.7. (38) Прочитайте пример, приведенный в книге [30, с. 2] и проанализируйте его. См. также *Am. J. Phys.* **47**, 379 (1979).

III. Гравитация

Полуклассические модели

Теперь мы располагаем набором математических «инструментов», необходимых для изложения теории тяготения, совместимой со специальной теорией относительности. Но не будем торопиться. Прежде чем приступить к такой сложной проблеме, как гравитация, немного попрактикуемся в использовании этих инструментов для решения других задач, которые понадобятся нам в дальнейшем. Как уже говорилось в разд. 19, наша математическая модель частицы — это высокочастотный волновой пакет. Такая модель называется полуклассической, ибо применяемый в ней способ описания лежит где-то посередине между методами классической и квантовой механики. Доказательство же того, что поведение составных систем, построенных из множества отдельных частиц (например, таких систем как Земля), следует тем же самым правилам, выходит далеко за пределы наших возможностей и фактически до сих пор находится в стадии разработки.

Изучение волновых пакетов приведет нас к конкретной реализации идеи о встроенных часах, которая в гл. I рассматривалась как некоторое первичное понятие. Кроме того, мы построим модель, воплощающую в себе представление о лоренц-инвариантности. В результате окажется, что такое простое явление, как распространение волн на глубокой воде, имеет непосредственное отношение к релятивистской симметрии.

Свободные частицы

Шаг за шагом мы наконец доберемся до гравитации и прежде всего займемся однородными полями тяготения. В итоге выяснится, что специальной теорией относительности по-прежнему можно пользоваться даже в присутствии таких полей, изменив определение свободной частицы. Теперь под свободной частицей нужно понимать частицу, на которую не действуют никакие силы, помимо, возможно, силы тяготения. Так как в малых областях все гладкие гравитационные поля выглядят однородными, такое расширение сферы применимости тео-

рии означает, что локально геометрию пространства-времени можно приближенно считать геометрией специальной теории относительности, т. е. геометрией Минковского. Ясно, что такое рассмотрение без труда можно распространить на неоднородные поля тяготения — достаточно лишь предположить, что метрический тензор не остается постоянным, а является функцией положения.

В конце главы мы обсудим несколько простых приложений теории тяготения. Сначала мы разработаем некоторые дополнительные методы, необходимые для изучения распространения волновых пакетов в неоднородных полях. Затем с их помощью покажем, что используемая в ОТО модель пространства-времени в окрестности массивного тела описывает знакомые нам гравитационные эффекты. Мы кратко остановимся на представлении о кривизне пространства-времени и на уравнениях Эйнштейна, однако их глубокий количественный анализ выходит за рамки этой книги.

27. Распространение волновых пакетов

Модель, с помощью которой мы будем описывать динамику частиц, основывается на представлении о волновом пакете. Чтобы получить простой и удобный способ описания важнейшей характеристики движения волновых пакетов, их групповой скорости¹⁾ или скорости распространения энергии, необходимо уделить внимание геометрическому смыслу этой величины. Проведенное в разд. 19 обсуждение страдает определенным недостатком, ибо координата времени в нем играет выделенную роль. В этом и в двух следующих разделах мы ограничимся изучением прямолинейного движения волновых пакетов, распространяющихся в однородных средах, где однородность следует понимать в 4-мерном смысле. В общем случае движение будет рассмотрено в разд. 30.

Начнем с описания волнового уравнения с помощью дисперсионного уравнения. В случае двух измерений дисперсионное уравнение имеет вид

$$W(k, \omega, x, t) = 0. \quad (27.1)$$

При распространении пакета в среде, однородной как в пространстве, так и во времени, функция W не должна явно зави-

Дисперсионное уравнение

[Для простоты мы обсуждаем только двумерный случай. Обобщение очевидно.]

¹⁾ См. примечание на стр. 149. — *Прим. ред.*

сеть от x и t . Каждый высокочастотный волновой пакет всюду обладает волновым числом $k(x, t)$ и частотой $\omega(x, t)$, удовлетворяющими соотношению

$$W[k(x, t), \omega(x, t)] = 0, \quad (27.2)$$

частные производные от которого по x и t также должны быть равны нулю. Эти производные можно найти с помощью цепного правила дифференцирования сложной функции:

$$\frac{\partial W}{\partial k} \frac{\partial k}{\partial x} + \frac{\partial W}{\partial \omega} \frac{\partial \omega}{\partial x} = 0, \quad (27.3)$$

$$\frac{\partial W}{\partial k} \frac{\partial k}{\partial t} + \frac{\partial W}{\partial \omega} \frac{\partial \omega}{\partial t} = 0. \quad (27.4)$$

Напомним (см. стр. 155), что поскольку k и ω — компоненты градиента фазы, то

$$\frac{\partial k}{\partial t} = \frac{\partial \omega}{\partial x}. \quad (27.5)$$

Следовательно, уравнения (27.3) и (27.4) можно представить в виде

$$\frac{\partial W}{\partial k} \frac{\partial k}{\partial x} + \frac{\partial W}{\partial \omega} \frac{\partial k}{\partial t} = 0, \quad (27.6)$$

$$\frac{\partial W}{\partial k} \frac{\partial \omega}{\partial x} + \frac{\partial W}{\partial \omega} \frac{\partial \omega}{\partial t} = 0. \quad (27.7)$$

Групповая скорость

Отсюда видно, что величины k и ω остаются постоянными в направлении вектора

[Повторите ту часть разд. 25, где обсуждались характеристики.]

$$\frac{\partial W}{\partial k} \frac{\partial}{\partial x} + \frac{\partial W}{\partial \omega} \frac{\partial}{\partial t}, \quad (27.8)$$

который называется вектором групповой скорости.

Пример 1

Волны на воде

Для волн на глубокой воде выполняется соотношение

$$W(k, \omega) = \omega^4 - k^2 = 0, \quad (27.9)$$

так что k и ω постоянны в направлении

$$4\omega^3 \frac{\partial}{\partial t} - 2k \frac{\partial}{\partial x}. \quad (27.10)$$

Ранее на стр. 155 мы нашли вектор групповой скорости для волн на глубокой воде:

$$\frac{\partial}{\partial x} - 2\omega \frac{\partial}{\partial t}. \quad (27.11)$$

Обратите внимание, что эти векторы имеют одинаковые направления в пространстве-времени.

Групповая скорость, как она определялась до сих пор, — это лишь то направление в пространстве-времени, вдоль которого должны оставаться постоянными волновое число и частота. Длина же вектора групповой скорости остается неопределенной. Имеется несколько подходящих нормировок длины этого вектора. Одна из них сводится к требованию, что компонента $\partial/(\partial t)$ должна быть равна единице. Совершенно очевидно, однако, что тогда координата t становится выделенной. Существует также инвариантная нормировка

Нормировка

$$d\theta \cdot v = 1, \quad (27.12)$$

где $d\theta$ — градиент фазы и v — вектор групповой скорости, изображенные на рис. 27.1. Если воспользоваться формулой (27.8) для определения направления вектора групповой скорости, а в качестве условия нормировки выбрать (27.12), то мы придем к соотношению

$$v = \frac{\frac{\partial W}{\partial k} \frac{\partial}{\partial x} + \frac{\partial W}{\partial \omega} \frac{\partial}{\partial t}}{k \frac{\partial W}{\partial k} + \omega \frac{\partial W}{\partial \omega}}. \quad (27.13)$$

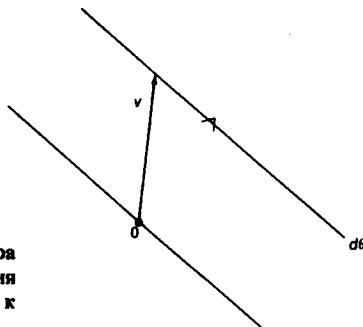


Рис. 27.1.

Естественная нормировка вектора групповой скорости. Линии уровня 1-формы $d\theta$ в 2π раз ближе друг к другу, чем гребни волны.

Асимптоты

Особый интерес представляет ситуация, когда знаменатель дроби (27.13) обращается в нуль. В этом случае дисперсия отсутствует и волны с различными частотами распространяются с одинаковой скоростью. Таким свойством обладают, например, световые волны. Однако волновые пакеты, представляющие наши частицы, ведут себя совершенно иначе.

Волновая диаграмма

Каждой паре величин k и ω , удовлетворяющих интересующему нас дисперсионному уравнению, соответствует единственный вектор групповой скорости, описывающий распространение волнового пакета в пространстве-времени. Другими словами, каждой точке, удовлетворяющей дисперсионному уравнению, соответствует единственный касательный вектор. Таким образом, в касательном пространстве есть кривая, каждая точка которой соответствует дисперсионному уравнению.

[Продолжение примера 1.]

В случае волн на глубокой воде кривую, соответствующую дисперсионному уравнению, можно параметризовать с помощью волнового числа k :

$$d\theta = \sqrt{|k|} dt + k dx. \quad (27.14)$$

Каждому значению k соответствует вектор групповой скорости

$$v = \frac{4\omega^3 \frac{\partial}{\partial t} - 2k \frac{\partial}{\partial x}}{4\omega^4 - 2k^2}. \quad (27.15)$$

С помощью дисперсионного уравнения эту формулу можно упростить:

$$v = \frac{2}{\sqrt{|k|}} \frac{\partial}{\partial t} - \frac{1}{k} \frac{\partial}{\partial x}. \quad (27.16)$$

В таком виде она описывает однопараметрическое семейство касательных векторов, определяемых дисперсионным уравнением.

При изображении систем векторов удобнее всего провести их из начала системы координат, а концам поставить в соответствие отдельные точки. Тогда каждый вектор будет изображаться точкой в касательном пространстве. В этом пространстве можно ввести координаты, которые мы договоримся обозначать $\dot{x}$ и $\dot{t}$. Соотношение

$$v = \dot{x} \frac{\partial}{\partial x} + \dot{t} \frac{\partial}{\partial t} \quad (27.17)$$

показывает, как от координат $(\dot{x}, \dot{t})$ перейти к вектору v , а соотношение

$$d\theta = k dx + \omega dt \quad (27.18)$$

показывает, как от координат k (волновое число) и ω (частота) перейти к 1-форме градиента фазы.

Наше однопараметрическое семейство векторов групповой скорости описывается уравнениями

$$\begin{aligned} \dot{x} &= -\frac{1}{k}, \\ \dot{t} &= \frac{2}{\sqrt{|k|}}. \end{aligned} \quad (27.19)$$

Исключив параметр k , получим следующее неявное уравнение для множества векторов групповой скорости:

$$\dot{t}^2 + 4\dot{x} = 0 \quad (27.20)$$

для одной ветви дисперсионного соотношения и

$$\dot{t}^2 - 4\dot{x} = 0 \quad (27.21)$$

для другой ветви. Эти множества векторов изображены на рис. 27.2.

В дальнейшем множество всех нормированных векторов групповой скорости окажется чрезвычайно полезным для нас инструментом. Хотя эта конструкция уже использовалась математиками в исследованиях по вариационному исчислению, она никогда прежде не применялась физиками для изучения диспергирующих волн. Назовем ее *волновой диаграммой*. Теперь продемонстрируем одно важное геометрическое соотношение, на котором, собственно, и зиждется полезность волновой диаграммы. 1-форма градиента фазы, соответствующая любому вектору групповой скорости, касательна к волновой диаграмме в точке, изображающей данный вектор. Указанная связь иллюстрируется рис. 27.3.

Волновая диаграмма для волн на глубокой воде, отражающая упомянутую связь, изображена на рис. 27.4. Мы сразу же видим, что волновой пакет движется медленнее (фактически в два раза), чем гребни волны. На рис. 27.5 представлена диаграмма пространства-времени, на которой показан волновой пакет,

[Достаточно беглого взгляда на уравнения интегральных кривых в разд. 25, чтобы стало ясно, почему выбраны именно эти обозначения. Как легко обнаружит читатель, знакомый с лагранжевой механикой, это то же самое, что взять q и $\dot{q}$ в качестве независимых координат.]

[Продолжение примера 1.]

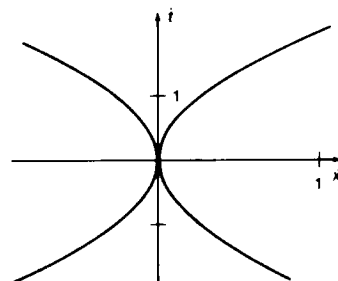


Рис. 27.2
Волновая диаграмма для волн на глубокой воде.

[См., например, [4].]

Дуальность

[Продолжение примера 1.]

[Не забывайте, что в касательном пространстве есть единственный нулевой вектор. Это очень существенный момент в нашей диаграмме.]

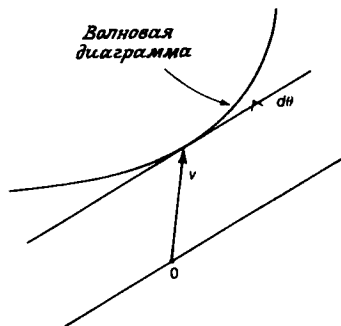


Рис. 27.3

Установление связи между групповой скоростью и градиентом фазы с помощью волновой диаграммы.

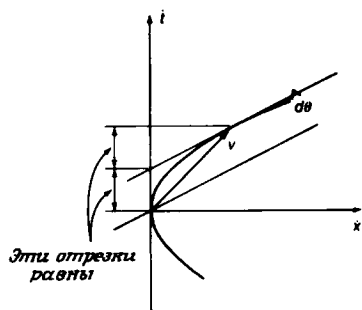


Рис. 27.4

Градиент фазы и групповая скорость для волн на воде.

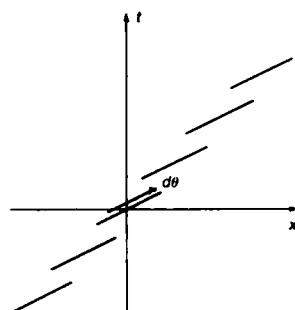


Рис. 27.5

Пространственно-временная диаграмма для волнового пакета, соответствующая волновой диаграмме на рис. 27.4. Масштаб 1:10 по отношению к предыдущему рисунку. Обратите внимание на градиент фазы в центре диаграммы. Наклонные линии — гребни волн. На конечность размеров волнового пакета указывает ограниченная длина мировых линий этих гребней. См. рис. 19.6 и 19.7.

представленный на рис. 27.4.

[Это доказательство можно пропустить без серьезного ущерба для понимания дальнейшего материала.]

Чтобы показать, что линии уровня 1-формы $d\theta$ на самом деле касаются волновой диаграммы, рассмотрим вектор групповой скорости v и ассоциированный с ним градиент фазы $d\theta$. В силу нашего условия нормировки имеем

$$d\theta \cdot v = 1; \quad (27.22)$$

это означает, что единичный контур, связанный с 1-формой $d\theta$, проходит через конец вектора v . Рассмотрим близкий к исходному вектор групповой скорости v' и соответствующий ему градиент фазы $d\theta'$, который можно записать в виде

$$d\theta' = (k + \Delta k) dx + (\omega + \Delta\omega) dt. \quad (27.23)$$

Новый вектор групповой скорости

$$v' = v + \Delta v \quad (27.24)$$

и 1-форма $d\theta'$ должны удовлетворять соотношению

$$d\theta' \cdot v' = 1. \quad (27.25)$$

Отсюда в пределе при Δk , $\Delta\omega$ и Δv , стремящихся к нулю, следует (после отбрасывания членов, квадратичных по приращениям)

$$(\Delta k dx + \Delta\omega dt) \cdot v + d\theta \cdot \Delta v = 0. \quad (27.26)$$

Выпишем первый член из левой части этого соотношения и подставим в него явное выражение для вектора v :

$$\frac{(\Delta k dx + \Delta\omega dt) \cdot \left(\frac{\partial W}{\partial k} \frac{\partial}{\partial x} + \frac{\partial W}{\partial \omega} \frac{\partial}{\partial t} \right)}{\left(k \frac{\partial W}{\partial k} + \omega \frac{\partial W}{\partial \omega} \right)} = \frac{\Delta k \frac{\partial W}{\partial k} + \Delta\omega \frac{\partial W}{\partial \omega}}{\left(k \frac{\partial W}{\partial k} + \omega \frac{\partial W}{\partial \omega} \right)}. \quad (27.27)$$

Градиент фазы $d\theta'$ должен удовлетворять дисперсионному уравнению

$$W(k + \Delta k, \omega + \Delta\omega) = 0, \quad (27.28)$$

откуда после разложения в ряд Тейлора получаем

$$\frac{\partial W}{\partial k} \Delta k + \frac{\partial W}{\partial \omega} \Delta\omega = 0. \quad (27.29)$$

Таким образом, первый член в левой части соотношения (27.26) равен нулю, следовательно,

$$d\theta \cdot \Delta v = 0. \quad (27.30)$$

Как нетрудно заметить, взглянув на рис. 27.6, это условие как раз и означает, что связанный с $d\theta$ единичный контур касателен к волновой диаграмме.

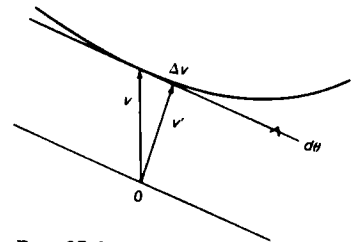


Рис. 27.6

Волновая диаграмма очень полезна при решении различных задач теории диспергирующих волн. Она является обобщением гиперболы, возникающей в СТО для соответствующего волнового уравнения.

Гипербола СТО

Пример 2

Построим волновую диаграмму для уравнения

$$\frac{\partial^2 \psi}{\partial t^2} - \frac{\partial^2 \psi}{\partial x^2} + \psi = 0. \quad (27.31)$$

Теперь дисперсионное уравнение имеет вид

$$W = \omega^2 - k^2 - 1 = 0, \quad (27.32)$$

а нормированный вектор групповой скорости можно записать в форме

$$\begin{aligned} v &= \frac{\frac{\partial W}{\partial \omega} \frac{\partial}{\partial t} + \frac{\partial W}{\partial k} \frac{\partial}{\partial x}}{\omega \frac{\partial W}{\partial \omega} + k \frac{\partial W}{\partial k}} = \\ &= \frac{\omega \frac{\partial}{\partial t} - k \frac{\partial}{\partial x}}{\omega^2 - k^2} = \\ &= \omega \frac{\partial}{\partial t} - k \frac{\partial}{\partial x}. \end{aligned} \quad (27.33)$$

Выражая вектор v через его компоненты $\dot{t}$ и $\dot{x}$, получаем

$$v = \dot{t} \frac{\partial}{\partial t} + \dot{x} \frac{\partial}{\partial x}, \quad (27.34)$$

так что

$$\begin{aligned} \dot{t} &= \omega, \\ \dot{x} &= -k. \end{aligned} \quad (27.35)$$

Следовательно, волновая диаграмма определяется уравнением

$$(\dot{t})^2 - (\dot{x})^2 = 1, \quad (27.36)$$

действительно задающим гиперболу.

Градиент фазы и вектор групповой скорости движущегося волнового пакета изображены на рис. 27.7. Обратите внимание на специфический наклон волновых фронтов по отношению к вектору групповой скорости. Это тот самый наклон, о кото-

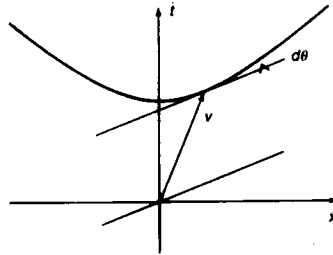


Рис. 27.7

Градиент фазы и групповая скорость волнового пакета, который описывается обычным волновым уравнением.

ром шла речь при обсуждении одновременности. Таким образом, наше определение оказалось полезным. Если бы мы начали с уравнений Максвелла и построили аналогичным способом волновую диаграмму, то получили бы гиперболу. Тогда определение одновременности появилось бы совершенно естественным образом. Однако такой подход слишком сложен для того, чтобы с него можно было начинать изложение теории.

В случае прямолинейного движения волнового пакета в однородной среде между пространством-временем (x, t) и касательным пространством $(\dot{x}, \dot{t})$ имеется лишь небольшое различие (за исключением положения начала координат). При переходе к неоднородным полям, когда дисперсионное уравнение явно зависит от x и t , волновую диаграмму следует строить в линейном пространстве, а для этого необходимо касательное пространство. С касательным пространством легче работать, чем с дуальным пространством или с градиентами фазы, ибо касательное пространство — это линейная аппроксимация самого многообразия. Поэтому волновую диаграмму удобнее использовать, чем дисперсионное уравнение.

Роль метрического тензора пространства-времени сводится к описанию волновой диаграммы. Составляющие векторов групповой скорости для волнового уравнения (27.31), обладающего лоренцевой симметрией СТО, должны удовлетворять соотношению

$$\dot{t}^2 - \dot{x}^2 = 1, \quad (27.37)$$

которое можно записать в виде

$$\mathcal{N} \cdot (t, v) = -1, \quad (27.38)$$

Прямолинейное движение

[Напомним, что касательное пространство — это пространство касательных векторов в точке.]

Метрический тензор

где $\mathcal{N}$ — метрический тензор пространства Минковского

$$\mathcal{N} = dx^2 - dt^2. \quad (27.39)$$

Кроме того, ассоциированный с вектором скорости v градиент фазы $d\theta$ можно задать с помощью формулы

$$d\theta = -\mathcal{N} \cdot v. \quad (27.40)$$

[Объяснение этого обозначения дано на стр. 165.]

Амплитуда волны

Теперь, не вдаваясь в подробности, выясним, почему движение максимума волнового пакета происходит в направлении групповой скорости. Мы получили дисперсионное уравнение, оставив в разложении для почти плоских волн

$$\psi \sim A \cos \theta \quad (27.41)$$

[Мы продолжаем рассуждение, начатое в разд. 19.]

только главные члены. Рассмотрим теперь случай, когда не только k и ω , но и амплитуда A медленно меняются в зависимости от координат x и t . Присутствие в дисперсионном уравнении такой величины, как ω^4 , отражает наличие в исходном дифференциальном уравнении в частных производных члена $(\partial^4 \psi)/(\partial t^4)$. При четырехкратном дифференцировании фазы θ — величины, изменяющейся быстрее всех остальных, — в дисперсионном уравнении появляется ω^4 . Члены следующего, более высокого порядка малости появляются, когда одно из четырех дифференцирований относится к любой из величин k , ω или A . В последнем случае мы будем иметь член $4\omega^3[(\partial A)/(\partial t)]$, где множитель ω^3 возникает в результате выполнения трех оставшихся дифференцирований фазы θ . Коэффициент 4 объясняется тем, что к A может относиться любое из четырех дифференцирований. Таким образом, в результате дифференцирования амплитуды A получаем уравнение

$$\frac{\partial W}{\partial \omega} \frac{\partial A}{\partial t} + \frac{\partial W}{\partial k} \frac{\partial A}{\partial x} =$$

$$= (\text{Некоторое известное выражение}). \quad (27.42)$$

Уравнение (27.42) как раз и показывает, что поведение величины A определенным образом связано с направлением вектора групповой скорости. Так как волны могут «растягиваться», их амплитуда не обязательно должна оставаться постоянной. Основываясь на вышеприведенной схеме, можно было бы предпринять и более тщательный анализ рассматриваемой проблемы. Однако такой анализ не входит в наши планы, и мы ограничимся проведенным, в какой-то мере интуитивным, обсуждением.

Основная идея заключается в том, что каждое касательное пространство обладает своей волновой диаграммой. Иногда можно подобрать такие координаты, в которых волновые диаграммы одинаковы во всех касательных пространствах. В большей части наших примеров дело обстоит именно таким образом. Важно следующее: в каждом касательном пространстве должно существовать правило построения волновой диаграммы. Волновая диаграмма — это хорошо определенный геометрический объект, но не тензор и не обязательно линейный оператор. В этом отношении особняком стоит ОТО. Соответствующую ей волновую диаграмму можно описать на тензорном языке.

ЗАДАЧИ

27.1. (18) Введите подходящие единицы измерения и дайте численные оценки для ситуации, изображенной на рис. 27.4.

27.2. (32) Параболу можно построить как огибающую семейства прямых, связанных с фиксированной точкой S и осью L (рис. 27.8). Опишите это построение на нашем геометрическом языке, рассматривая прямые как ко векторы. Огибающая — это и есть волновая диаграмма. Докажите, что она на самом деле представляет собой параболу. См. задачу 23.1.

27.3. (21) Начертите волновую диаграмму для дисперсионного уравнения

$$\omega = k - k^3.$$

27.4. (21) Дисперсионное уравнение для капиллярных волн на глубокой воде имеет вид

$$\omega^2 = k^3.$$

Начертите соответствующую волновую диаграмму.

27.5. (21) Постройте волновую диаграмму для дисперсионного уравнения

$$\omega = k(1 + \alpha k^2).$$

27.6. (19) Докажите, что для двумерной пространственно-временной диаграммы скорость волнового пакета равна $\partial\omega/\partial k$, а скорость распространения гребней волн равна ω/k .

28. Теория относительности и волны на воде

Теория распространения волн на поверхности воды обнаруживает поразительное сходство со специальной теорией относи-

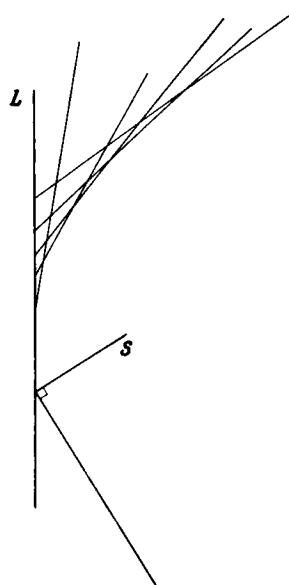


Рис. 27.8

[В этом и следующих разделах рассматриваются довольно интересные вопросы. Их обсуждение позволит вам набить руку в обращении с векторами, 1-формами и волновыми диаграммами. Однако полученные результаты (разумеется, помимо приобретенного опыта) не найдут в дальнейшем особых применений.]

[Краткое введение в теорию волн на воде можно найти в Фейнмановских лекциях по физике [11]. Более полно эти вопросы рассматриваются в книге Лайтхилла [20] и в статье Синга [38].]

Дисперсионные уравнения

[В действительности при использовании обычных единиц измерения дисперсионное уравнение имеет вид $\omega^4 = g^2 k^2$.]

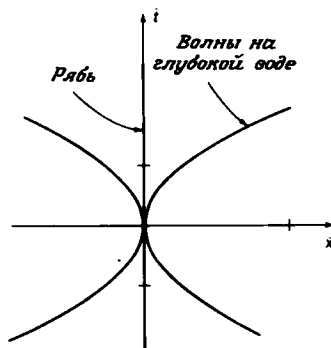


Рис. 28.1

Рябь

тельности. Поэтому прежде чем переходить к обсуждению тяготения, мы, чтобы попрактиковаться в использовании волновой диаграммы, коснемся некоторых вопросов физики волн на воде. В результате выяснится, что волны на глубокой воде обладают симметрией, удивительно похожей на лоренц-инвариантность. Такой подход к релятивистской симметрии приведет нас не только к новой точке зрения на СТО, но и позволит овладеть навыками в выполнении расчетов. Кроме того, используемые нами обозначения позволяют самым естественным и очевидным образом переходить от одной системы координат к другой, тогда как этот вопрос обычно вызывает затруднения и в большинстве книг по тензорному анализу ему уделяется много внимания.

В теории волн на воде, описываемых дисперсионным уравнением

$$\omega^4 = k^2, \quad (28.1)$$

не отражены многие свойства реальных волн. Эта теория учитывает лишь отношение между инерционными свойствами воды и гравитацией. В такой теории не принимается в расчет ни влияние дна, ни поверхностное натяжение. Кроме того, это линейизованная теория, так что в ней не рассматривается ни взаимодействие, ни разрушение волн. При учете поверхностного натяжения дисперсионное уравнение принимает вид

$$\omega^4 = k^2 \left(g + \frac{T}{\rho} k^2 \right)^2, \quad (28.2)$$

где T — коэффициент поверхностного натяжения, а ρ — плотность. Пренебрегая не представляющими интереса постоянными, мы получаем дисперсионное уравнение

$$\omega^4 = k^2 (1 + \alpha k^2)^2, \quad (28.3)$$

которое описывает волны с учетом влияния поверхностного натяжения. При больших k инерция жидкости уравнивается уже не силой тяжести, а самим поверхностным натяжением. Волны, описываемые уравнением (28.3), называются рябью. Рябь распространяется намного медленнее волн на глубокой воде и имеет малую длину волны. Мы воздержимся от сколько-нибудь детального описания ряби и представим их на волновой диаграмме в виде единственной вертикальной линии, как на рис. 28.1. Этот рисунок соответствует предельному случаю пренебрежимо малого коэффициента поверхностного натя-

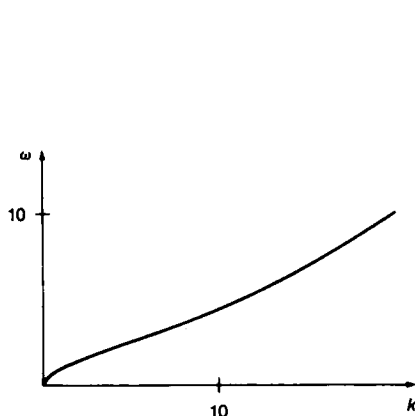


Рис. 28.2

Дисперсионное уравнение для волн на воде при учете поверхностного натяжения. Однако здесь не принято во внимание влияние дна, связанное с конечностью глубины слоя воды. Кривая построена для $\alpha = 0,01$.

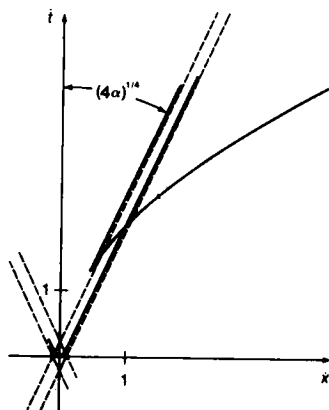


Рис. 28.3

Волновая диаграмма для дисперсионного уравнения, изображенного на рис. 28.2. Обратите внимание на множество дополнительных элементов, появившихся на этой волновой диаграмме.

жения. Точное дисперсионное уравнение показано на рис. 28.2, а соответствующая волновая диаграмма, которая, естественно, имеет более сложный вид, — на рис. 28.3. В дальнейшем мы будем использовать слегка модифицированную теорию волн на воде, которая учитывает волны ряби, но не рассматривает никаких особенностей их движения помимо того, что их групповая скорость принимается равной нулю.

Теперь перейдем к изучению теории волн на воде. Вообразите, что вы летите над поверхностью воды, скажем, в гондole дирижабля и видите только поверхность воды и распространяющиеся по ней волны. Дирижабль движется над водой с неизменной скоростью. Тогда возникает ряд вопросов. Например: можно ли определить скорость дирижабля только путем наблюдений волн на воде? Интересно, что такая теория волн на воде оказывается удивительно похожей на СТО. Чтобы в этом убедиться, рассмотрим наблюдателя в гондole дирижабля с тех же позиций, с которых в гл. I рассматривался наблюдатель в пространстве-времени.

Начнем с обсуждения волновых пакетов, образованных волнами на воде. В нашей теории они будут играть роль «свободных частиц». (Думаю, что в аналогичной ситуации Синг не

Наблюдения

Свободные частицы

Часы

стал бы возражать против названия «гидроны» для таких частиц.) Разумеется, подобные «частицы» нельзя считать ни точечными, ни обычными частицами, которыми оперирует ньютонова механика. А что можно сказать о часах? Подходя к вопросу с точки зрения наблюдателя, находящегося в гондole дирижабля, мы можем выяснить, что в данном случае следует понимать под встроенными часами. Это не должны быть обычные физические часы, использование которых могло бы лишь сбить с толку. Действительно, принцип действия наших часов должен быть связан с волнами на воде. Такими часами мы будем располагать, если станем следить за тем волновым пакетом, групповая скорость которого равна скорости дирижабля. Этот волновой пакет будет сопутствовать наблюдателю как единое целое. «Приготовить» такой пакет очень просто — достаточно бросить с дирижабля в воду какой-нибудь предмет и немного подождать. Через некоторое время волны, групповая скорость которых отличается от скорости наблюдателя, уйдут, а сопутствовать ему, естественно, будут лишь волны с групповой скоростью, совпадающей со скоростью дирижабля. Поскольку волны на воде испытывают дисперсию, гребни будут распространяться с иной скоростью, в данном случае превышающей групповую скорость. В результате вы увидите, что поверхность воды в любой точке, расположенной непосредственно под вами, совершает колебания вверх и вниз. Такие колебания и дают нам точное представление о том, что именно следует считать встроенными часами.

Пример

Рассмотрим наблюдателя, движущегося вправо со скоростью v . Волновая диаграмма для волн, распространяющихся вправо, описывается формулой

$$i^2 = 4\dot{x}, \quad (28.4)$$

или в параметрической форме

$$\begin{aligned} \dot{x} &= -\frac{1}{k}, \\ i &= \frac{2}{\sqrt{-k}}, \end{aligned} \quad (28.5)$$

где k — волновое число (в данном случае отрицательное). Вол-

новому числу k соответствует скорость

$$v = \frac{\dot{x}}{\dot{t}} = \frac{1}{2\sqrt{-k}}. \quad (28.6)$$

Таким образом, вектор групповой 4-скорости, соответствующий скорости v , можно представить в виде

$$\begin{aligned} \dot{x} &= 4v^2, \\ \dot{t} &= 4v. \end{aligned} \quad (28.7)$$

Наблюдатель, движущийся вместе с волнами, обнаружит, что новые гребни волны проходят мимо него через промежутки времени

$$\Delta t = 8\pi v. \quad (28.8)$$

[Снова эти надоевшие 2π ; см. разд. 19.]

Обратите внимание на рис. 28.4. На нем помимо всего прочего показан градиент фазы рассматриваемой волны

$$d\theta = \frac{1}{2v} dt - \frac{1}{4v^2} dx. \quad (28.9)$$

Чтобы завершить перечень первичных понятий, необходимо ввести специальный класс волн, который можно было бы назвать «световые сигналы». В качестве такого класса волн мы выберем рябь. Теперь у нас есть набор первичных элементов, сходный с набором элементов, заложенных в основу фундамен-

Световые сигналы

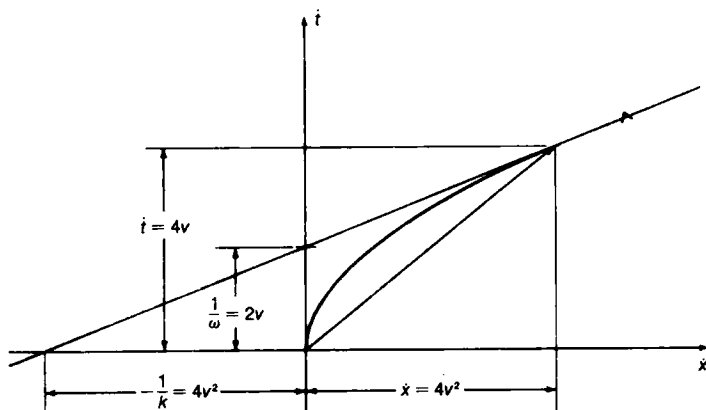


Рис. 28.4

Более подробное изображение волнового пакета, обладающего групповой скоростью v . Обратите внимание, что гребни волн движутся в два раза быстрее самого волнового пакета.

тальной структуры пространства-времени. Таким образом, распространение волн на воде дает нам еще одну реализацию интересующей нас абстрактной структуры.

Перевод терминов, описывающих волны на воде, на язык теории пространства-времени: волновой пакет, образованный волнами на воде — свободная частица; колебания поверхности воды, вызванные волнами, движущимися со скоростью, равной скорости наблюдателя — часы; система обычных координат (x, t) — инерциальная система отсчета; рябь — световые сигналы.

Поскольку наблюдатель движется с отличной от нуля скоростью, эти первичные элементы удовлетворяют всем постулатам, касающимся структуры пространства-времени, за исключением того, что вместо световых сигналов, распространяющихся как влево, так и вправо, теперь имеется единственная система волн ряби.

Инерциальные системы отсчета

Обычная система координат (x, t) образует инерциальную систему отсчета. Что можно сказать о канонической системе отсчета? Какие линейные преобразования должен выполнить данный наблюдатель, чтобы упростить волновую диаграмму? Предположим, что сделанных им наблюдений достаточно для построения полной волновой диаграммы. В двумерном случае линейное преобразование обладает четырьмя степенями свободы. Рассмотрим процедуру преобразования к канонической системе отсчета, основанную на выполнении последовательности из четырех простых преобразований, упрощающих волновую диаграмму данного наблюдателя.

Канонические системы отсчета

[Здесь стоит вернуться к разд. 6, где обсуждалась каноническая система отсчета в СТО.]

Будем считать теперь, что наблюдатель не движется, а покоится по отношению к волнам ряби.

На рис. 28.5 показана исходная ситуация, которую увидит такой наблюдатель. Обсудим преобразование к канонической системе отсчета именно этой диаграммы. Первое требование к канонической форме волновой диаграммы заключается в том, что мы должны перейти к системе отсчета, в которой наблюдатель покоится, т. е. к системе отсчета, где его мировая линия вертикальна, как на рис. 28.6. Второе требование к каноническому представлению сводится к выбору шкалы времени, приспособленной к ходу встроенных часов. Так, на рис. 28.7 точка пересечения волновой диаграммы с вертикальной осью

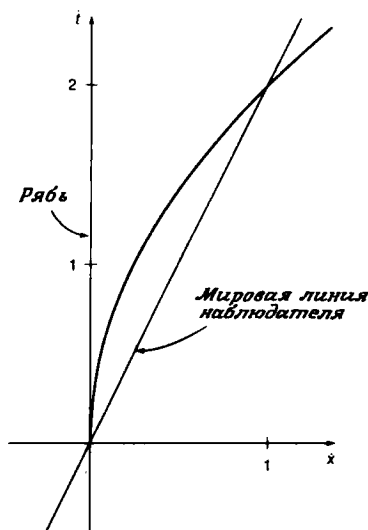


Рис. 28.5

Волновая диаграмма и мировая линия наблюдателя, движущегося со скоростью $1/2$.

соответствует единице времени. Далее, нужно добиться, чтобы волновая диаграмма пересекала вертикальную ось под прямым углом, как на рис. 28.8, — в этом состоит третье упрощение. Наконец, можно выбрать на горизонтальной оси такой масштаб, чтобы мировая линия для ряби шла под углом в 45° , как на рис. 28.9. Все эти диаграммы — эквивалентные представления волн на глубокой воде. Достоинства только что введенных упрощений проявятся в полной мере, когда мы перейдем к обсуждению симметрии волн на воде.

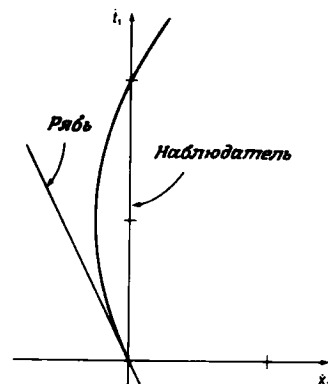
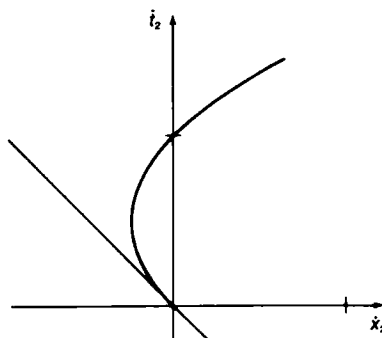


Рис. 28.6

В результате горизонтального сдвига мировая линия наблюдателя становится вертикальной.

Рис. 28.7

В результате растяжения вертикальной шкалы единичный пространственный интервал становится равным интервалу между началом координат и точкой пересечения волновой диаграммы с мировой линией наблюдателя.



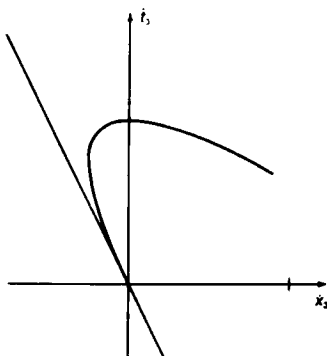


Рис. 28.8

В результате вертикального сдвига угол между волновой диаграммой и мировой линией наблюдателя становится прямым.

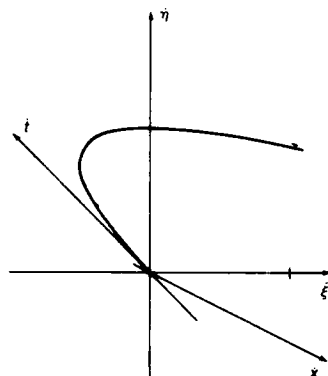


Рис. 28.9

В результате горизонтального сжатия угол наклона мировой линии ряби становится равным 45° .

Пример

Рассмотренные выше рисунки соответствуют точке зрения наблюдателя, движущегося со скоростью $v = 1/2$ в используемых безразмерных единицах. Первое из примененных нами преобразований — это сдвиг:

$$(\dot{x}, t) \mapsto \left(\dot{x} - \frac{t}{2}, t \right) \equiv (\dot{x}_1, t_1). \quad (28.10)$$

Уравнение, описывающее преобразованную волновую диаграмму, можно получить путем поочередного преобразования всех векторов, имеющих на этой диаграмме. Проще всего это сделать, исходя из параметрической формы

$$\begin{aligned} \dot{x} &= -\frac{1}{k}, \\ t &= \frac{2}{\sqrt{-k}} \end{aligned} \quad (28.11)$$

и применяя преобразования (28.10). В новых координатах $(\dot{x}_1, t_1)$ получаем

$$\begin{aligned} \dot{x}_1 &= -\frac{1}{k} - \frac{1}{\sqrt{-k}}, \\ t_1 &= \frac{2}{\sqrt{-k}}. \end{aligned} \quad (28.12)$$

Обычно кривую легче преобразовать, если она задана в параметрической форме, однако обсуждать ее свойства проще, когда ее уравнение записано в явном виде. В данном случае трудно исключить параметр k , что приводит к уравнению

$$4\dot{x}_1 = i_1(i_1 - 2). \quad (28.13)$$

Теперь волновая диаграмма пересекает ось i_1 в точке

$$i_1 = 2, \quad (28.14)$$

что согласуется с расчетами, проделанными на стр. 227. Соответствующая ситуация изображена на рис. 28.6. Второе преобразование заключается в сжатии в два раза оси i_1 ; в результате точка пересечения волновой диаграммы с осью окажется расположенной на единичной высоте. После выполнения преобразования

[Из уравнения (28.7) видно, что i -компонента вектора групповой скорости, соответствующая значению $v = 1/2$, равна 2 в согласии с (28.14).]

$$(\dot{x}_1, i_1) \mapsto \left(\dot{x}_1, \frac{i_1}{2}\right) \equiv (\dot{x}_2, i_2) \quad (28.15)$$

уравнение волновой диаграммы примет вид

$$\dot{x}_2 = i_2(i_2 - 1). \quad (28.16)$$

Соответствующая кривая показана на рис. 28.7.

Третье преобразование — это сдвиг, позволяющий сделать волновую диаграмму горизонтальной в точке ее пересечения с осью i_2 . Вблизи точки $\dot{x}_1 = 0$ имеем

$$\dot{x}_2 \approx (i_2 - 1), \quad (28.17)$$

откуда видно, что для приведения диаграммы к надлежащему виду необходимо произвести сдвиг

$$(\dot{x}_2, i_2) \mapsto (\dot{x}_2, i_2 - \dot{x}_2) \equiv (\dot{x}_3, i_3). \quad (28.18)$$

Новая диаграмма показана рис. 28.8. Чтобы получить уравнение этой кривой, можно сначала переписать (28.16) в параметрической форме

$$\begin{aligned} \dot{x}_2 &= \alpha(\alpha - 1), \\ i_2 &= \alpha, \end{aligned} \quad (28.19)$$

а затем перейти к координатам $(\dot{x}_3, i_3)$, что дает

$$\begin{aligned} \dot{x}_3 &= \alpha(\alpha - 1), \\ i_3 &= \alpha(2 - \alpha). \end{aligned} \quad (28.20)$$

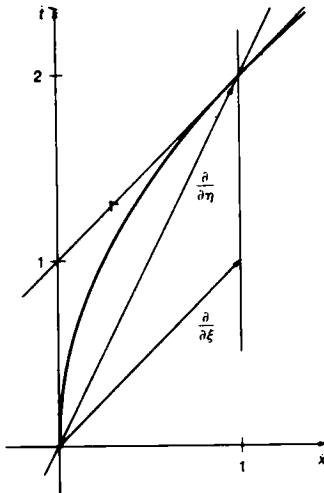


Рис. 28.10
Базисные векторы канонической системы отсчета, изображенные в системе отсчета, неподвижной относительно ряби и воды.

Прямой путь

Теперь осталось выполнить последнее преобразование, позволяющее расположить мировую линию волн ряби под углом 45° . Выполняя преобразования, мы не следили за изменениями направления этой линии. Но ее направление легко восстановить, если вспомнить, что мировая линия ряби проходит через начало координат — точку, инвариантную относительно всех наших линейных преобразований, и что она касается в этой точке волновой диаграммы, а касательная при линейных преобразованиях остается касательной. В параметрической форме записи началу координат соответствует значение параметра

$$\alpha = 0. \quad (28.21)$$

Тогда уравнение касательной

$$\begin{aligned} \dot{x}_3 &\approx -\alpha, \\ \dot{i}_3 &\approx 2\alpha \end{aligned} \quad (28.22)$$

можно получить, разложив (28.20) в ряд по малым α и сохранив только главные члены разложения. Теперь ясно, что мировая линия ряби пойдет под углом 45° , если ось $\dot{x}_3$ растянуть в два раза. Итак, мы приходим к окончательному преобразованию

$$(\dot{x}_3, \dot{i}_3) \mapsto (2\dot{x}_3, \dot{i}_3) \equiv (\dot{\xi}, \dot{\eta}). \quad (28.23)$$

Последние координаты нам еще пригодятся. Поэтому мы используем для них более удобные обозначения, чем $\dot{x}_4$ и $\dot{i}_4$. Таким образом, после выполнения всех преобразований уравнение волновой диаграммы в параметрической форме принимает вид

$$\begin{aligned} \dot{\xi} &= 2\alpha(\alpha - 1), \\ \dot{\eta} &= \alpha(2 - \alpha). \end{aligned} \quad (28.24)$$

Использованный в рассмотренном примере метод последовательных преобразований можно применять при построении канонической волновой диаграммы и в случае, когда наблюдатель движется с произвольной скоростью. Однако вместо того, чтобы приспосабливать этот метод к произвольному наблюдателю, полезнее рассмотреть другой, более короткий способ приведения волновой диаграммы к каноническому виду. Каноническая система отсчета определяется двумя базисными векторами; обозначим их $\partial/\partial \xi$ и $\partial/\partial \eta$. В нашей исходной системе отсчета, покоящейся по отношению к воде, эти векторы расположены, как на рис. 28.10. При построении этого рисунка были

выполнены все четыре требования, предъявляемые к канонической диаграмме. Вектор $\partial/\partial\eta$ по определению представляет собой вектор групповой 4-скорости волнового пакета, сопутствующего наблюдателю. Тогда 4-вектор $\partial/\partial\xi$ должен быть параллелен касательной к волновой диаграмме в точке, совпадающей с концом вектора $\partial/\partial\eta$. Длина же вектора $\partial/\partial\xi$ подбирается так, чтобы мировая линия волн ряби шла под углом 45° .

Опишем теперь все сказанное на точном геометрическом языке. С помощью уравнения (28.7) вектор групповой 4-скорости можно представить в виде следующей функции скорости v :

$$\frac{\partial}{\partial\eta} = 4v \left(\frac{\partial}{\partial t} + v \frac{\partial}{\partial x} \right). \quad (28.25)$$

Выражение для градиента фазы имеет вид

$$2d\theta = -\frac{1}{2v^2}dx + \frac{1}{v}dt. \quad (28.26)$$

Условие параллельности векторов $\partial/\partial\xi$ и $d\theta$ записывается следующим образом:

$$d\theta \cdot \frac{\partial}{\partial\xi} = 0. \quad (28.27)$$

Как уже говорилось, мировая линия волн ряби должна идти под углом 45° . Следовательно, ей должны быть параллельны как вектор $\partial/\partial t$, так и вектор $(\partial/\partial\eta - \partial/\partial\xi)$, что выражается соотношением

$$dx \cdot \left(\frac{\partial}{\partial\eta} - \frac{\partial}{\partial\xi} \right) = 0. \quad (28.28)$$

Компоненты вектора $\partial/\partial\xi$ можно найти из соотношений (28.27) и (28.28). Из (28.28) следует, что

$$\xi^x = \eta^x = 4v^2, \quad (28.29)$$

а из (28.27), что

$$-4v^2 + 2v\xi^t = 0, \quad (28.30)$$

откуда

$$\xi^t = 2v. \quad (28.31)$$

Итак, новый базис связан со старым соотношениями

$$\frac{\partial}{\partial\xi} = 2v \left(\frac{\partial}{\partial t} + 2v \frac{\partial}{\partial x} \right), \quad (28.32)$$

[Это скалярное произведение является собой пример естественной операции умножения ковектора на вектор. Для выполнения этой операции не требуется ни евклидовой метрики, ни метрики Минковского. См. разд. 15.]

[Часто соотношения, содержащие неизвестные коэффициенты пропорциональности, можно преобразовать в точные равенства, содержащие 1-формы.]

$$\frac{\partial}{\partial \eta} = 4v \left(\frac{\partial}{\partial t} + v \frac{\partial}{\partial x} \right). \quad (28.33)$$

Чтобы найти уравнение волновой диаграммы в новых координатах, достаточно выразить векторы, определяемые этой диаграммой, через их новые компоненты. Эти компоненты (обозначим их $\dot{\xi}$ и $\dot{\eta}$) определяются формулой

$$u = \dot{\xi} \frac{\partial}{\partial \xi} + \dot{\eta} \frac{\partial}{\partial \eta} \quad (28.34)$$

для любого вектора u . Принимая во внимание (28.32) и (28.33), последнюю формулу можно представить в виде

$$u = \dot{\xi} 2v \left(\frac{\partial}{\partial t} + 2v \frac{\partial}{\partial x} \right) + \dot{\eta} 4v \left(\frac{\partial}{\partial t} + v \frac{\partial}{\partial x} \right), \quad (28.35)$$

откуда после приведения подобных членов получаем

$$u = (2v\dot{\xi} + 4v\dot{\eta}) \frac{\partial}{\partial t} + (4v^2\dot{\xi} + 4v^2\dot{\eta}) \frac{\partial}{\partial x}. \quad (28.36)$$

Компоненты вектора u в координатах (x, t) определяются из соотношения

$$u = \dot{x} \frac{\partial}{\partial x} + \dot{t} \frac{\partial}{\partial t}. \quad (28.37)$$

[Учебники тензорного анализа для физиков полны сложными и громоздкими выражениями, описывающими эти преобразования координат. По-моему, от них мало пользы. Обычно проще и вернее получить их прямым путем, как это сделано здесь.]

Сравнение (28.37) с (28.36) показывает, что

$$\dot{x} = 4v^2(\dot{\xi} + \dot{\eta}), \quad (28.38)$$

$$\dot{t} = 2v(\dot{\xi} + 2\dot{\eta}).$$

На основании этих формул уравнение волновой диаграммы

$$\dot{t}^2 = \pm 4\dot{x} \quad (28.39)$$

можно записать в виде

$$(\dot{\xi} + 2\dot{\eta})^2 = \pm (\dot{\xi} + \dot{\eta}). \quad (28.40)$$

Именно это уравнение описывает кривую, изображенную на рис. 28.9.

Все наблюдатели равноправны

Волновая диаграмма на рис. 28.9 представляет собой соответствующий волнам на воде аналог спецрелятивистской гиперболы. В частности, обратите внимание на замечательный факт: канонические волновые диаграммы выглядят для различных наблюдателей одинаково! Волновая диаграмма, определяе-

мая уравнением (28.40), не зависит от v . Точно так же в СТО каждый наблюдатель получит ту же гиперболу. Случай волн на воде отличается только тем, что каждый получит одну и ту же параболу. Таким образом, если наблюдатель, движущийся над водой, следит только за волновым узором на ее поверхности и при его интерпретации пользуется только встроенными часами, то он не сможет установить свою скорость по отношению к воде. Все наблюдатели равноправны — чем не относительность! Напомню, что ход наших часов, измеряющих абсолютное ньютоново время, тоже характеризуется симметрией, а именно галилеевой симметрией. Так что такую симметрию нельзя считать ни редкостью, ни чем-то совершенно исключительным.

Мы могли бы повторить здесь все расчеты, выполненные в главе, посвященной СТО. Тогда мы получили бы преобразования между каноническими системами отсчета, аналогичные преобразованиям Лоренца, формулу доплеровского смещения частоты и формулу сложения скоростей. Выше мы показали, что мировая линия волн ряби в каждой канонической системе отсчета идет под углом 45° и что для каждого наблюдателя существует каноническая система отсчета. Все это по-видимому означает, что бессмысленно рассчитывать двигаться с такой же скоростью, как и волны ряби. Следует признать, что хотя в СТО такого сорта рассуждение выглядит вполне приемлемо, здесь вывод о невозможности догнать волны ряби кажется парадоксальным. И все-таки в некотором узком смысле он справедлив. Невозможно замедлить движение волнового пакета вплоть до полной остановки. Невозможно догнать волны ряби, ибо они никуда не движутся. Однако на самом деле наша теория — это лишь приближение к истине и не дает корректного описания медленных волн на воде. Из более точной теории следует, что при замедлении волнового пакета, образованного волнами на воде, он *превращается* в рябь.

Описанную выше модель следовало бы воспринимать лишь как шутку. Однако проблема, возникающая при рассмотрении ряби на воде, наводит на мысль, что к утверждениям о невозможности движения быстрее света следует относиться несколько обдуманнее. Чтобы приспособить СТО к описанию таких процессов, в нее нужно было бы внести совсем незначительные изменения (хотя учебники пришлось бы переделать).

ЗАДАЧИ

28.1. (08) Почему в примере на стр. 226 волновое число k отрицательно?

28.2. (16) Опишите несколько волновых пакетов в координатах, указанных на рис. 28.9.

28.3. (18) Докажите, что в канонической системе отсчета для волн на воде (рис. 28.9) гребни волны и энергия распространяются в противоположных направлениях.

28.4. (22) Парабола на рис. 28.9 не прерывается в точке пересечения с осью ξ , а простирается дальше, так что создается впечатление, будто возможно путешествие обратно во времени. Объясните, почему это не так.

28.5. (28) Опишите парадокс близнецов для часов, ассоциированных с волнами на воде.

28.6. (30) Найдите матричное представление симметрии волн на воде.

28.7. (25) Предложите способ построения кривой на рис. 28.9, как огибающей.

28.8. (29) Сделайте то же самое для кривой на рис. 28.2.

28.9. (40) Дисперсионное уравнение для волн на воде конечной глубины D имеет вид

$$\omega^2 = gk \operatorname{th}(kD).$$

Исследуйте его.

28.10. (23) Докажите, что асимптоты кривой на рис. 28.3 описываются уравнением

$$\frac{d\omega}{dk} = \frac{\omega}{k}.$$

28.11. (20) Каков смысл точки возврата кривой на рис. 28.3? Какой точке на рис. 28.2 она соответствует?

28.12. (34) Найдите векторное поле, генерирующее преобразования симметрии для волн на глубокой воде.

28.13. (36) Используя симметрию для волн на глубокой воде, докажите, что угол, который образует с направлением движения лодки остающийся за ней волновой след, не зависит от ее скорости. Для доказательства не нужно проводить вычисление этого угла.

28.14. (36) Исследуйте симметрию волн, описываемых дисперсионным уравнением

$$\omega = k - k^3$$

из задачи 27.3.

28.15. (36) Исследуйте симметрию волн, описываемых дисперсионным уравнением

$$\omega^2 = k^3$$

из задачи 27.4.

28.16. (36) Исследуйте симметрию (если она есть) волн, описываемых дисперсионным уравнением

$$\omega = k(1 + \alpha k^2)$$

из задачи 27.5.

29. Взаимодействие волновых пакетов

Дисперсионное уравнение и ассоциированное с ним линейное дифференциальное уравнение в частных производных описывают движение заданного волнового пакета. Однако обычно существует какое-то взаимодействие волновых пакетов. Кроме того, должны существовать источники волн. Чтобы учесть и то и другое, дифференциальное уравнение в частных производных дополняют другими членами. Источники описываются линейными членами, взаимодействия — нелинейными. Оказывается, можно узнать удивительно много, рассматривая только геометрические аспекты взаимодействий и не вдаваясь больше ни в какие подробности. С аналогичной ситуацией мы встречаемся в оптике, где используются как закон преломления Снеллиуса, так и уравнения Френеля, позволяющие получать более подробное описание амплитуды и поляризации световых волн. Опираясь на соответствующий геометрический подход, мы обобщим закон Снеллиуса на все волновые поля, получим обобщенный вариант спецрелятивистского закона сохранения 4-импульса и обсудим трудности, связанные с проблемой тахионов.

Простейший способ разрешить вопрос о возникновении волнового пакета — это не решать его вовсе, отодвинув момент возникновения пакета так далеко назад, чтобы этим процессом можно было пренебречь. Взамен мы воспользуемся известной информацией о волне в некоторый начальный момент времени и посмотрим, как она будет распространяться в дальнейшем.

Задача с начальными значениями

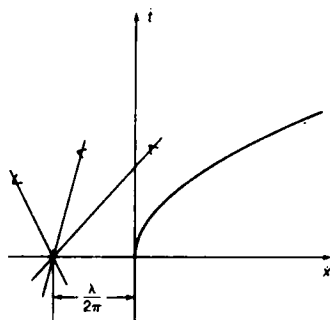


Рис. 29.1
Градиенты фазы, удовлетворяющие дисперсионному уравнению для волн на воде.

Бескоординатный подход

Такая задача называется задачей с начальными значениями. Используемый в ней геометрический подход является общим для всех проблем, рассматриваемых в данном разделе. Чтобы описать состояние волнового пакета, необходимо задать градиент фазы. В случае двух измерений 1-форма градиента фазы должна удовлетворять двум условиям:

1) ее единичный контур должен касаться волновой диаграммы (другими словами, градиент фазы должен удовлетворять дисперсионному уравнению);

2) ее линии уровня должны располагаться в соответствии с начальными условиями.

На рис. 29.1 и 29.2 изображены волновые диаграммы для волн на воде, причем на рис. 29.1 градиент фазы удовлетворяет только первому условию, а на рис. 29.2 — только второму. На рис. 29.3 показано, как следует построить градиент фазы, чтобы оба условия выполнялись одновременно. Наконец, на рис. 29.4 изображена пространственно-временная диаграмма последующей эволюции такого волнового пакета. В большинстве случаев волны распространяются как вправо, так и влево. Чтобы их различать, для полного описания волн требуется дополнительная информация, подобная заданию, скажем, значения производной $\partial\psi / \partial t$.

Эти начальные условия можно записать в бескоординатной форме, если воспользоваться операцией ограничения на поверхность, определенной в разд. 20. Начальные данные определяют структуру волны на некоторой поверхности. Поэтому мы должны выполнить операцию ограничения градиента фазы волны, согласующую его с градиентом начальной фазы, задан-

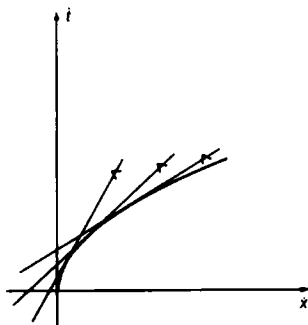


Рис. 29.2
Градиенты фазы с определенной длиной волны.

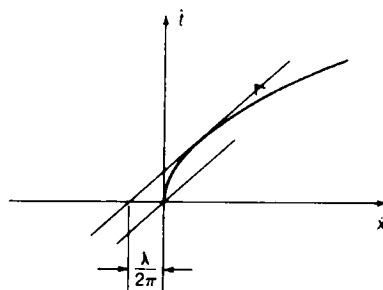


Рис. 29.3
Градиент фазы, удовлетворяющий дисперсионному уравнению и обладающий определенной длиной волны.

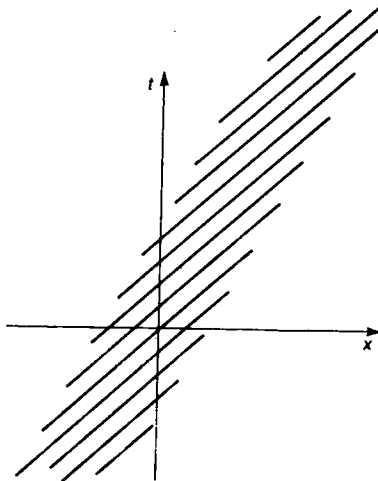


Рис. 29.4
Пространственно-временная диаграмма для конкретного волнового пакета, определенного на рис. 29.3. (Масштаб уменьшен.)

ной на этой поверхности. Градиенты должны иметь одинаковые величины и направления, но не обязательно одинаковые знаки.

Существует и другой подход к проблеме возникновения волновых пакетов: можно считать, что волна порождается некоторым устройством, которое взаимодействует с полем; будем называть такое устройство излучателем. Можно задать, например, частоту излучателя; при этом следует иметь в виду особенности каждого отдельного случая. У малого (по сравнению с длиной волны) излучателя отсутствует направленность, а его эволюция описывается единственной мировой линией. Любой излучатель взаимодействует лишь с теми волнами, которые все время находятся в фазе с его колебаниями. Другими словами, он связан только с теми волнами, градиенты фазы которых при их ограничении на мировую линию излучателя согласуются с его собственной частотой. Последнее утверждение не зависит от выбора координат и позволяет без особых дополнительных усилий решать задачи о распространении волн, создаваемых движущимся источником.

Антенны

[Напомним, что дисперсионное уравнение и соответствующая ему волновая диаграмма дают полное описание волн, причем это описание имеет силу не только в тех координатах, в которых оно дано.]

Движущийся источник

[Эта ситуация подобна той, которая довольно часто возникает в общей теории относительности, где смысл тех или иных координат косвенным образом содержится в геометрических объектах, определенных на рассматриваемом пространстве. Таким объектом может быть, например, метрика.]

[Определение операции ограничения на подпространство см. в разд. 20.]

Пример 1

Пусть источник волн на воде движется вдоль пути

$$s \mapsto (t, x, y) = (s, vs, 0) \quad (29.1)$$

и колеблется с частотой ω_0 , рассчитанной по отношению к параметру s . Из вида дисперсионного уравнения для волн на воде

$$\omega^4 = k_x^2 + k_y^2 \quad (29.2)$$

следует, что t , x и y — это уже знакомые нам координаты, связанные с системой отсчета, по отношению к которой вода покоится. Запишем формулу градиента фазы

$$d\theta = \omega dt + k_x dx + k_y dy. \quad (29.3)$$

Операция ограничения на траекторию источника сводится к выполнению отображения

$$\begin{aligned} t^* &= s, \\ x^* &= vs, \\ y^* &= 0. \end{aligned} \quad (29.4)$$

Принимая во внимание, что

$$\begin{aligned} dt^* &= ds, \\ dx^* &= v ds, \\ dy^* &= 0; \end{aligned} \quad (29.5)$$

получаем отсюда

$$d\theta^* = (\omega + vk_x) ds. \quad (29.6)$$

Так как частота колебаний источника, рассчитанная по отношению к параметру s , равна ω_0 , должно выполняться равенство

$$\omega_0 ds = (\omega + vk_x) ds. \quad (29.7)$$

Эта ситуация изображена на рис. 29.5. Итак, в нашем распоряжении имеются два уравнения (29.2) и (29.7) для трех переменных, т.е. существует однопараметрическое семейство волн, испущенных в определенных направлениях.

Размеры излучателей в одном или двух измерениях могут достигать многих длин волн. Такие устройства, способные концентрировать излучение в пучки, обычно называются антен-

Концентрация излучения в пучок

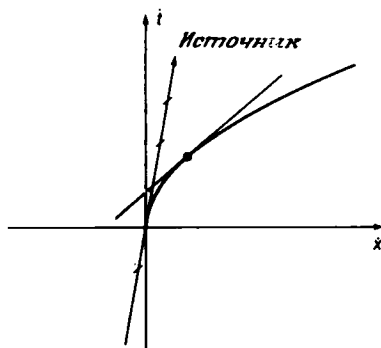


Рис. 29.5

Волновая диаграмма (в одном пространственном и одном временном измерении) для волн на воде, испущенных движущимся периодическим источником.

нами. Например, радиоантенны одномерны и имеют большую длину, что, скажем, для вертикальных антенн позволяет избежать потерь энергии на излучение волн вверх. Наоборот, антенна радара двумерна и концентрирует излучение в узкий пучок, распространяющийся в каком-то одном направлении. При расчете таких антенн речь должна идти об операции ограничения на двух- и трехмерные гиперповерхности. Полученный таким путем градиент фазы волны должен согласовываться (с точностью до знака) с фазой, заданной на гиперповерхности. Для этого необходимо, чтобы наряду с дисперсионным уравнением выполнялись еще два или три уравнения. Следует отметить, что размеры источника могут иметь порядок многих длин волн во всех трех пространственных измерениях. В этом случае вообще могут существовать только волны, соответствующие определенной фазировке излучателя. Такая ситуация возникает, например, при дифракции рентгеновских лучей, а соответствующие условия называются уравнениями Лауэ.

Получить распределенный источник с правильно согласованными фазами можно, возбудив элементарные излучатели падающим на него волновым пакетом. Простейший пример тому — отражение электромагнитной волны от зеркала. Падающая волна индуцирует в проводящей поверхности зеркала электрические токи, которые служат источниками отраженной волны. Подобное рассуждение применимо и к явлению преломления волны на поверхности раздела, где резко меняется величина показателя преломления. В обоих этих случаях можно извлечь важную информацию из требования согласованности фаз всевозможных волн на отражающей поверхности либо на поверхности раздела. Если α — градиент фазы падающей волны,

Отражение и преломление

а β — отраженной, то должно выполняться равенство

$$\alpha^* = \beta^*, \quad (29.8)$$

означающее, что ограничения величин α и β на отражающую поверхность должны быть равны.

Пример 2

Волновой пакет, образованный волнами на воде, дисперсионное уравнение для которых имеет вид

$$\omega^4 = k^2, \quad (29.9)$$

Движущийся отражатель

падает на отражатель, движущийся вдоль мировой линии $x = vt$. Найдем отраженную волну. Геометрия этой задачи в касательном пространстве представлена на рис. 29.6.

Пусть частота падающей волны ω_0 и ее волновое число k_0 удовлетворяют дисперсионному уравнению, а мировая линия отражателя определяется отображением

$$u \mapsto (t, x) = (u, vu). \quad (29.10)$$

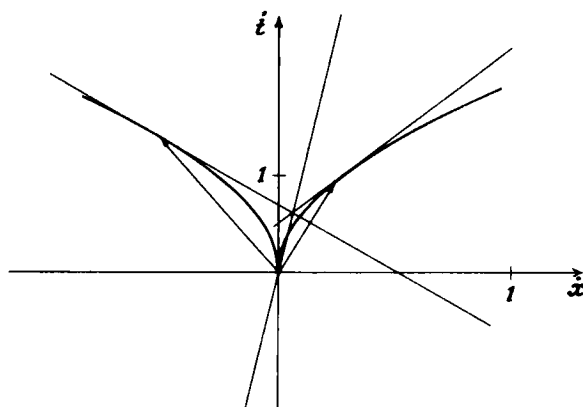


Рис. 29.6

Геометрия в касательном пространстве для волн на воде, отраженных от подвижной границы. В данном случае $v = 0,10$, $\omega = 1,40$, $k_0 = 1,97$, $\omega_1 = 2$ и $k_1 = 4$. Обратите внимание на различие масштабов по осям координат.

Ограничения базисных 1-форм на эту мировую линию имеют вид

$$dt^* = du, \quad (29.11)$$

$$dx^* = v du,$$

так что ограничение градиента фазы падающей волны можно записать в форме

$$\alpha^* = (\omega_0 + vk_0) du. \quad (29.12)$$

В результате, принимая во внимание выражение для градиента фазы отраженной волны

$$\beta = \omega_1 dt + k_1 dx, \quad (29.13)$$

получаем уравнения

$$\omega_0 + vk_0 = \omega_1 + vk_1, \quad (29.14)$$

$$k_1 = -\omega_1^2. \quad (29.15)$$

Здесь выбрано отрицательное значение волнового числа k_1 , откуда следует, что отраженная волна распространяется вправо. На рис. 29.7 показано одно из решений этой задачи. Такой

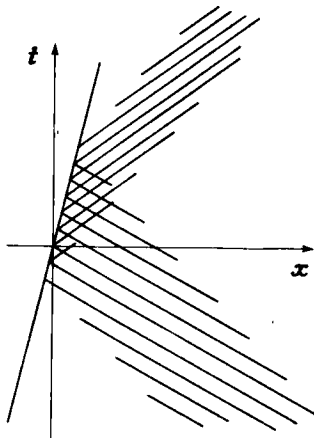


Рис. 29.7

Пространственно-временная диаграмма, на которой показаны падающая и отраженная волны, определенные на рис. 29.6. (Масштаб уменьшен в два раза.)

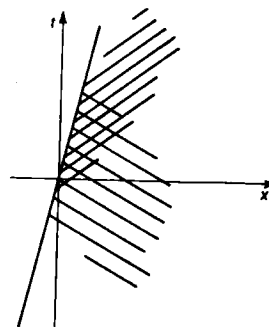


Рис. 29.8

Отражение со сдвигом фаз, равным π .

геометрический подход не позволяет установить, как связаны между собой фазы падающей и отраженной волн, если не считать, что нужно заранее согласовать их градиенты. Например, на рис. 29.8 показана ситуация, в которой фазы волн сдвинуты

Закон Снеллиуса

[Обратите внимание также на пример в разд. 33.]

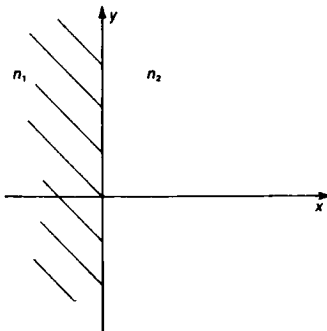


Рис. 29.9

[Наиболее полно преимущества, которые дает использование операции ограничения на подпространство, проявляются в тех случаях, когда приходится иметь дело с трехмерными или четырехмерными пространствами.]

Взаимодействия

по отношению друг к другу на фиксированную величину. Чтобы выяснить, какое решение является единственно верным, необходимо подробное исследование исходных дифференциальных уравнений в частных производных. Однако во многих случаях фаза волны не представляет интереса.

Пример 3

Световые волны в изотропной среде описываются дисперсионным уравнением

$$k_x^2 + k_y^2 = \left(\frac{n}{\lambda}\right)^2. \quad (29.16)$$

Показатель преломления n является функцией координат. Если n медленно меняется на расстояниях, сравнимых с длиной волны, то траектория волнового пакета определяется на основе принципа Гюйгенса, который нам еще предстоит обсудить в разд. 33. Если же показатель преломления быстро меняется на расстояниях порядка длины волны, то на границе раздела возникает отраженная волна, а проходящая волна резко изменяет направление распространения. Посмотрим, например, что происходит на границе раздела, совпадающей с поверхностью $x = 0$ на рис. 29.9.

Запишем градиент фазы падающей волны

$$\alpha = k_x dx + k_y dy. \quad (29.17)$$

Введем в эту формулу угол ϕ_1 , определяющий, как показано на рис. 29.10, направление распространения волны. Тогда получаем

$$\alpha = \frac{n_1 \cos \phi_1}{\lambda} dx + \frac{n_1 \sin \phi_1}{\lambda} dy. \quad (29.18)$$

Соответствующее ограничение 1-формы dx здесь равно нулю, так что падающая и прошедшая волны должны иметь одинаковые значения k_y . Таким образом

$$n_1 \sin \phi_1 = n_2 \sin \phi_2, \quad (29.19)$$

а это знакомый нам закон преломления Снеллиуса.

Существенно, что волновые пакеты могут еще и взаимодействовать между собой. На рис. 29.11 показаны два волновых пакета, проходящих через одну и ту же область пространства-времени. Предположим, что они слабо воздей-

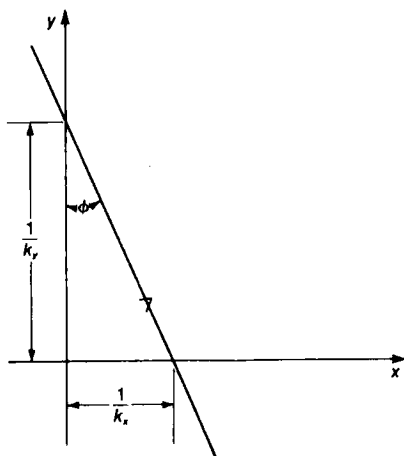


Рис. 29.10

должна быть больше там, где велики первичные волны, т.е. в отмеченных на рис. 29.12 точках пересечения гребней обеих волн. Взаимодействие первичных волн приведет к возникновению новой волны только в том случае, если такие точечные источники остаются в фазе с новой волной всюду в пространстве-времени, где перекрываются падающие волны. Волна, совпадающая по фазе с такими точечными источниками, показана на рис. 29.13. Она просто связана с падающими волнами друг на друга. Пусть это взаимодействие настолько слабо, что не разрушает волны, но в то же время достаточно сильно для того, чтобы появились некоторые дополнительные волны. Если интенсивность источника вторичных волн зависит от результата взаимного влияния двух исходных волн, то она

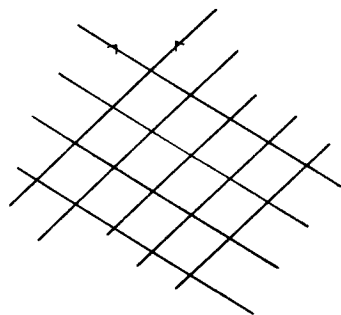


Рис. 29.11

Два перекрывающихся волновых пакета. Прямые представляют гребни волн.

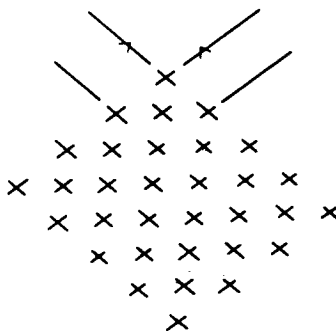


Рис. 29.12

Крестики соответствуют максимумам при сложении двух волновых полей. Показанные здесь 1-формы представляют собой градиенты фазы, деленные на 2π .



Рис. 29.13

Волновой пакет, который находится в фазе с точечными источниками, показанными на предыдущем рисунке. Градиент фазы этого пакета равен сумме 1-форм градиентов фаз распространяющихся волн.

[Обратите внимание, что сила и импульс всегда появляются в виде 1-форм.]

Пример 4

Волновые пакеты частиц в пространстве-времени описываются дисперсионным уравнением

$$\omega^2 = k^2 + \frac{m^2}{\hbar^2}, \quad (29.20)$$

где m — масса частицы. Соответствующая волновая диаграмма определяется уравнениями

$$\begin{aligned} \dot{x} &= -\frac{2\omega\hbar^2}{m^2}, \\ \dot{t} &= \frac{2k\hbar^2}{m^2}, \end{aligned} \quad (29.21)$$

откуда следует, что скорость волнового пакета равна

$$v = -\frac{\omega}{k}. \quad (29.22)$$

В рассматриваемом случае связь между градиентом фазы и спецрелятивистским 4-импульсом по определению имеет вид

$$E = \hbar\omega, \quad (29.23)$$

$$p = -\hbar k. \quad (29.24)$$

Из нашего правила суммирования градиентов фаз при наличии взаимодействия волновых пакетов следует, что

$$E_1 + E_2 = E_3, \quad (29.25)$$

$$p_1 + p_2 = p_3. \quad (29.26)$$

Эти соотношения представляют собой использованный в разд. 14 закон сохранения полного 4-импульса. Подобные законы сохранения можно написать для любых волновых полей.

Пример 5

Тахионы

Тахионы — это гипотетические частицы, описываемые лоренц-инвариантным дисперсионным уравнением, однако их групповая скорость превосходит скорость света. Теперь дисперсион-

ное уравнение должно иметь вид

$$\omega^2 = k^2 - \frac{m^2}{\hbar^2}, \quad (29.27)$$

где m — некоторая константа. Волновая диаграмма таких частиц описывается уравнением

$$(\dot{x})^2 = (\dot{t})^2 + \left(\frac{\hbar}{m}\right)^2; \quad (29.28)$$

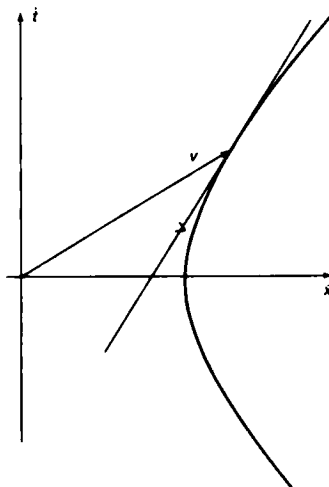


Рис. 29.14

Волновая диаграмма тахионного волнового пакета в касательном пространстве.

она отображена на рис. 29.14. Соответствующий волновой пакет представлен на рис. 29.15.

Все это напоминает вполне согласованную теорию — обычную теорию частиц, в которой пространство и время поменялись местами. Раз пространство и время входят в теорию симметрично, то может ли здесь что-нибудь быть не так? Существует ли хоть какое-нибудь различие между пространством и временем? Существует, но оно коренится не в уравнениях, а в граничных условиях, ибо вопросы, на которые отвечает наука, формулируются следующим образом: что произойдет, если заданы такие-то и такие-то условия? Такого сорта вопросы выделяют временное направление в качестве направления предсказания. Так что в пространстве-времени есть направление, определяемое вопросами, которые мы задаем, и теми состояниями, к которым мы приходим.

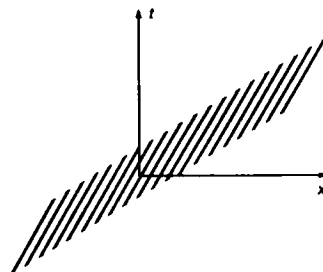


Рис. 29.15

Пространственно-временная диаграмма для волнового пакета, изображенного на рис. 29.14. (Масштаб уменьшен.)

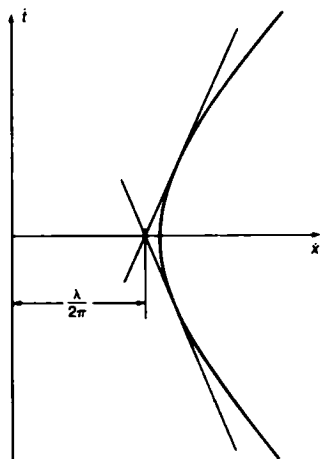


Рис. 29.16
Две тахионные моды, связанные с пространственно-периодическим источником с длиной волны λ .

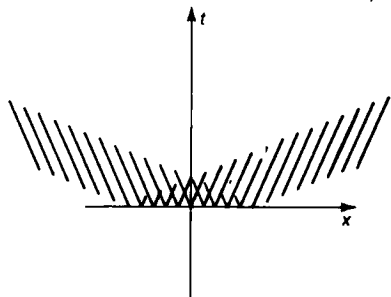


Рис. 29.17
Пространственно-временная диаграмма для изображенных на рис. 29.16 двух тахионных волновых пакетов, расходящихся от источника конечной протяженности. Один пакет распространяется влево, а другой вправо. (Масштаб уменьшен.)

Тахионы представляют интерес только в том случае, если они могут взаимодействовать с известными нам физическими полями. Допустим, что в нашем распоряжении оказалось поле, способное взаимодействовать с тахионным полем. Пусть источник в момент времени $t = 0$ обладает, для простоты, периодической пространственной конфигурацией. Такой источник, как показано на рис. 29.16 и 29.17, будет связан с отдельной парой тахионов. Пока что все в порядке. Но что произойдет, если длина волны источника возрастет? Это несложно установить. Обратите внимание, что в ситуации, представленной на рис. 29.18, две тахионные моды исчезли. Дело в том, что дисперсионное уравнение имеет вид квадратного уравнения, а частота определяется дисперсионным уравнением (29.27). При малых k это уравнение не имеет *действительных* решений. Частота становится чисто мнимой. В результате осциллирующий режим сменяется экспоненциальным. Возникает пара экспонент, одна из которых нарастает. Таким образом, тахионная волна неустойчива по отношению к длинноволновым возмущениям. Никому еще не удалось найти средство, позволяющее избавиться от этой неустойчивости. Это смертельный порок теории тахионов, который известен уже довольно давно. Как мне кажется, люди, которые занимаются тахионами, все еще надеются, что рано или поздно будет найден какой-нибудь красивый способ разрешить эту проблему.

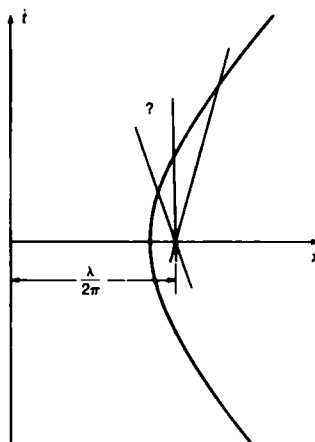


Рис. 29.18
По мере увеличения пространственной длины волны тахионные моды исчезают.

ЗАДАЧИ

29.1. (14) По отношению к каким возмущениям, коротковолновым или длинноволновым, неустойчивы волны, которые описываются волновой диаграммой, изображенной на рис. 29.19? Какова критическая длина волны, разделяющая устойчивые и неустойчивые моды?

29.2. (12) Изобразите распространяющийся вправо волновой пакет, для которого длина волны и волновая диаграмма показаны на рис. 29.20.

29.3. (20) На рис. 29.21 отмечены положения *гребней* распространяющейся вправо волны в момент времени $t = 0$. Начертите мировые линии этих гребней, воспользовавшись изображенной на том же рисунке волновой диаграммой.

29.4. (26) Докажите, что из вида уравнения

$$\frac{\partial^2 \psi}{\partial t^2} - \frac{\partial^2 \psi}{\partial x^2} - \mu^2 \psi = 0$$

сразу следует неустойчивость его длинноволновых решений.

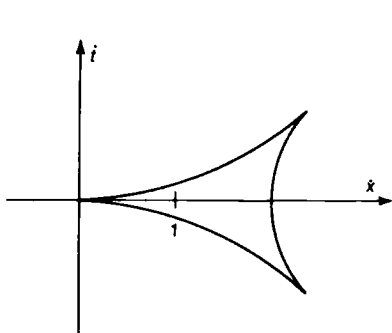


Рис. 29.19

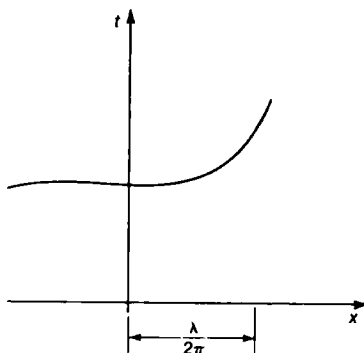


Рис. 29.20

29.5. (29) Рассмотрите граничную задачу для обычного волнового уравнения

$$\frac{\partial^2 \psi}{\partial t^2} - \frac{\partial^2 \psi}{\partial x^2} + \mu^2 \psi = 0,$$

когда функция ψ задана на мировой линии $x = 0$. Это так называемая задача сигнализации. Она выглядит подобно граничной задаче для тахиона, если соответствующую диаграмму по-

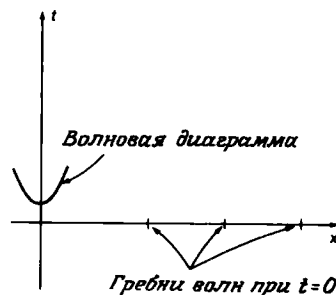


Рис. 29.21

вернуть на 90° . Почему в данном случае не возникает неустойчивость?

29.6. (31) (Для изучавших электричество и магнетизм.) Опишите с помощью волновых диаграмм возникновение критической частоты волновода.

29.7. (26) В примере 2 настоящего раздела при некоторых скоростях может не быть отраженной волны. Покажите, что условием этого является неравенство

$$(u + v) < \sqrt{2}u.$$

29.8. (34) Пусть в примере 2 настоящего раздела движущийся отражатель не только отражает волны, но и частично пропускает их. Докажите, что при некоторых условиях возможны две дополнительные преломленные волны.

29.9. (30) В предыдущей задаче на определенной стадии теряются две волновые моды. Однако в отличие от примера с тахионами эта потеря вовсе не означает, что в задаче присутствует неустойчивость. Почему?

30. Тяготение

*Нужно было быть Ньютоном,
чтобы заметить, что падает Луна,
хоть всякий видит —
не падает она*

Поль Валери

Тяготением мы до сих пор пренебрегали. Это ведет к нарушению стройности наших логических построений — развитая в предыдущих главах специальная теория относительности может оказаться неприменимой на поверхности Земли. Конечно, когда речь идет о живущем всего лишь несколько микросекунд μ -мезоне, тяготение безусловно можно игнорировать; однако оно оказывает возмущающее воздействие, вследствие чего все основные посыпки теории удовлетворяются лишь приближенно. Из-за влияния тяготения нам приходится иметь дело с несвободными частицами. В отличие от любой другой силы, тяготение невозможно исключить или скомпенсировать с помощью экранирования. Ни один известный нам объект не может не испытывать действия гравитационного притяжения.

Выйти из этой сложной ситуации удастся неожиданным и в то же время красивым способом. Откажемся от соблазна просто так, механически ввести тяготение в нашу теорию. Вместо этого вернемся назад и «встроим» его непосредственно в основание теории. Ключ к этой операции содержится в задаче 1.8, где вам было предложено доказать, что отображение

$$(x, t) \mapsto (X, T) = (x - \frac{1}{2}t^2, t) \quad (30.1)$$

(или любой его линейный вариант) переводит семейство парабол

$$x = a + bt + \frac{1}{2}t^2 \quad (30.2)$$

в прямые линии

$$X = a + bT. \quad (30.3)$$

Поведение объектов, падающих в однородном гравитационном поле, описывается как раз таким семейством парабол; следовательно, можно так изменить наше определение свободной частицы, чтобы оно относилось к частицам в однородном поле тяготения и при этом не нарушались постулаты, касающиеся свободных частиц. Такой подход оказывается полезным и плодотворным, так что впредь мы будем пользоваться именно им.

Свободная частица — это частица, на которую не действуют никакие силы, кроме, возможно, силы тяготения.

Перефразируя эпиграф, можно сказать, что нужно было обладать гениальностью Эйнштейна, чтобы заметить, что Луна движется по «прямой», тогда как всякий видит совсем иное.

Опираясь на такое представление о свободной частице, специальную теорию относительности можно применять даже в однородных гравитационных полях. Но в пределах малых областей все гладкие поля однородны, поэтому СТО теперь оказывается пригодной для изучения локального поведения частиц, часов и световых сигналов вблизи каждой точки. Это огромный прогресс по сравнению с возможностью применения СТО лишь в тех случаях, когда тяготение отсутствует. Именно это имеют в виду, когда говорят, что СТО локально аппроксимирует геометрию пространства-времени, точно так же, как плоскость аппроксимирует локальную геометрию любой гладкой искривленной поверхности.

Решив ввести инерциальные системы отсчета с помощью приведенного выше определения свободной частицы, мы дол-

Новая точка зрения на свободную частицу

[Долсон и Постон [9] воспользовались моделью искривленного пространства для описания распространения света и возникновения миражей. На стр. 423 их книги имеется отличный рисунок.]

Локальная теория — это по-прежнему СТО

жны изучить локальное поведение свободных частиц в различных гравитационных полях. Ведь может оказаться, что дисперсионное уравнение для частиц уже не изображается спецрелятивистской гиперболой. Исследовать этот вопрос экспериментально очень сложно, ибо по существу мы «привязаны» к одному значению гравитационного поля. Тем не менее благодаря суточному вращению Земли, а также благодаря эксцентриситету ее орбиты имеют место суточные и годовые вариации поля Солнца. Предпринимались тщательные поиски эффектов, обладающих такими периодами, однако пока не получено никаких данных, указывающих на необходимость какой-либо модификации дисперсионного уравнения СТО. Тем не менее специалисты, занимающиеся планированием космических экспериментов, рассчитывают послать часы поближе к Солнцу, чтобы иметь возможность попристальнее изучить такие эффекты.

Несколько метафизическое рассуждение приводит нас к выводу, что присутствие полей тяготения не должно сказываться на дисперсионном уравнении СТО. Пусть в некоторой теории имеется безразмерный параметр θ , причем в результате проведенных в некоторых условиях измерений установлено, что он равен $1,473 \pm 0,001$. Наверняка каждый согласится, что при других, но близких к первоначальным условиям этот параметр может иметь какое-то чуть-чуть иное значение. Теперь предположим, что измеренное значение θ равно $2,001 \pm 0,001$. Тогда было бы вполне разумно предположить, что значение $\theta = 2$ фактически является точным и при изменении условий не должно изменяться, пусть даже слабо. В ряду дисперсионных уравнений дисперсионное уравнение СТО стоит особняком, как стоят особняком целые числа в множестве всех действительных чисел. Дисперсионному уравнению СТО соответствует кривая второго порядка; следовательно, его можно выразить через тензор второго ранга. Это колоссальное упрощение! Ведь то, что для тензора верно здесь и сейчас, в силу его свойств верно всюду и всегда.

Общая теория относительности

Общая теория относительности — это теория тяготения, в которой движение волн и частиц в окрестности любой точки определяется тензорным полем

$$\mathcal{G} = g_{\alpha\beta} dx^\alpha \otimes dx^\beta, \quad (30.4)$$

а свободные частицы массой m описываются волновой диаграммой

$$g_{\alpha\beta} \dot{x}^\alpha \dot{x}^\beta = -m^2 \quad (30.5)$$

или эквивалентным ей дисперсионным уравнением. Наличие тяготения приводит к изменению $\mathcal{G}$ от точки к точке, а это изменение обусловлено распределением материи в пространстве-времени. Влияние материи на тяготение описывается уравнениями, которые называются уравнениями Эйнштейна. Собственно, вот и все, что содержится в ОТО. Нам, разумеется, пришлось основательно потрудиться, чтобы прийти к точке зрения, позволяющей дать столь простое описание ОТО. Математическое описание влияния распределенной в пространстве-времени материи на тензорное поле $\mathcal{G}$ довольно формально, но требует навыков работы с дополнительным набором математических «инструментов». Развитие таких навыков может занять у нас недели, так что оставим эту задачу специальным курсам общей теории относительности. Здесь же мы будем рассматривать то или иное пространство-время как заданное и займемся только изучением его свойств. Это даст возможность во всех подробностях исследовать и обсудить стандартные космологические модели.

Пример

Внешнее асимптотически плоское пространство-время невращающейся черной дыры описывается метрическим тензором

$$\mathcal{G} = -\left(1 - \frac{2m}{r}\right) dt^2 + \frac{dr^2}{\left(1 - \frac{2m}{r}\right)} + r^2(d\theta^2 + \sin^2 \theta d\phi^2). \quad (30.6)$$

Наиболее сложное свойство таких 4-мерных общерелятивистских пространств заключается в том, что смысл используемых для их описания координат заранее неизвестен. Понять, что оз-

ТАБЛИЦА 30.1. ГЕОМЕТРИЧЕСКИЕ ЕДИНИЦЫ

<i>Длины</i>	
Фут	$1,02 \cdot 10^{-9} \text{ с}$
Метр	$3,34 \cdot 10^{-9} \text{ с}$
Радиус Земли	$2,12 \cdot 10^{-2} \text{ с}$
Радиус Солнца	$2,32 \text{ с}$
Астрономическая единица длины	499 с
Световой год	$3,16 \cdot 10^7 \text{ с}$
Парсек	$1,03 \cdot 10^8 \text{ с}$
<i>Массы:</i>	
Килограмм	$2,48 \cdot 10^{-36} \text{ с}$
Масса Земли	$1,48 \cdot 10^{-11} \text{ с}$
Масса Солнца	$4,92 \cdot 10^{-6} \text{ с}$
<i>Другие величины:</i>	
Плотность воды	$6,67 \cdot 10^{-8} \text{ с}^{-2}$
$\hbar$	$1,82 \cdot 10^{-86} \text{ с}^2$
$(50 \text{ км/с} \cdot \text{Мпс})^{-1}$	$6,18 \cdot 10^{17} \text{ с}$

дах, а угловая скорость — в (рад/с). В геометрических единицах масса Солнца равна $4,92 \cdot 10^{-6}$ секунд, масса Земли $1,48 \cdot 10^{-11}$ секунд, а масса, соответствующая одному килограмму, составляет $2,48 \cdot 10^{-36}$ секунд (табл. 30.1).

Световые сигналы

[Обсуждение процедуры обращения см. на стр. 189.]

Что можно сказать по поводу световых сигналов? Обсуждая поведение световых сигналов с точки зрения СТО, мы не делали различия между групповой и фазовой скоростями. Как сейчас будет показано, аналогичная ситуация имеет место и в ОТО. Дисперсионное уравнение для волнового пакета светового сигнала представляет собой тензорное уравнение

$$g^{\alpha\beta} p_\alpha p_\beta = 0, \quad (30.8)$$

где $g^{\alpha\beta}$ — компоненты тензора, обратного метрическому, а p_α — компоненты градиента фазы

$$p = d\theta. \quad (30.9)$$

Уравнение (30.8) можно рассматривать либо как постулат, либо как результат экспериментальных наблюдений, либо как следствие уравнений Максвелла. Для нас оно будет играть роль удобной отправной точки.

Пример

Для спецрелятивистской метрики

$$\mathcal{G} = -dt^2 + dx^2 \quad (30.10)$$

обратный тензор $\mathcal{G}^{-1}$ имеет вид

$$\mathcal{G}^{-1} = -\frac{\partial}{\partial t} \otimes \frac{\partial}{\partial t} + \frac{\partial}{\partial x} \otimes \frac{\partial}{\partial x}. \quad (30.11)$$

[Этот тип обращения вы найдете на стр. 181.]

Если написать

$$p = k dx + \omega dt, \quad (30.12)$$

то наше дисперсионное уравнение

$$\mathcal{G}^{-1} \cdot (p, p) = 0 \quad (30.13)$$

можно переписать в индексных обозначениях. Из формул

$$\mathcal{G}^{-1} = g^{\alpha\beta} \frac{\partial}{\partial x^\alpha} \otimes \frac{\partial}{\partial x^\beta} \quad (30.14)$$

и

$$p = p_\alpha dx^\alpha \quad (30.15)$$

следует, что

$$\begin{aligned} g^{\alpha\beta} p_\gamma p_\delta \frac{\partial}{\partial x^\alpha} \otimes \frac{\partial}{\partial x^\beta} \cdot (dx^\gamma, dx^\delta) = \\ = g^{\alpha\beta} p_\alpha p_\beta = 0, \end{aligned} \quad (30.16)$$

т.е. мы снова имеем уравнение (30.8). Записав его через k и ω , получим хорошо знакомое нам дисперсионное уравнение

$$\mathcal{G}^{-1} \cdot (p, p) = k^2 - \omega^2 = 0. \quad (30.17)$$

Это дисперсионное уравнение обладает особенностью, которая заставляет нас вернуться назад и повторить вывод выражения для групповой скорости. Чтобы найти групповую скорость σ , продифференцируем дисперсионное уравнение

Групповая скорость света

[Вычисление этой производной — хорошее упражнение в жонглировании индексами.]

$$\frac{\partial}{\partial x^\gamma} (g^{\alpha\beta} p_\alpha p_\beta) = \frac{\partial g^{\alpha\beta}}{\partial x^\gamma} p_\alpha p_\beta + 2g^{\alpha\beta} p_\alpha \frac{\partial p_\beta}{\partial x^\gamma}. \quad (30.18)$$

Так как p — это градиент, имеем

$$\frac{\partial p_\beta}{\partial x^\gamma} = \frac{\partial p_\gamma}{\partial x^\beta}, \quad (30.19)$$

ибо частные производные коммутируют. Таким образом, приходим к уравнению

$$\frac{\partial p_\gamma}{\partial x^\beta} (g^{\alpha\beta} p_\alpha) = -\frac{1}{2} \frac{\partial g^{\alpha\beta}}{\partial x^\gamma} p_\alpha p_\beta. \quad (30.20)$$

Левая часть здесь представляет собой производную величин p_γ вдоль направления, которое мы считаем направлением групповой скорости

$$\sigma = g^{\alpha\beta} p_\alpha \frac{\partial}{\partial x^\beta}. \quad (30.21)$$

Особое свойство рассматриваемых волн заключается в том, что

$$p \cdot (\sigma) = 0. \quad (30.22)$$

Нормировка невозможна

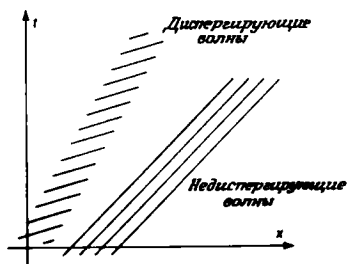


Рис. 30.1
Различие между диспергирующими и недиспергирующими волнами.

Последнее соотношение не позволяет построить для световых волн обычную волновую диаграмму, ибо она превращается в луч, выходящий прямо из начала координат. Любая попытка нормировки приводила бы к его вырождению в точку. Этот луч и есть та самая штриховая линия, которой мы договорились изображать мировую линию светового сигнала. Энергия световых волн распространяется с той же скоростью, что и гребень волны. На рис. 30.1 наряду со световыми волновыми пакетами для сравнения изображены обычные волновые пакеты, соответствующие диспергирующим волнам. Так как для света (в пустоте) дисперсия отсутствует, в качестве светового сигнала можно рассматривать либо весь волновой пакет целиком, либо отдельный гребень волны. И тот и другой распространяются с одинаковой скоростью. Именно благодаря этому вполне законна процедура, использованная в разд. 13 для вычисления доплеровского смещения. Там мы вычислили сдвиг частоты, исходя из величины интервала между мировыми линиями двух световых сигналов, что допустимо только для волн, не испытывающих дисперсии. Аналогичную процедуру мы применим в разд. 32 при вычислении сдвига частоты в гравитационном поле Земли и в разд. 40 при расчете сдвига частоты в расширяющейся Вселенной.

ЗАДАЧИ

- 30.1. (10) Убедитесь в справедливости записанного в геометрических единицах закона Кеплера для системы Земля — Солнце.
- 30.2. (12) Покажите, что величина массы m_g , выраженная в геометрических единицах, связана с массой m_k , выраженной в килограммах, формулой

$$m_g = \frac{G m_k}{c^3}.$$

- 30.3. (23) Евклидова метрика $\mathcal{E}$, записанная в полярных координ-

натах, имеет вид

$$\mathcal{G} = dr^2 + r^2 d\theta^2.$$

Это простой пример метрического тензора с зависящими от координат компонентами. Выведите эту формулу, опишите метрику в общих чертах и обсудите ее.

30.4. (28) Покажите, что двумерное пространство (u, v) с метрикой

$$\mathcal{G} = du^2 - u^2 dv^2$$

представляет собой не что иное, как пространство с метрикой Минковского, записанной в непривычных координатах. Воспользуйтесь преобразованием

[См. задачу 1.5.]

$$x = u \operatorname{ch} v,$$

$$t = u \operatorname{sh} v.$$

31. Пространство-время в окрестности Земли

В нескольких следующих разделах мы будем изучать внешнее гравитационное поле сферически-симметричного тела, подобного Земле. В результате мы увидим, как искривленное пространство-время в действительности описывает тяготение, узнаем, каков физический смысл кривизны пространства-времени и познакомимся с экспериментом, позволяющим проверить справедливость ОТО.

Метрика $\mathcal{G}$ вне Земли или любого другого сферически-симметричного тела приближенно определяется выражением

Метрика пространства-времени

$$\mathcal{G} = -dt^2 + dr^2 + r^2(d\theta^2 + \sin^2 \theta d\phi^2) + \frac{2m}{r}(dt^2 + dr^2). \quad (31.1)$$

На поверхности Земли отличие компонент метрики (31.1) от спецрелятивистских значений составляет около $1 \cdot 10^{-9}$. Мы пренебрегли здесь многими малыми поправками к этому, уже и так малому члену. К их числу относятся поправки на сплюснутость Земли (погрешность порядка 10^{-6}), массу атмосферы (около $1 \cdot 10^{-15}$), влияние Солнца и Луны (порядка $1 \cdot 10^{-17}$). Кроме того, не учтены квадратичные члены типа $(m/r)^2$ (погрешность порядка $1 \cdot 10^{-18}$). Нашей подготовки пока недостаточно для вывода метрики (31.1) из уравнений ОТО. Из него следует, что пространство-время вне любого сферически-симметричного тела заданной массы единственно и не зависит от его внутренней структуры, будь это планета или черная ды-

[Здесь намеренно используется приближенное выражение, хотя точное едва ли сложнее. Таким путем я хочу подчеркнуть, что есть много астрофизических поправок, более важных, чем отброшенные члены. Точность решения уравнений общей теории относительности — это бессмысленная точность. Редкое упоминание астрофизических поправок отчасти связано с тем, что общей теорией относительности в основном занимались математики.]

Интерпретация координат

ра. Выражение (31.1) — это упрощенный вариант приведенной в разд. 30 метрики, описывающей пространство-время вне черной дыры, в котором опущены члены $(m/r)^2$.

Все, что необходимо для физического истолкования пространства-времени ОТО, содержится в нем самом. Обнаружив незнакомую вам фотографию, вы можете тщательно изучить ее и на основании признаков самой фотографии узнать, где, когда и как она сделана. А что можно сказать о пространстве-времени с метрикой (31.1), если его исследовать в том же духе?

Положив в (31.1) $m = 0$, мы приходим к рассмотренной в разд. 23 метрике пространства-времени Минковского, записанной в полярных координатах. Таким образом, в отсутствие тела мы имеем пространство-время Минковского.

Доказательство

Покажем, что метрика

$$\mathcal{N} = -dt^2 + dx^2 + dy^2 + dz^2, \quad (31.2)$$

переписанная в полярных координатах

$$\begin{aligned} z &= r \cos \theta, \\ x &= r \sin \theta \cos \phi, \\ y &= r \sin \theta \sin \phi, \end{aligned} \quad (31.3)$$

принимает форму метрики (31.1) с $m = 0$. Каждая из старых координат (x, y, z) является функцией полярных координат (r, θ, ϕ) . Градиенты этих функций представляют собой 1-формы вида

$$\begin{aligned} dz^* &= \cos \theta dr - r \sin \theta d\theta, \\ dx^* &= \sin \theta \cos \phi dr + r \cos \theta \cos \phi d\theta - \\ &\quad - r \sin \theta \sin \phi d\phi, \\ dy^* &= \sin \theta \sin \phi dr + r \cos \theta \sin \phi d\theta + \\ &\quad + r \sin \theta \cos \phi d\phi. \end{aligned} \quad (31.4)$$

Метрика $\mathcal{N}$ записана через 1-формы dx , dy и dz . Чтобы преобразовать ее к новым координатам, мы воспользуемся формулами (31.4), выражающими старые 1-формы через новые. Тогда

после подстановки получим

$$\begin{aligned} \mathcal{N} = & -dt^2 + (\cos \theta dr - r \sin \theta d\theta)^2 + \\ & + (\sin \theta \cos \phi dr + r \cos \theta \cos \phi d\theta - \\ & - r \sin \theta \sin \phi d\phi)^2 + \\ & + (\sin \theta \sin \phi dr + r \cos \theta \sin \phi d\theta + \\ & + r \sin \theta \cos \phi d\phi)^2, \end{aligned} \quad (31.5)$$

откуда после приведения подобных членов следует искомый результат

$$\mathcal{N} = -dt^2 + dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2. \quad (31.6)$$

Важнейшее свойство пространства-времени — его симметрии. В рассматриваемом случае временная переменная t не входит в компоненты метрики $\mathcal{S}$, так что преобразование

$$(t, r, \theta, \phi) \mapsto (t + a, r, \theta, \phi) \quad (31.7)$$

при любом a оставляет метрику $\mathcal{S}$ неизменной. Векторное поле, генерирующее такое преобразование — это просто $\partial/\partial t$. Любое пространство-время с такого рода симметрией по отношению к сдвигу во времени называется стационарным. Метрика $\mathcal{S}$ не меняется также и при отражениях $t \rightarrow -t$, $dt \rightarrow -dt$. Стационарное пространство-время, обладающее такой дополнительной симметрией, называется статическим. Наконец, угловые координаты θ и ϕ входят в рассматриваемой метрике в члены, совокупность которых описывает геометрию на поверхности сферы, так что наше пространство-время еще и сферически-симметрично.

Рассмотрим одну из таких сферически-симметричных поверхностей, задаваемую уравнениями $r = r_0$ и $t = t_0$. Линейный размер такой сферической поверхности, измеренный с помощью инструментов, расположенных на ней самой, определяется величиной r_0 , поскольку $(r_0)^2$ представляет собой коэффициент при части метрики $(d\theta^2 + \sin^2 \theta d\phi^2)$, описывающей поверхность сферы.

Пример

Рассчитаем длину наибольшего замкнутого контура, натянутого на такую сферу, т.е. длину экватора или окружности, ограничивающей любой другой большой круг. Экватор — это кри-

Симметрии

Симметрия по отношению к сдвигу во времени

Сферическая симметрия

Сферическая оболочка

вая, задаваемая отображением

$$u \mapsto [T(u), R(u), \Theta(u), \Phi(u)] = (t_0, r_0, \frac{\pi}{2}, u), \quad (31.8)$$

где $0 \leq u \leq 2\pi$. Короткий отрезок кривой, заключенный между значениями параметра u и $u + \Delta u$, описывается вектором $\Delta u (\partial/\partial\phi)$, а длина этого вектора определяется с помощью метрики $\mathcal{G}$ и равна

$$(\Delta l)^2 = \mathcal{G} \cdot \left(\Delta u \frac{\partial}{\partial\phi}, \Delta u \frac{\partial}{\partial\phi} \right) = (\Delta u)^2 r_0^2. \quad (31.9)$$

Суммирование длин всех таких отрезков, лежащих на кривой, дает

$$l = \int_0^{2\pi} \Delta l = \int_0^{2\pi} r_0 du = 2\pi r_0. \quad (31.10)$$

Именно потому, что длина окружности большого круга оказалась равной $2\pi r_0$, мы и говорим, что измеренный внутренним образом линейный размер сферы определяется величиной r_0 . Из тех же соображений следует, что площадь сферы равна $4\pi r_0^2$. Обратите внимание, что диаметр сферы мы измерить не можем. При попытке сделать это мы будем вынуждены проникнуть внутрь Земли, где метрика (31.1) уже не верна. Вот почему для расчета линейного размера сферы мы пользовались внутренними измерениями. Аналогичная ситуация имеет место и для четырехмерного пространства-времени. При изучении четырехмерных пространств мы тоже будем иметь дело только с внутренней геометрией. Мы не станем обсуждать вопрос о том, погружено ли исследуемое пространство-время в некоторое пространство большего числа измерений.

[Здесь все довольно просто. Вникнув в суть вопроса, вы поймете, что ответ непосредственно следует из вида метрики $\mathcal{G}$.]

[Слова «внутренним образом» означают «оставаясь на поверхности сферы».]

Кривизна пространства-времени

Помимо измерений на поверхности сферы можно измерить и расстояние между двумя сферическими слоями $r = \text{const}$. Два слоя, имеющие радиусы r_0 и $r_0 + \Delta r$, соединяются вектором $\Delta r (\partial/\partial r)$, длина которого в пределе при $\Delta r \rightarrow 0$ дает расстояние Δl между этими слоями. Воспользовавшись метрикой (31.1), мы получим

$$(\Delta l)^2 = \mathcal{G} \cdot \left(\Delta r \frac{\partial}{\partial r}, \Delta r \frac{\partial}{\partial r} \right) = \left(1 + \frac{2m}{r_0} \right) (\Delta r)^2, \quad (31.11)$$

откуда с оговоренной нами степенью точности следует, что

$$\Delta l = \left(1 + \frac{m}{r_0} \right) \Delta r. \quad (31.12)$$

Такой результат не может получиться в евклидовой геометрии. Так, две концентрические сферы с радиусами r_0 и $r_0 + \Delta r$ были бы расположены друг от друга на расстоянии Δr . Обнаруженное отклонение от евклидовой геометрии — это проявление кривизны пространства-времени. Следовательно, наши вычисления позволяют утверждать, что пространство-время в окрестности Земли искривлено.

Довольно большое число промежуточных шагов в предыдущем рассуждении могло бы создать впечатление, что кривизна пространства-времени — понятие в некотором смысле таинственное. Однако это не так. Повторим наши рассуждения, но уже в другой, более конкретной, форме. Представим себе такую ситуацию: некоторый наблюдатель, находящийся на мировой линии $r = \text{const}$, $\theta = \text{const}$ и $\phi = \text{const}$, посылает в момент события E во всех направлениях световые сигналы. Где окажутся эти сигналы спустя короткий промежуток собственного времени $\Delta\tau$? Сначала рассмотрим только радиальное направление. В этом случае можно начертить двумерную диаграмму пространства-времени, которая и приведена на рис. 31.1. Для событий, близких к событию E , координаты r и t образуют инерциальную, но не каноническую, систему отсчета. Отклонение от канонической системы отсчета определяется членами, линейными по (m/r) . События, которые наш наблюдатель считает происшедшими одновременно, через промежуток времени $\Delta\tau$ после E образуют единичный контур 1-формы

$$\omega = \left(1 - \frac{m}{r_0}\right) \frac{dt}{\Delta\tau}. \quad (31.13)$$

Для упрощения квадратного корня мы здесь снова воспользовались теоремой о разложении в биномиальный ряд. Вектор σ , описывающий мировую линию светового сигнала с момента излучения до события, происшедшего спустя время $\Delta\tau$, должен удовлетворять двум условиям:

$$\omega \cdot (\sigma) = 1, \quad (31.14)$$

$$\mathcal{G} \cdot (\sigma, \sigma) = 0. \quad (31.15)$$

Пусть

$$\sigma = \Delta t \frac{\partial}{\partial t} + \Delta r \frac{\partial}{\partial r}, \quad (31.16)$$

тогда, согласно этим условиям, получаем

$$\left(1 - \frac{m}{r_0}\right) \Delta t = \Delta\tau, \quad (31.17)$$

Другая точка зрения на кривизну

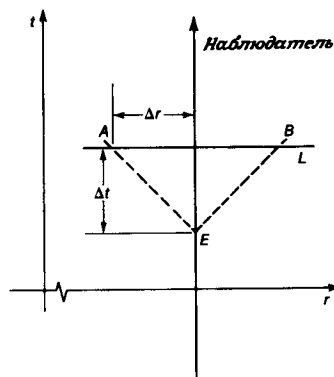


Рис. 31.1

Прямая L одновременных событий, которые неподвижный наблюдатель видит через промежуток собственного времени $\Delta\tau$ после события E . Мировые линии двух радиальных световых сигналов пересекают эту прямую в точках A и B . Обратите внимание, что начало отсчета радиальной координаты находится вне нашего поля зрения.

$$-\left(1 - \frac{2m}{r_0}\right)(\Delta t)^2 + \left(1 + \frac{2m}{r_0}\right)(\Delta r)^2 = 0, \quad (31.18)$$

[Это еще один пример, показывающий, почему 1-форму dr следует отличать от малого приращения координаты Δr .]

откуда следует, что

$$\Delta r = \pm \left(1 - \frac{m}{r}\right) \Delta \tau. \quad (31.19)$$

Не составляет особого труда повторить это рассуждение для случая, когда имеется еще одно пространственное измерение. Будем рассматривать экваториальную плоскость $\theta = \pi/2$, а ϕ используем в качестве добавочной пространственной координаты. События, соответствующие приходу световых сигналов через время $\Delta \tau$ после их испускания, теперь определяются вектором

Световые сигналы

$$\sigma = \Delta t \frac{\partial}{\partial t} + \Delta r \frac{\partial}{\partial r} + \Delta \phi \frac{\partial}{\partial \phi} \quad (31.20)$$

с компонентами, удовлетворяющими условиям

$$\left(1 - \frac{2m}{r_0}\right)(\Delta t)^2 = (\Delta \tau)^2, \quad (31.21)$$

$$\left(1 + \frac{2m}{r_0}\right)(\Delta r)^2 + r_0^2 (\Delta \phi)^2 = (\Delta \tau)^2. \quad (31.22)$$

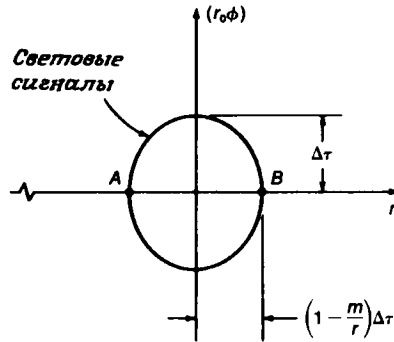


Рис. 31.2

Пространственное сечение, содержащее одновременные события, которые произошли через время $\Delta \tau$ после события E . Изображенная на рис. 31.1 прямая L на этом сечении совпадает с осью r , на которой указаны события A и B . Азимутальный масштаб выбран таким образом, чтобы при $m \rightarrow 0$ геометрическое место точек, образованных световыми сигналами, испущенными в момент события E , представляло собой окружность.

Для наглядности представим полученный результат в графической форме. На рис. 31.2 приведена пространственная диаграмма в двух измерениях, на которой нанесены только события, происшедшие спустя время Δt после события E . Все события, соответствующие приходу световых сигналов, изображаются эллипсом на этой диаграмме. Такой способ описания световых сигналов весьма полезен.

Диаграмме, изображенной на рис. 31.2, можно дать непосредственное физическое истолкование: свет распространяется в рассматриваемом пространстве-времени так, как будто показатель преломления анизотропен. Поскольку расстояния в пространстве-времени измеряются с помощью световых сигналов, анизотропия проявляется и при рассмотрении геометрии и приводит к неевклидовым эффектам, которые мы считаем обусловленными кривизной пространства-времени.

Пример

Взгляните на рис. 31.2 и установите, на какие его особенности указывает уравнение (31.22). Из (31.12) следует, что при неизменных r и t близлежащие концентрические сферы отстоят друг от друга дальше, чем можно ожидать согласно евклидовой геометрии. Как видно из рис. 31.2, сечения метрической фигуры — это эллипсы, сплюснутые в радиальном направлении. Однако здесь все согласуется, поскольку наш эталон длины определяется расстоянием, проходимым светом за некоторый стандартный промежуток времени, а уменьшение длины эталона приводило бы к увеличению значения измеряемого расстояния.

Существование симметрии по отношению к сдвигам во времени в значительной мере упростило проведенное в этом разделе обсуждение геометрии пространства-времени в окрестности Земли. Не будь ее, мы не смогли бы отдельно рассматривать геометрию пространственного сечения. В следующем разделе эта симметрия поможет нам упростить обсуждение гравитационного красного смещения.

ЗАДАЧИ

31.1. (14) Докажите, что, согласно нашим определениям, вращающийся диск стационарен, но не статичен. Придумайте такое пространство-время, чтобы оно было стационарным, но не статическим.

[Обратите внимание, как в ходе этого рассуждения мы то и дело переходим от касательных векторов к близким событиям и обратно. Так можно поступать, поскольку определение касательных векторов неразрывно связано с локальным поведением кривых, состоящих из событий. Но все это имеет силу только в пределе $\Delta t \rightarrow 0$.]

31.2. (22) Опираясь на идеи, изложенные в этом разделе, покажите, что сфера искривлена, а плоскость нет.

31.3. (18) В примере на стр. 259 вместо длины экватора найдите длину окружности, измеренную вдоль меридиана.

31.4. (25) Оцените точность эксперимента, необходимую для измерения кривизны пространства вблизи поверхности Земли.

31.5. (20) Дайте вывод порядка величины трех первых поправок, указанных на стр. 257.

31.6. (27) Докажите, что метрика

$$\mathcal{G} = \frac{dr^2}{1-r^2} + r^2 d\theta^2$$

описывает геометрию на сфере, но в непривычных координатах. Выясните, в каких пределах может изменяться координата r .

32. Гравитационное красное смещение

[Это будет ваш самый первый расчет, касающийся общей теории относительности!]

С концептуальной точки зрения в простейшем эксперименте в искривленном пространстве-времени должны использоваться только часы. В рамках нашего формализма часы — более простой инструмент, нежели жесткие стержни, хотя на практике это и не так. Сравнение скоростей хода часов, расположенных в разных местах, позволяет определить влияние поля тяготения на их наблюдаемый ход. Проявление такого влияния называется гравитационным красным смещением. Смещение может происходить как в сторону высоких, так и в сторону низких частот. Впервые явление смещения наблюдалось при исследовании света, приходящего к нам от фотосфер белых карликов. Спектральные линии оказались смещенными в красную часть спектра; отсюда и произошло название. Описанный эффект наблюдался также в лабораторных опытах с гамма-излучением и в эксперименте с ракетой, оборудованной прецизионными часами.

Симметрия и сравнение скоростей хода часов

Научившись осторожности в обращении даже со старыми представлениями, вы теперь наверняка ожидаете обсуждения вопроса о способе сравнения хода часов, движущихся с различными скоростями. Если дело обстоит действительно так, то все в порядке. Для сравнения хода можно было бы использовать, например, обмен световыми сигналами между часами. Такой метод применим почти всегда. Однако в данном случае можно ввести еще более широкий класс сравнений, причем все

они приводят к одному и тому же результату: ход часов можно сравнить, опираясь на симметрию пространства-времени по отношению к сдвигу во времени.

Будем считать сначала, что часы покоятся. (Смысл, который здесь вкладывается в слово «покоятся», связан с существованием симметрии по отношению к сдвигу во времени.) В этом случае векторное поле $\partial/\partial t$, генерирующее симметрию, дает нам естественную единицу измерения времени вдоль каждой мировой линии часов.

[Конечно, можно было бы обойтись без этого «тяжеловесного» утверждения. Но оно не бесполезно, ибо позволяет вплотную подойти к представлению о симметрии. Я специально отклонился от основной темы, чтобы ввести это абстрактное понятие, в надежде, что практическое применение делает его более осязаемым.]

Пример

Предположим, что для сравнения хода часов вы решили воспользоваться почтовой связью. Последней свойственна симметрия относительно сдвига во времени (письма во все времена шли, идут и будут идти одинаково медленно), поэтому ход часов можно сравнить, пошлав два письма. В первом вы напишите: «Начинайте отсчет». Во втором: «Прекращайте отсчет. Я насчитал N ударов». Тогда вашему адресату для сравнения хода часов потребуются лишь сравнить число N с числом ударов, отсчитанных им самим. Для повышения точности с отправлением второго письма можно подождать подольше.

Мы будем вычислять гравитационное красное смещение с помощью процедуры обмена световыми сигналами. При этом следует иметь в виду, что существование симметрии относительно сдвига во времени не только упрощает расчет, но и позволяет распространить его на любой другой процесс сравнения, если ему свойственна симметрия во времени. Знакомство с этим разделом позволит вам не только составить представление о гравитационном красном смещении, но и расширить понятие о симметрии и ее применениях.

Во избежание недоразумений следует отметить, что предстоящий расчет неприменим непосредственно к наземным экспериментам, о которых здесь будет идти речь. Дело в том, что используемые в них часы нельзя рассматривать как покоящиеся; напротив, следует считать (и совершенно обоснованно), что они вращаются вместе с Землей. Чтобы внести в ход часов поправки на это движение, достаточно учесть лишь спецрелятивистское доплеровское смещение. Процедура такой корректировки не настолько поучительна, чтобы о ней говорить в основном тексте; поэтому ее обсуждение перенесено в задачи.

Вращение не учитывается

[Сейчас самое время еще раз посмотреть расчет доплеровского смещения в разд. 13, поскольку здесь будут проделаны аналогичные вычисления.]

Геометрия пространства-времени

Для сравнения хода двух часов с помощью световых сигналов естественнее всего действовать следующим образом: сначала отыскать мировые линии световых сигналов, а затем найти способ, с помощью которого локально покоящиеся наблюдатели могут описать эти сигналы, используя только свои часы. Для простоты рассмотрим лишь случай, когда мировые линии часов вертикальны. Тогда световые сигналы будут распространяться только в плоскости (r, t) при неизменных значениях угловых координат θ и ϕ . Любой такой сигнал описывается параметризованной кривой

$$u \mapsto [u, R(u), \theta_0, \phi_0] \quad (32.1)$$

с касательным вектором

$$\sigma = \frac{\partial}{\partial t} + \frac{dR}{du} \frac{\partial}{\partial r}. \quad (32.2)$$

Поскольку специальная теория относительности по-прежнему локально справедлива, этот вектор должен удовлетворять условию

$$\mathcal{G} \cdot (\sigma, \sigma) = 0, \quad (32.3)$$

Отсюда получается обыкновенное дифференциальное уравнение

$$\frac{dR}{du} = 1 - \frac{2m}{r}. \quad (32.4)$$

[Здесь всюду безжалостно отбрасываются члены, квадратичные по (m/r) .]

Симметрия мировых линий световых сигналов

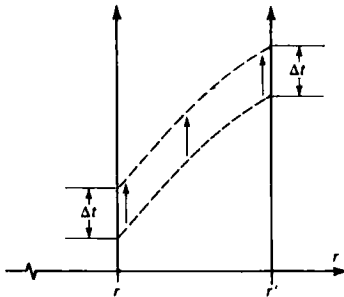


Рис. 32.1

Его решения точно определяют мировую линию светового сигнала. Обратите внимание, что симметрия по отношению к сдвигу во времени остается в силе: если $R(u)$ — любое решение уравнения (32.4), то и $R(u - u_0)$ — тоже решение этого уравнения при любом фиксированном значении u_0 . Это показано на рис. 32.1. Векторное поле $\partial/\partial t$ переводит одну мировую линию светового сигнала в другую. Так как использованная здесь координата t имеет непосредственное отношение к симметрии сдвига во времени, координатный интервал Δt между двумя мировыми линиями световых сигналов постояен.

Фактически, после того как мы выяснили, что координатный интервал между мировыми линиями двух световых сигналов постояен, нам больше не нужна никакая дополнительная информация о мировых линиях. Интервал собственного времени $\Delta\tau$, измеренный наблюдателем, покоящимся в точке, где радиус равен r , и соответствующий координатному интервалу Δt , определяется формулой

$$(\Delta\tau)^2 = -\mathcal{G} \cdot \left(\Delta t \frac{\partial}{\partial t}, \Delta t \frac{\partial}{\partial t} \right) = \left(1 - \frac{2m}{r} \right) (\Delta t)^2. \quad (32.5)$$

Различные наблюдатели, расположенные в точках r и r' , обнаружат, что отношение скоростей хода их часов равно

$$\frac{\Delta\tau'}{\Delta\tau} = 1 + \frac{m}{r} - \frac{m}{r'}. \quad (32.6)$$

В лабораторных экспериментах r и r' почти совпадают. Если записать

$$r' = r + H, \quad (32.7)$$

где разность уровней H удовлетворяет неравенству

$$H \ll r, \quad (32.8)$$

то предыдущее выражение можно заменить приближенной формулой

$$\frac{\Delta\tau'}{\Delta\tau} = 1 - \frac{mH}{r^2}. \quad (32.9)$$

В следующем разделе мы узнаем, что m — это не что иное, как масса гравитирующего тела, измеренная в геометрических единицах. Ускорение свободного падения

[По поводу этих единиц см. табл. 30.1. Масса Земли равна $1,48 \cdot 10^{-11}c$.]

$$g = \frac{m}{r^2} \quad (32.10)$$

теперь равно величине, обратной времени, необходимому, чтобы тело, двигаясь с таким ускорением, достигло скорости, сравнимой со скоростью света, а его единица измерения — (секунда)⁻¹. Ускорение свободного падения на поверхности Земли — это (примерно) год⁻¹. Теперь сдвиг частоты можно представить в виде

$$\frac{\Delta\tau'}{\Delta\tau} = 1 - gH. \quad (32.11)$$

При разности уровней H порядка 9 м, типичной для лабораторного эксперимента, произведение gH примерно равно 10^{-15} . Как мы видим, этот эксперимент осуществить не просто!

[Неплохой геометрической единицей является фут. Он равен приблизительно 10^{-9} световых секунд.]

Тем не менее такой эксперимент был поставлен. Роль световых сигналов в нем играли отдельные гребни волн гамма-излучения, испущенного при распаде ядра атома, а для сравнения частот использовался эффект Мессбауэра. Наблюдавшийся сдвиг частоты совпал с предсказаниями ОТО с точностью около 1%, равной точности эксперимента. (В действительности сначала был получен результат, который плохо согласовывался с теорией. Но затем было учтено замедление хода «встроен-

Реальный эксперимент

ных» в ядра часов, обусловленное их тепловым движением, что потребовало тщательного контроля за этим эффектом и внесения соответствующих поправок. Таким образом, был обнаружен не просто парадокс близнецов для ядерных часов, а сторонний эффект, который следовало учесть в других экспериментах.)

Следствия

Приятно сознавать, что результаты эксперимента согласуются с предсказаниями. Но можно ли на основании этого факта самого по себе считать ОТО единственно правильной теорией тяготения? Конечно, нет. Ведь может оказаться, что имеются конкурирующие теории, которые предсказывают точно такое же гравитационное красное смещение. Поскольку мы не располагаем всеми возможными теориями тяготения, остается лишь тщательно проанализировать, насколько наш вывод зависит от деталей, присущих именно ОТО. К сожалению, оказывается, что не очень сильно. Мы рассмотрим еще два способа вывода формулы гравитационного красного смещения, в которых несколько сокращено число исходных предположений. В результате выяснится, что почти любая теория тяготения предсказывает один и тот же сдвиг частоты.

Сохранение энергии

Для вывода формулы гравитационного красного смещения вполне достаточно закона сохранения энергии. Мы не доказывали, что этот закон выполняется в ОТО. В расширяющейся модели Вселенной вопрос о сохранении энергии на самом деле оказался очень сложным. Но в статическом пространстве-времени, с которым мы здесь имеем дело, никаких проблем не возникает.

Гравитационная машина

[Здесь используется спецрелятивистское соотношение между массой и энергией. Предполагается, что оно верно как для инертной, так и для гравитационной масс. В некоторых теориях тяготения наш подход считается неприемлемым.]

Рассмотрим две небольшие системы, каждая из которых способна накапливать энергию и излучать световые сигналы. При желании эти системы можно считать атомными ядрами с массами M и m ; тогда роль световых сигналов будут играть гамма-лучи. Каждая система должна быть способна поглощать гамма-излучение; в результате перехода в возбужденное состояние массы системы будут возрастать; обозначим их M^* и m^* .

Теперь обсудим замкнутый цикл, схематически представленный на рис. 32.2. Начнем с состояния 1, когда система m возбуждена, а M нет. Пусть M находится выше m . В состоянии 2 нижнее ядро освобождает в виде гамма-кванта энергию

$$h\nu = m^* - m. \quad (32.12)$$

Затем (состояние 3) этот гамма-квант достигает ядра M , имея уже новую частоту ν' . Предположим, что верхнее ядро M способно поглотить новую энергию кванта и перейти в возбуж-

[В этой формуле h — постоянная Планка, численно равная $1,83 \cdot 10^{-86} \text{ с}^2$.]

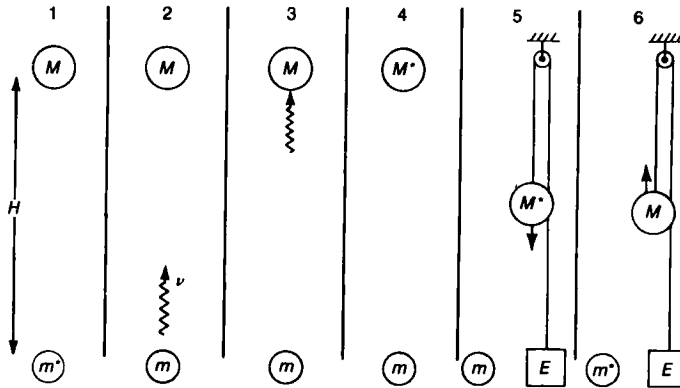


Рис. 32.2

К выводу формулы для гравитационного красного смещения.

денное состояние с массой M^* (состояние 4). Таким образом,

$$h\nu' = M^* - M. \quad (32.13)$$

В состоянии 5 мы медленно опускаем возбужденное ядро M^* , собирая его гравитационную энергию E_{down} , выделяющуюся при падении

$$E_{\text{down}} = M^* g H, \quad (32.14)$$

в какой-нибудь коллектор. (Высота обозначена прописной буквой H , чтобы ее нельзя было спутать с постоянной Планка). В самом низу отберем у M^* энергию возбуждения и поднимем M в исходное состояние. Но теперь это ядро стало легче, так что в результате внизу накопится энергия

$$h\nu' + E_{\text{down}} - E_{\text{up}} = (M^* - M)gH + h\nu', \quad (32.15)$$

где E_{up} — энергия, затрачиваемая на подъем. Энергии, накопленной внизу, должно быть ровно столько, чтобы вернуть ядро m в первоначальное возбужденное состояние. В противном случае, используя так или иначе этот цикл, мы могли бы получить неисчерпаемый источник энергии. Следовательно, должно выполняться равенство

$$h\nu = h\nu' (1 + gH), \quad (32.16)$$

откуда

$$\frac{\nu}{\nu'} = 1 + gH. \quad (32.17)$$

Красное смещение как следствие сохранения энергии

Таким образом, мы пришли к тому же результату, что и раньше.

Отсюда следует вывод, что всякая теория, в которой выполняются принцип эквивалентности инерциальной и гравитационной масс и локальный закон сохранения энергии, должна давать правильное описание гравитационного сдвига частоты. Такой вывод может либо разочаровать, если вы рассчитывали получить подтверждение ОТО как единственно правильной теории тяготения, либо ободрить, если вам нужна формула, описывающая сдвиг частоты, но вы не испытываете уверенности в справедливости общей теории относительности. Пока не существует никаких практических применений гравитационного сдвига частоты, в нас, естественно, преобладает чувство разочарования.

ЗАДАЧИ

32.1. (15) Докажите, что не существует гравитационного красного смещения между часами, расположенными в разных местах, но на одном и том же расстоянии от центра Земли.

32.2. (11) Найдите гравитационное красное смещение для часов, расположенных не на одной вертикальной прямой.

32.3. (33) Рассчитайте влияние вращения на лабораторный эксперимент, связанный с исследованием красного смещения.

32.4. (31) Рассчитайте гравитационное красное смещение с помощью точной метрики, приведенной в разд. 30. При этом величину (m/r) не следует считать малой.

32.5. (27) Объясните, как с помощью волновой диаграммы рассчитать доплеровское смещение для обычных диспергирующих волн, таких, как волны на воде или волны упругой деформации. Вместо введенных в разд. 28 встроенных часов воспользуйтесь обычными часами.

33. Принцип Гюйгенса и падающее яблоко

Теперь мы готовы показать, что искривленное пространство-время с метрикой

$$\mathcal{G} = \mathcal{N} + \frac{2m}{r} (dr^2 + dt^2) \quad (33.1)$$

Гравитация — это притяжение

(где $\mathcal{N}$ — метрика Минковского) описывает тяготение. Собственно, нам следует показать, что в таком пространстве

времени ускорения всех свободных тел направлены вниз. Сделаем это, опираясь на геометрический подход. Помимо доказательства падения тел такой подход позволит убедиться, что параметр m в (33.1) действительно представляет собой гравитационную массу притягивающего тела. Кроме того, будет показано, что любая теория тяготения, основанная на представлениях о волнах и волновых пакетах, приводит к гравитационному смещению частоты, совпадающему с только что описанным.

Назовем геометрический способ описания движения волновых пакетов принципом Гюйгенса, несмотря на то что они несколько отличаются друг от друга. Кроме того, ограничимся изучением высокочастотного предела, тогда как принцип Гюйгенса был предложен для обсуждения самого общего случая. Если речь идет об электронах, то высокочастотный предел означает, что гравитационное поле не должно быстро меняться на расстояниях порядка комптоновской длины волны ($3 \cdot 10^{-11}$ см); такое условие заведомо выполняется. У более тяжелых частиц комптоновская длина волны еще меньше.

Движение волнового пакета считается заданным, если в каждой точке известны вид волновой диаграммы или дисперсионное уравнение. Геометрический способ описания движения основан на применении волновой диаграммы. Предположим, что в каждой точке известна фаза волны. Чтобы 1-форма градиента фазы была решением исходного дифференциального уравнения в частных производных, она должна в каждой точке касаться волновой диаграммы. Такая фазовая функция изображена на рис. 33.1, а на рис. 33.2 и 33.3 показаны волновые диаграммы и градиенты фазы для событий A и B . Поскольку касательные пространства на рис. 33.2 и 33.3 представляют со-

Полуклассический подход

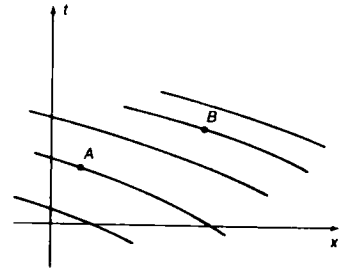


Рис. 33.1
Пространственно-временная диаграмма для гребней выбранного наугад цуга волн.

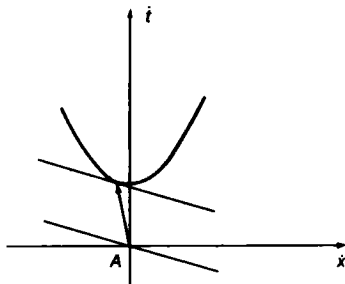


Рис. 33.2
Касательное пространство в точке A .

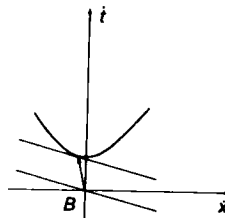


Рис. 33.3
Касательное пространство в точке B .

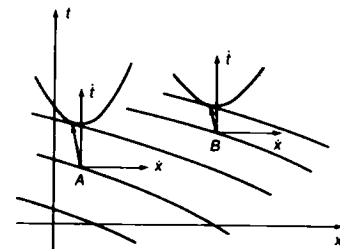


Рис. 33.4

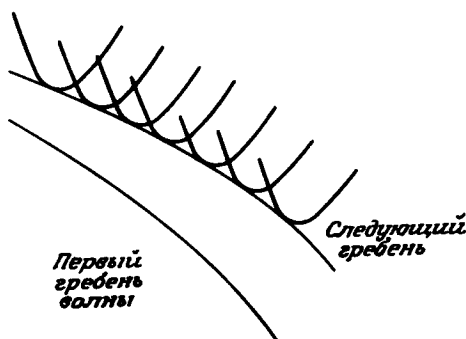


Рис. 33.5
Построение фронта волны с помощью принципа Гюйгенса.

[Для лучшего понимания описанных здесь геометрических построений полезно еще раз обратиться к рис. 26.7.]

Построение Гюйгенса

бой локальные аппроксимации рассматриваемого многообразия, их можно нанести прямо на само многообразие, что и сделано на рис. 33.4.

Построение Гюйгенса в некотором смысле обратно предыдущему. Его начинают с единственного заданного фронта волны. Чтобы построить следующий волновой фронт, нужно начертить волновые диаграммы для точек, лежащих на первом фронте, и провести их огибающую (рис. 33.5). Эта огибающая и будет новым фронтом волны. Чаще всего такое построение будет использоваться в ситуациях, когда имеется лишь минимальная косвенная информация. Если волновые пакеты олицетворяют собой частицы, то они не должны быть слишком растянутыми. Рассмотрим на рис. 33.6 только два близких события. Начнем с события A и проходящего через него гребня волны, помеченного на рисунке номером 1. Выберем на этом гребне близкое событие A' и начертим для каждого из них волновую диаграмму. Новый гребень волны, который мы обозначим

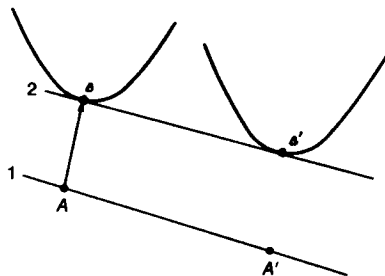


Рис. 33.6
Применение принципа Гюйгенса на малых расстояниях.

номером 2, согласно нашему определению, представляет собой общую касательную к этим двум диаграммам. Теперь можно построить вектор групповой скорости v и найти положение следующего события B на мировой линии волнового пакета. В результате в B мы будем иметь ту же информацию, что и в A ; следовательно, процедуру построения можно продолжить. Принципу Гюйгенса соответствует предельный процесс такого построения: событие A' следует выбрать столь близко к A , чтобы результат построения нового гребня не зависел от того, насколько тесно расположены эти события.

В общем случае при таком построении меняется направление вектора групповой скорости волнового пакета. Но изменение скорости означает, что у пакета есть ускорение. Таким образом, мы нашли способ описания динамики волновых пакетов, с помощью которого можно показать, что ускорение волновых пакетов в искривленном пространстве-времени вблизи Земли действительно представляет собой знакомое нам ускорение свободного падения. Но сначала испытаем свои силы на другом хорошо знакомом примере.

Ускорение

Пример

Попытаемся объяснить миражи. Дисперсионное уравнение для света, распространяющегося в среде с показателем преломления, зависящим от пространственных координат, и соответствующий градиент фазы имеют вид

Искривление лучей света

$$k_x^2 + k_y^2 + k_z^2 = n^2 \omega^2, \quad (33.2)$$

$$d\theta = k_x dx + k_y dy + k_z dz + \omega dt. \quad (33.3)$$

Это ковекторное уравнение определяет величины k_x , k_y , k_z и ω . Показатель преломления n может зависеть от x , y , z и ω . Если вы изучали оптику, то, возможно, знаете, как выводится это уравнение. (В противном случае его следует принять как вполне разумную исходную посылку.) Кроме того, договоримся, что в дальнейшем речь будет идти только о монохроматическом свете. Настоящий мираж — это смесь множества цветов. Он может иметь окраску, как в случае зеленого луча ¹⁾, а

¹⁾ Зеленый луч — вспышка зеленого света над диском Солнца при его заходе, наблюдаемая в течение нескольких секунд в момент исчезновения верхнего края солнечного диска за горизонтом. Это явление объясняется рефракцией солнечных лучей в верхних слоях атмосферы. См. Миннарт М. Свет и цвет в природе. — М.: Физматгиз, 1959. — Прим. ред.

может и не иметь. Частота монохроматического света ω всюду одинакова, даже если показатель преломления меняется от точки к точке. Следует также иметь в виду, что зависимость от времени решений линейного уравнения с коэффициентами, которые не являются функциями времени, может быть перенесена в показатель экспонент.

Так как с течением времени ничего нового не происходит, мы забудем о времени и рассмотрим только пространственный аспект нашей задачи. В результате получится усеченная задача, сходная по своей структуре с уже обсуждавшимися нами задачами о распространении волн, но в данном случае о распространении в пространстве, а не в пространстве-времени. Такой усеченной задаче соответствует дисперсионное уравнение

$$k_x^2 + k_y^2 + k_z^2 = \frac{n^2}{\lambda^2}, \quad (33.4)$$

[Проверьте это утверждение; см. задачу 15.1.]

где постоянная λ — это длина волны в вакууме. Проанализируем уравнение (33.4) в двух пространственных измерениях. Из него следует, что волновая диаграмма представляет собой окружность с радиусом, пропорциональным (n/λ) . Мираж возникает, когда вследствие более интенсивного прогревания и соответствующего расширения воздуха вблизи поверхности земли уменьшается его показатель преломления. С помощью волновых диаграмм эта ситуация изображена на рис. 33.7. Используя принцип Гюйгенса, можно показать, что лучи света, представляющие собой траектории волновых пакетов, в данном случае должны отклоняться вверх. Это хорошо видно из рис. 33.8. В силу такого искривления лучей человек, смотрящий вдаль, фактически увидит свет, пришедший с неба; ему будет казаться, что свет отразился от поверхности воды. Мираж такого типа хорошо знаком водителям автомобилей.

Воспользуемся теперь нашим геометрическим подходом, чтобы удостовериться в том, что в пространстве-времени с метрикой (33.1) ньютоново яблоко действительно будет падать и что входящий в эту формулу параметр m представляет собой массу притягивающего тела. На самом деле мы собираемся показать лишь, что падают отдельные частицы. Вопрос же о падении составного тела требует дополнительного обсуждения. Массы падающих тел, которые мы здесь будем рассматривать, предполагаются столь малыми, что они не влияют на гравитационное поле. Кроме того, такие малые массы (их называют пробными телами) не должны влиять и друг на друга. Соглас-

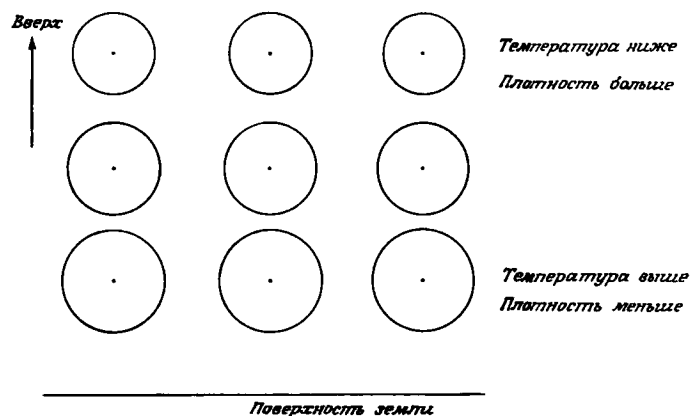


Рис. 33.7

Волновые диаграммы для света, распространяющегося над нагретой поверхностью (масштаб сильно преувеличен).

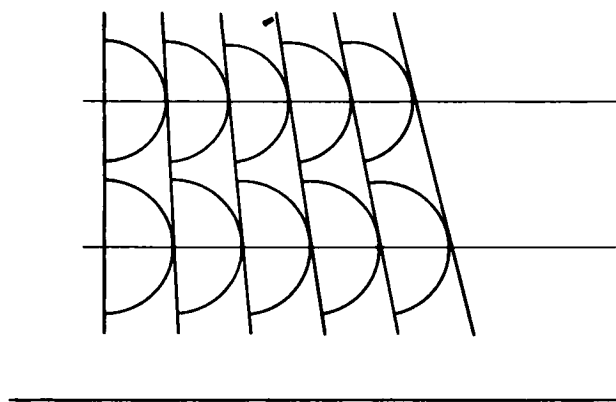


Рис. 33.8

Использование принципа Гюйгенса для определения наклона гребней волны.

но проведенным ниже расчетам, все пробные частицы падают с одинаковыми ускорениями, а значит и составное тело, построенное из таких частиц, будет падать с тем же ускорением. Сейчас мы постараемся записать наши геометрические рассуждения в виде алгебраических соотношений.

Локальные свойства пространства-времени

Начнем с изучения локального поведения волновых диаграмм для пространства-времени с метрикой (33.1). Будем рассматривать только радиальное движение, не обращая внимания на угловые координаты. В этом случае нам понадобится лишь часть исходной метрики, а именно

$$\mathcal{G} = -\left(1 - \frac{2m}{r}\right) dt^2 + \left(1 + \frac{2m}{r}\right) dr^2. \quad (33.5)$$

Эта метрика определяет вид волновой диаграммы, описывающей волновые пакеты. По мере перемещения вдоль радиуса волновые диаграммы претерпевают определенные изменения, которые легко изобразить схематически, если уяснить, что единичные 1-формы

$$\left(1 - \frac{m}{r}\right) dt, \quad \left(1 + \frac{m}{r}\right) dr \quad (33.6)$$

дуальны единичным векторам

$$\left(1 + \frac{m}{r}\right) \frac{\partial}{\partial t}, \quad \left(1 - \frac{m}{r}\right) \frac{\partial}{\partial r} \quad (33.7)$$

[Напомним, что у численно большей 1-формы линии уровня расположены ближе друг к другу и наоборот.]

(здесь опущены члены, квадратичные по m/r). При удалении от начала координат вдоль радиуса радиальный вектор растягивается, а времениподобный вектор сокращается. 1-формы ведут себя аналогичным образом. Все это показано на рис. 33.9. Теперь метрику (33.6) можно переписать через единичные 1-формы (33.7):

$$\begin{aligned} \mathcal{G} = & (-) \left(1 - \frac{m}{r}\right) dt \otimes \left(1 - \frac{m}{r}\right) dt + \\ & + \left(1 + \frac{m}{r}\right) dr \otimes \left(1 + \frac{m}{r}\right) dr. \end{aligned} \quad (33.8)$$

Соответствующая метрическая фигура изображена на рис. 33.10.

Размер лаборатории диктует характер приближения

В лабораторном эксперименте с «установкой», имеющей размер в несколько метров, величина r в пределах рассматриваемой области пространства меняется очень мало, что позволяет еще сильнее упростить наши расчеты. Пусть в точке, соответствующей на рис. 33.6 событию A , единичные 1-формы (33.6) имеют вид

$$\left(1 - \frac{m}{R}\right) dt, \quad \left(1 + \frac{m}{R}\right) dr, \quad (33.9)$$

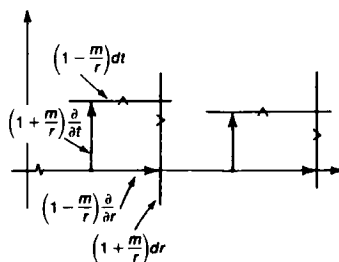


Рис. 33.9

1-формы $[1 + (m/r)]dr$ и $[1 - (m/r)]dt$ и векторы $[1 - (m/r)](\partial/\partial r)$, $[1 + (m/r)](\partial/\partial t)$. Масштабы по осям координат выбраны так, чтобы левая фигура была квадратом. Отклонение формы правой фигуры от квадратной сильно преувеличено.

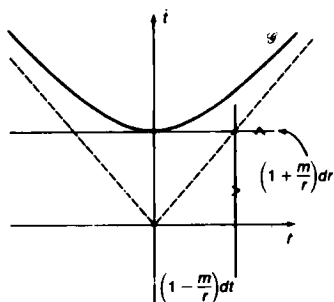


Рис. 33.10

Метрическая фигура и мировые линии световых сигналов на волновой диаграмме, выполненной с помощью единичных 1-форм.

где R — некоторое исходное значение радиуса. В точке с радиусом

$$r = R + H, \quad (33.10)$$

где $H \ll R$, соответствующей близкому событию B , единичные 1-формы (33.6) с помощью биномиального разложения можно записать в виде

$$\left(1 - \frac{m}{R} + \frac{mH}{R^2}\right)dt, \quad \left(1 + \frac{m}{R} - \frac{mH}{R^2}\right)dr. \quad (33.11)$$

Нам нужно найти ускорение, т.е. линейную часть приращения скорости со временем. Следовательно, чтобы вычислить ускорение, все расчеты требуется выполнять лишь с точностью до членов, линейных по H . Сказанное схематически представлено на рис. 33.11.

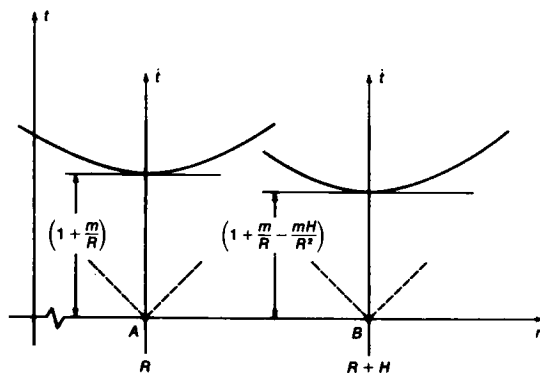


Рис. 33.11

Линейное изменение метрики в радиальном направлении.

[Еще раз обратите внимание на обратную зависимость между размером изображения и величиной 1-формы.]

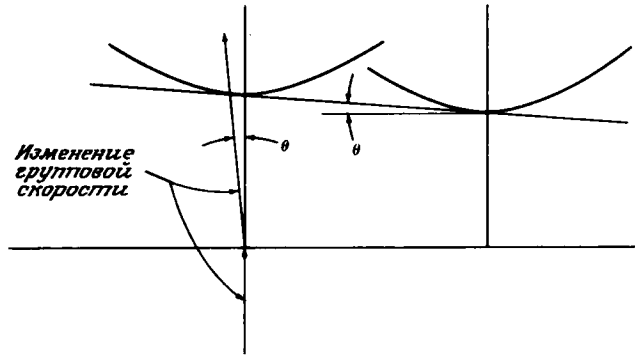


Рис. 33.12.

Ускорение волнового пакета в пространстве-времени, представленном на рис. 33.11.

Движение волнового пакета

Чтобы узнать, что произойдет с волновым пакетом в рассматриваемом пространстве-времени, воспользуемся геометрическим построением. Обсудим только простой случай, когда волновой пакет начинает движение из состояния покоя. Один гребень волны такого пакета должен располагаться на диаграмме совершенно горизонтально. Тогда следующий гребень будет касательной к волновым диаграммам, связанным с событиями A и B и расположенным на предыдущем гребне, как показано на рис. 33.12.

Не представляет особого труда рассчитать положение нового гребня в пространстве-времени. Вспомним, что

$$\frac{m}{R} \ll 1. \quad (33.12)$$

Наклон гребней волны

Тогда угол наклона θ гребня волны равен просто частному от деления разности вертикальных размеров волновых диаграмм, которая здесь равна mH/R^2 , на расстояние H между событиями A и B :

$$\theta = \frac{m}{R^2}. \quad (33.13)$$

[Заметьте, что H не входит в эту формулу. Именно это мы подразумевали на стр. 273 под словами: «... событие A' следует выбрать столь близко к A , чтобы результат ... не зависел от того, насколько тесно расположены эти события».]

Это угол наклона спустя единицу времени после начала движения пакета. Однако длина волны, с которой нам здесь придется иметь дело, должна быть намного короче одной световой секунды, так что наклон спустя один период волны Δt после начала движения будет соответственно меньше

$$\theta = \frac{m \Delta t}{R^2}. \quad (33.14)$$

Поскольку $m/R \ll 1$, малым различием между координатным и собственным временами всюду можно пренебречь.

Все вроде бы неплохо, но создается впечатление, что волновой пакет уходит вверх, т.е. в сторону больших значений радиуса, тогда как ему полагалось бы отклоняться вниз. Однако не будем торопиться с выводами; здесь надо быть осторожнее. Ведь для того, чтобы сказать, куда движется волновой пакет, следует пользоваться групповой скоростью, а не гребнями волн. Групповую скорость можно найти с помощью волновой диаграммы. При этом нужно иметь в виду, что здесь, как и в СТО, углы ведут себя весьма специфическим образом. Так как наши диаграммы отличаются от спецрелятивистских всего на одну миллиардную долю, мы вполне можем воспользоваться теми сведениями о волновых диаграммах в СТО, которыми располагаем. В частности, мы знаем, что вектор групповой 4-скорости наклонен под таким же углом, как и волновой фронт, но в противоположном направлении. Таким образом, отклонение волнового фронта вверх соответствует падающим частицам. Изменение скорости волнового пакета за единицу времени равно

$$\frac{\Delta v}{\Delta t} = \frac{m}{R^2}, \quad (33.15)$$

а это и есть ускорение свободного падения

$$g = \frac{m}{R^2}. \quad (33.16)$$

Заметьте, что ускорение свободного падения не зависит от фактической длины волны волнового пакета. Частицы с разными массами обладают разными длинами волн, но ускорение свободного падения для них одинаково. Этот фундаментальный результат называется принципом эквивалентности. В его пользу с чрезвычайно высокой точностью свидетельствуют результаты многочисленных экспериментов. Различие в ускорениях свободного падения для объектов лабораторных масштабов не превышает $10^{-12} g$. Один из первых экспериментов такого рода был выполнен Этвешем и носит его имя. (Согласно его собственному результату, полученному в 1905 г., различие ускорений не превышает 10^{-8} .)

Теперь вернемся немного назад и обсудим сложившуюся ситуацию. Для локального выполнения СТО необходимо, чтобы всякая теория тяготения опиралась на волновые диаграммы, близкие к спецрелятивистским. Далее, чтобы частицы падали,

Ускорение

[Это обсуждалось в примере на стр. 220.]

Эксперимент Этвеша

Обсуждение

волновые фронты должны быть наклонены, причем под совершенно определенным углом. Это в свою очередь требует сжатия волновых диаграмм в направлении оси времени по мере смещения от начала координат вдоль радиуса. Но именно такое сжатие необходимо для получения гравитационного красного смещения. Итак, теперь мы располагаем еще более общим выводом красного смещения. Вообще, любая теория тяготения, в которой частицы могут описываться волновыми уравнениями, приводит к правильному значению гравитационного красного смещения.

ЗАДАЧИ

33.1. (19) Если мираж образуется над нагретым шоссе, то вполне можно считать, что $n \approx 1$, а градиент $dn/dz \approx \approx 0,00005 \text{ м}^{-1}$. Найдите, на какой угол отклонится свет, прошедший вдоль дороги расстояние порядка километра.

33.2. (27) Обратите внимание, что в расчет гравитационного красного смещения входит только tt -компонента метрического тензора. Докажите, что rr -компоненту можно изменить с помощью преобразований координат, сохраняющих сферическую симметрию пространства-времени и одновременно симметрию по отношению к сдвигу во времени, тогда как tt -компонента при таких преобразованиях не меняется.

33.3. (19) На сфере можно ввести так называемые стереографические координаты, которые получаются, если спроецировать точки сферы из южного полюса на плоскость, касательную к северному полюсу, как показано на рис. 33.13. В таких координатах волновые диаграммы представляют собой окружности с радиусами, пропорциональными $(1 + r^2)$. Исследуйте графически геодезические этого пространства.

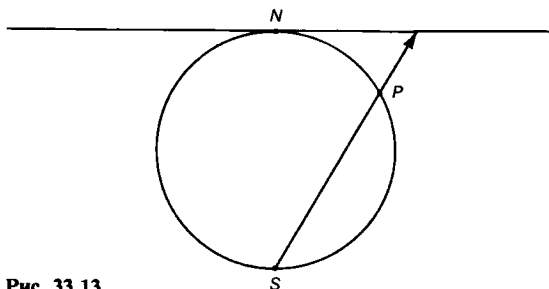


Рис. 33.13

33.4. (28) Найдите приближенный вид дисперсионного уравнения свободной частицы

$$g^{\alpha\beta}p_\alpha p_\beta + b^2 = 0.$$

[Здесь величина b пропорциональна массе частицы.]

для малых расстояний над поверхностью Земли. Изменив масштаб вертикальной координаты, запишите это уравнение в более простой форме

$$\left(1 + \frac{2mx}{R^2}\right)\omega^2 = k^2 + b^2,$$

где x — новая вертикальная координата.

33.5. (37) Найдите групповую скорость и волновую диаграмму для упрощенного дисперсионного уравнения из задачи 33.4. Найдите k и ω для $x < 0$, полагая, что при $x = 0$ скорость частицы равна нулю. Сделайте аккуратный набросок гребней волны и лучей.

33.6. (38) Распространите выполненное в этом разделе исследование падающей частицы на случай частиц, движущихся со скоростями, сравнимыми со скоростью света.

33.7. (18) Какое ускорение показано на рис. 33.12, если диаграмма метрики $\mathcal{S}$ соответствует интервалу времени 10^{-9} с?

33.8. (04) Куда направлен поток энергии на рис. 33.8?

33.9. (26) Исследуйте геодезические в верхней полуплоскости, т.е. на той части плоскости x, y , где $y > 0$, если метрические фигуры представляют собой окружности радиусом y .

34. Геодезические

Каждое пространство-время, используемое далее в космологических моделях, столь симметрично, что мы сможем найти мировые линии интересующих нас частиц, опираясь лишь на критерии симметрии и на представление о динамике волнового пакета, составленное с помощью принципа Гюйгенса. Дать описание мировых линий свободных частиц даже в произвольном пространстве-времени настолько просто, что не уделить ему внимание было бы по крайней мере неразумно. Кроме того, ниже с эвристических позиций будет доказано, что мировые линии свободных частиц должны удовлетворять принципу минимальности. Именно по этой причине их называют геодезиче-

скими, т.е. кривыми наименьшей длины¹⁾. Мы увидим, что уравнения для геодезических, являющиеся аналогами уравнений Гамильтона в классической механике, представляют собой естественное следствие нашего полуклассического подхода.

Частица в неоднородном гравитационном поле описывается дисперсионным уравнением

Дисперсионное уравнение

$$g^{\alpha\beta} p_\alpha p_\beta = -b^2. \quad (34.1)$$

Здесь и далее для удобства и для выработки у читателя практических навыков работы с такой формой записи используются индексные обозначения и правило суммирования по неммым индексам. Величины p_α — это компоненты градиента фазы

$$d\theta = p_\alpha dx^\alpha, \quad (34.2)$$

а $g^{\alpha\beta}$ — компоненты тензора, обратного $g_{\alpha\beta}$ и удовлетворяющего по определению соотношению

$$g^{\mu\alpha} g_{\alpha\nu} = \delta^\mu_\nu, \quad (34.3)$$

где $g_{\alpha\beta}$ — компоненты метрического тензора

$$\mathcal{G} = g_{\mu\nu} dx^\mu \otimes dx^\nu. \quad (34.4)$$

Как p_α , так и $g^{\alpha\beta}$ могут зависеть от положения в пространстве-времени. Постоянная b зависит от массы частицы и, очевидно, выпадает во всех вычислениях, связанных с выполнением операции дифференцирования. В оставшейся части этого раздела будем для простоты считать, что $b = 1$, поскольку, как уже было сказано, движение свободных частиц не зависит от их массы.

Пространственные геодезические

Пример 1

Сформулируем задачу об отыскании геодезических на поверхности сферы. С точностью до изменения нескольких знаков она совершенно аналогична задаче поиска мировых линий волновых пакетов в пространстве-времени. Метрика на поверхности

¹⁾ Определения геодезической линии в различных пространствах зависят от того, какая из структур (метрика, линейный элемент, линейная связность) лежат в основе геометрии рассматриваемого пространства. Определение автора соответствует случаю, когда метрика задана априори. — *Прим. ред.*

сферы имеет вид

$$\mathcal{G} = R^2(d\lambda^2 + \cos^2 \lambda d\phi^2), \quad (34.5)$$

где вместо обычной сферической координаты θ , которая измеряется от Северного полюса, используется координата λ — широта, измеряемая от экватора. Для этой метрики

$$\begin{aligned} g_{\lambda\lambda} &= R^2, \\ g_{\phi\phi} &= R^2 \cos^2 \lambda, \\ g_{\lambda\phi} &= g_{\phi\lambda} = 0, \end{aligned} \quad (34.6)$$

а

$$\begin{aligned} g^{\lambda\lambda} &= \frac{1}{R^2}, \\ g^{\phi\phi} &= \frac{1}{R^2 \cos^2 \lambda}, \\ g^{\lambda\phi} &= g^{\phi\lambda} = 0. \end{aligned} \quad (34.7)$$

Обозначим компоненты градиента фазы буквами l и n :

$$p = l d\lambda + n d\phi; \quad (34.8)$$

тогда

$$p_\lambda = l, p_\phi = n. \quad (34.9)$$

Следовательно, дисперсионное уравнение можно записать в форме

$$\frac{1}{R^2} \left(l^2 + \frac{n^2}{\cos^2 \lambda} \right) = 1. \quad (34.10)$$

[Этот пример будет продолжен после того, как мы несколько разовьем общую теорию.]

Поскольку p — градиент, т.е.

$$p_\alpha = \frac{\partial \theta}{\partial x^\alpha}, \quad (34.11) \quad \text{Групповая скорость}$$

должно выполняться условие симметрии его производных

$$\frac{\partial p_\alpha}{\partial x^\beta} = \frac{\partial p_\beta}{\partial x^\alpha}. \quad (34.12)$$

Примем во внимание, что p_α и $g^{\alpha\beta}$ — функции положения, и применим к дисперсионному уравнению (34.1) операцию градиента; тогда получим

$$\frac{\partial g^{\alpha\beta}}{\partial x^\sigma} p_\alpha p_\beta + 2g^{\alpha\beta} \frac{\partial p_\alpha}{\partial x^\sigma} p_\beta = 0. \quad (34.13)$$

Отсюда после соответствующей замены частных производных

следует

$$\frac{\partial p_\alpha}{\partial x^\alpha} (-g^{\alpha\beta} p_\beta) = \frac{1}{2} \frac{\partial g^{\alpha\beta}}{\partial x^\sigma} p_\alpha p_\beta. \quad (34.14)$$

[Производные по направлению обсуждались на стр. 139.]

Левая часть этого равенства представляет собой производную величин p_α по направлению

$$v = -g^{\alpha\beta} p_\beta \frac{\partial}{\partial x^\alpha}, \quad (34.15)$$

которое мы договорились считать направлением вектора групповой 4-скорости. Согласно дисперсионному уравнению,

$$p_\alpha v^\alpha = 1, \quad (34.16)$$

[Напомним, что постоянная b была положена равной единице.]

так что групповая 4-скорость уже имеет должную нормировку. Вектор v определяет в пространстве-времени кривые, которые являются мировыми линиями наших волновых пакетов. Если представить эти мировые линии в виде отображения

$$s \mapsto X^\alpha(s), \quad (34.17)$$

то должны выполняться равенства

$$\frac{dX^\alpha}{ds} = v^\alpha = -g^{\alpha\beta} p_\beta. \quad (34.18)$$

Таким образом, полученное выше уравнение для p_α определяет производную от этих величин вдоль линий (34.17):

Уравнение изменения групповой 4-скорости

$$\frac{dp_\alpha}{ds} = \frac{1}{2} \frac{\partial g^{\beta\gamma}}{\partial x^\alpha} p_\beta p_\gamma. \quad (34.19)$$

Вместе уравнения (34.18) и (34.19) образуют систему восьми обыкновенных дифференциальных уравнений, описывающую мировые линии волновых пакетов.

[Продолжение примера 1.]

В данной задаче мы оперируем локально-евклидовым типом метрики, из чего следует необходимость изменения знака v . Теперь нужно пользоваться уравнениями

$$\frac{dX^\alpha}{ds} = g^{\alpha\beta} p_\beta, \quad (34.20)$$

$$\frac{dp_\alpha}{ds} = -\frac{1}{2} \frac{\partial g^{\beta\gamma}}{\partial x^\alpha} p_\beta p_\gamma, \quad (34.21)$$

которые в исследуемом здесь случае примут вид

$$\frac{d\lambda}{ds} = \frac{l}{R^2}, \quad (34.22)$$

$$\frac{d\phi}{ds} = \frac{n}{R^2 \cos^2 \lambda}, \quad (34.23)$$

$$\frac{dl}{ds} = -\frac{n^2 \sin \lambda}{R^2 \cos^3 \lambda}, \quad (34.24)$$

$$\frac{dn}{ds} = 0. \quad (34.25)$$

Это четыре обыкновенных дифференциальных уравнения для четырех неизвестных.

Пусть мы хотим найти геодезическую, которая начинается в точке (λ_0, ϕ_0) и идет в таком направлении, что

$$\frac{d\lambda}{ds}(0) = 0. \quad (34.26)$$

Для решения полученной системы уравнений нужно располагать также начальными значениями величин l и n . Из первого уравнения следует, что

$$l_0 = 0, \quad (34.27)$$

а из дисперсионного уравнения имеем

$$n_0 = R \cos \lambda_0. \quad (34.28)$$

С помощью этих начальных значений теперь можно найти решения уравнений (34.22)—(34.25), а следовательно, и геодезические на поверхности сферы.

[Мы продолжим изучение этих дифференциальных уравнений в следующем разделе, когда приступим к обсуждению кривизны сферической поверхности.]

Распространяя только что описанную процедуру на уравнения (34.18), (34.19), можно найти траектории свободных частиц в любом пространстве-времени, если, конечно, удастся решить окончательную систему дифференциальных уравнений. С геометрической точки зрения на полученные нами восемь дифференциальных уравнений первого порядка можно смотреть как на векторное поле в восьмимерном пространстве с полным набором обобщенных координат q и p . Это весьма плодотворный подход, но его обсуждение увело бы нас слишком далеко от преследуемой цели.

Описывать динамику волновых пакетов можно также на основе принципа минимума. Тем, кто изучал классическую механику, этот принцип должен быть знаком. (Если вы прежде не сталкивались с ним, то достаточно беглого знакомства со следующими ниже рассуждениями.) Принципы минимума играют в физике очень важную роль и, несомненно, заслуживают изучения.

Уравнения геодезической

[Если вы знакомы с классической механикой, то узнаете в них уравнения Гамильтона. К тому же вы обнаружите здесь истолкование канонических импульсов, которые обычно предстают как нечто абстрактное. В рамках нашего геометрического подхода они являются компонентами градиента фазы волнового пакета.]

Принципы минимума

[Вернее их следовало бы называть принципами экстремума, поскольку они относятся не только к минимальным, но и к максимальным и стационарным значениям. Однако по историческим причинам их все же часто называют принципами минимума.]

Принципы минимума позволяют получить простые и эффективные формулировки физических законов. В качестве простого примера можно привести следующее утверждение: статическое равновесие частицы достигается в точках экстремума (т.е. там, где исчезают первые производные) потенциальной энергии. Все дело в том, что малые смещения системы в произвольном направлении от этих точек приводят, как показано на рис. 34.1, лишь к изменениям второго порядка малости в потенциальной энергии.

Динамика диспергирующих волновых пакетов не представляет исключения, ее вполне можно описать с помощью принципа минимума. На рис. 34.2 изображена мировая линия такого волнового пакета. В любой точке этой кривой можно построить 1-форму градиента фазы. Полное изменение фазы волнового пакета вдоль заданной мировой линии можно найти простым сложением всех изменений, претерпеваемых фазой вдоль этой линии. Такое полное изменение фазы называется

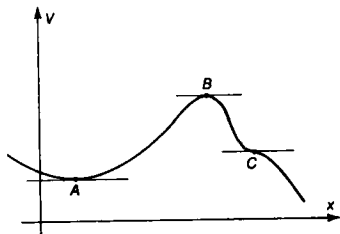


Рис. 34.1
В каждой из точек A , B и C достигается экстремум функции $V(x)$, но лишь точке A соответствует состояние устойчивого равновесия.

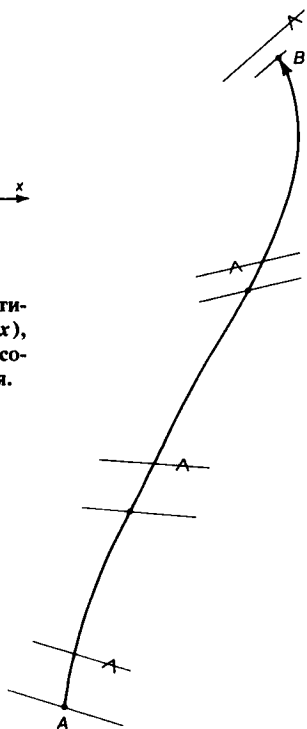


Рис. 34.2
Вычисление действия вдоль заданной кривой.

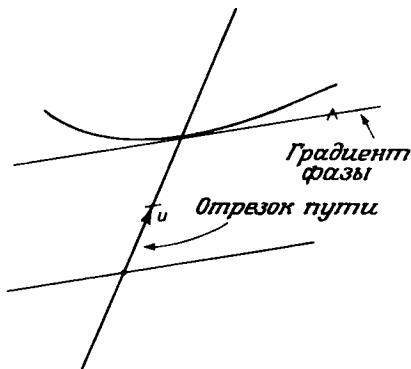


Рис. 34.3
Изменение фазы на коротком отрезке мировой линии u . Показанное здесь изменение фазы равно $1/2$.

действием и может быть вычислено следующим образом. Мысленно разобьем мировую линию на множество маленьких отрезков, каждый из которых будем считать прямой линией. Предположим, что эта процедура разбиения продолжается до бесконечности. Как найти изменение фазы вдоль одного из отрезков? В пределе любой такой отрезок можно рассматривать как вектор; назовем его u . Схематически ситуация представлена на рис. 34.3. Изменение фазы вдоль короткого отрезка мировой линии равно

$$\Delta\theta = d\theta \cdot u. \quad (34.29)$$

Пример 2

Рассчитаем изменение фазы вдоль прямолинейной мировой линии между точками $(0, 0)$ и $(t, x) = (2, 1)$ в случае, когда волновая диаграмма имеет вид

$$\dot{t}^2 - \dot{x}^2 = 1 \quad (34.30)$$

(рис. 34.4). Градиент фазы в каждой точке этой мировой линии определяется с помощью метрики Минковского $\mathcal{N}$ по формуле

$$d\theta = -\mathcal{N} \cdot v, \quad (34.31)$$

где v — групповая 4-скорость. Вектор v должен быть параллелен мировой линии; кроме того, его компоненты должны удовлетворять соотношению (34.30). Из названных условий следует, что

$$v = \frac{1}{\sqrt{3}} \left(\frac{\partial}{\partial x} + 2 \frac{\partial}{\partial t} \right). \quad (34.32)$$

Тогда градиент фазы имеет вид

$$\begin{aligned} d\theta &= (-dt^2 + dx^2) \cdot \frac{(-1)}{\sqrt{3}} \left(\frac{\partial}{\partial x} + 2 \frac{\partial}{\partial t} \right) = \\ &= \frac{1}{\sqrt{3}} (2 dt - dx). \end{aligned} \quad (34.33)$$

Так как мировая линия прямолинейна, а градиент фазы постоянен, в результате разбиения на малые отрезки и перехода к пределу ничего не изменится. Поэтому полное изменение фазы вдоль данной мировой линии, т.е. действие S , равно

$$\begin{aligned} S &= d\theta \cdot \left(\frac{\partial}{\partial x} + 2 \frac{\partial}{\partial t} \right) = \\ &= \frac{1}{\sqrt{3}} (4 - 1) = \sqrt{3}. \end{aligned} \quad (34.34)$$

Действие

[Еще раз обратите внимание на различие между $d\theta$, т.е. 1-формой $d\theta$, и $\Delta\theta$ — малым изменением фазы θ .]

[По этому поводу см. стр. 222.]

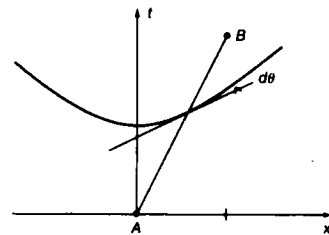


Рис. 34.4

Наименьшее действие

Фактически можно вычислить действие для какой угодно мировой линии, проходящей между любыми точками A и B . *Истинная мировая линия, вдоль которой распространяется волновой пакет, обладает тем свойством, что при ее малых вариациях действие остается неизменным.* Сейчас мы дадим графическое обоснование последнего утверждения, ибо подробный расчет требует более сложной техники, чем та, которую мы сочли необходимой здесь развивать.

Графическое построение

Рассмотрим два события и соединяющую их любую мировую линию. В каждой точке мировой линии можно построить градиент фазы, которым должен обладать волновой пакет, распространяющийся в данном направлении. Следовательно, в непосредственной окрестности этой линии можно начертить гребни волны. Если рассматриваемая мировая линия истинная, т.е. описывает историю волнового пакета, то локальные гребни можно построить, пользуясь принципом Гюйгенса. Чтобы перейти к действию, мы сначала упростим изображение исследуемой мировой линии, выполнив ряд соответствующих преобразований координат. Эта процедура совершенно аналогична линейным преобразованиям координат, которыми мы пользовались для упрощения задач в пространствах с геометриями Евклида и Минковского. Процесс упрощения показан на рис. 34.5. Вначале с помощью подходящего сдвига можно выпрямить мировую линию. Разумеется, в общем случае такой сдвиг не является линейной операцией. Выполнив дополнительные преобразования сдвига, можно добиться, чтобы все гребни

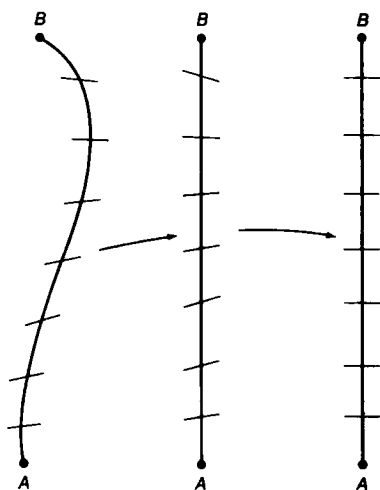


Рис. 34.5

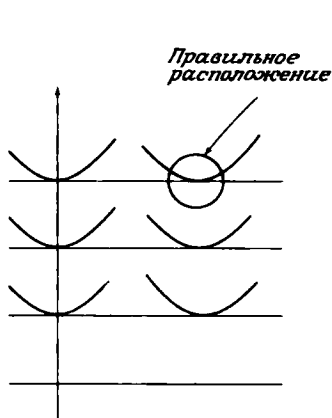


Рис. 34.6
«Правильное» поведение близких волновых диаграмм; проведены также поверхности постоянной фазы.

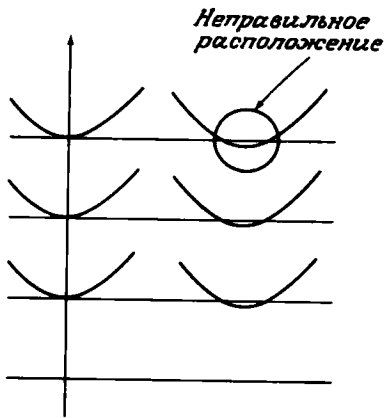


Рис. 34.7
«Неправильное» поведение близких волновых диаграмм.

волны выглядели перпендикулярными мировой линии. Наконец, с помощью дифференциального растяжения вдоль мировой линии можно сделать так, чтобы гребни располагались на одинаковом расстоянии друг от друга. Все эти упрощения не обязательны. При небольшом навыке можно проследить весь ход рассуждений, не прибегая к таким преобразованиям, но обычно бывает полезно максимально упростить представление кривой.

На рис. 34.6 показаны часть мировой линии, градиенты фазы и относящиеся к делу волновые диаграммы. Здесь всюду предполагается, что пространство-время точно определено; следовательно, в каждой точке можно построить волновую диаграмму. Кроме того, на рисунке изображены волновые диаграммы, расположенные близко к мировой линии; как мы видим, эта мировая линия в самом деле удовлетворяет принципу Гюйгенса. На рис. 34.7 показана аналогичная мировая линия, но не удовлетворяющая названному принципу. На обоих рисунках волновые диаграммы, изображенные слева, касаются гребней волны, что объясняется нашим способом построения этих гребней. Но, согласно принципу Гюйгенса, близкие волновые диаграммы тоже должны касаться гребней волны. Последнее имеет место на рис. 34.6, но не на рис. 34.7. На каждом рисунке кружком выделен один пример касания или его отсутствия.

Теперь займемся малыми вариациями мировой линии: сместим ее чуть-чуть в сторону и посмотрим, чему равно измене-

Простое представление

Вариация мировой линии

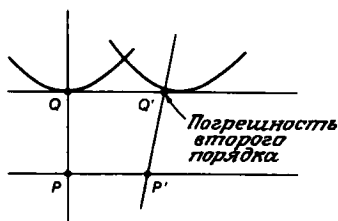


Рис. 34.8

Изменения фазы на отрезках путей PQ и $P'Q'$ совпадают с учетом членов, линейных по расстоянию PP' .

[Малые вариации волновых диаграмм уже использовались в разд. 33 при обсуждении падающего яблока.]

Кратчайшие линии

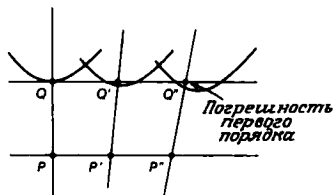


Рис. 34.9

Расхождение изменений фаз, линейно зависящих от вариации мировой линии.

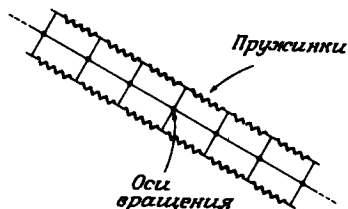


Рис. 34.10

Идеализированная модель упругого стержня.

ние фазы, линейно зависящее от вариации. Для определенности выберем какой-нибудь короткий отрезок мировой линии, скажем заключенный между точками P и Q на рис. 34.8, и сдвинем его в сторону. Чтобы вычислить изменение фазы между точками P' и Q' , нужно построить волновую диаграмму в точке P' . Если мировая линия удовлетворяет принципу Гюйгенса, то новая волновая диаграмма будет касаться старого гребня волны. Тогда с точностью до членов второго порядка малости изменение фазы между точками P' и Q' будет равно ее изменению между P и Q . Если же мировая линия не удовлетворяет этому принципу, как на рис. 34.9, то различие фаз, обусловленное вариацией, будет линейно по величине вариации. На рисунке показано также, что в два раза большему вариационному смещению соответствует удвоенная разность фаз. Само собой разумеется, что такого рода строгая линейность имеет место только при бесконечно малых смещениях мировой линии. Таким образом, действие на самом деле стационарно только для истинных мировых линий, что и предполагалось показать.

Подобный принцип наименьшего действия чрезвычайно полезен. С его помощью человек, знакомый с вариационным исчислением, может написать дифференциальные уравнения для мировых линий волновых пакетов, зная лишь суть поставленной задачи. Более того, принцип наименьшего действия дает интуитивное представление об этих мировых линиях как о кривых стационарной длины. Такие кривые называются геодезическими. Геодезические в пространстве-времени — это обычно линии наибольшего интервала времени (вспомним парадокс близнецов). Благодаря свойству экстремальности геодезические находят и много других приложений. На рис. 34.10 показана идеализированная модель упругого стержня. Сравнив рис. 34.10 с рис. 34.8 и 34.9, вы убедитесь, что состояния с минимальной энергией упругий стержень достигает только вдоль геодезической. Еще один пример. На рис. 34.11 изображена миниатюрная двухколесная тележка. Колеса жестко закреплены на оси, так что каждое должно проходить одно и то же расстояние. Тогда, если нет скольжения, тележка будет двигаться по поверхности вдоль геодезической.

ЗАДАЧИ

34.1. (14) Докажите, что уравнение геодезической можно рассматривать как «закон сохранения» дисперсионного уравнения. Другими словами, докажите, что дисперсионное уравнение дает одну константу движения.

34.2. (18) Предположим, что дисперсионное уравнение не зависит от какой-то определенной координаты. Покажите, что в этом случае возникает еще одна константа движения.

34.3. (30) Решите уравнения геодезической на евклидовой плоскости с метрикой

$$\mathcal{G} = dr^2 + r^2 d\phi^2,$$

записанной в полярных координатах.

34.4. (15) Запишите уравнения геодезической для метрики

$$\mathcal{G} = \frac{dx^2 + dy^2}{y^2}.$$

34.5. (33) Покажите, что геодезические в пространстве с метрикой из предыдущей задачи представляют собой окружности с центрами на оси x .

34.6. (15) Выпишите уравнения геодезической, как в примере на стр. 284, но для пространства с метрикой

$$\mathcal{G} = R^2(d\theta^2 + \sin^2 \theta d\phi^2).$$

34.7. (22) Сделайте то же самое, что и в задаче 34.6, но для метрики

$$\mathcal{G} = R^2(d\chi^2 + \text{sh}^2 \chi d\phi^2).$$

34.8. (37) Распространите условие задачи 31.5 на случай, когда необходимо учесть влияние Солнца и Луны. Не забывайте, что система координат, связанная с Землей, свободно падает на Солнце и Луну.

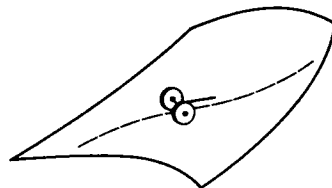


Рис. 34.11
Миниатюрная двухколесная тележка.

35. Кривизна пространства-времени

В качестве шага в развитии геометрии пространства-времени естественно обсудить поведение близких геодезических. Неевклидов характер их поведения свидетельствует о кривизне исследуемого пространства-времени. Для целей, которые мы преследуем в этой книге, вполне достаточно лишь интуитивного представления о кривизне. Чтобы научиться свободно обращаться с кривизной, необходимо овладеть сложным математическим аппаратом, затратив усилия, которые оправданы для учебника по ОТО, но не здесь. Мы и так уже рискнули уделить математике слишком большое внимание. Предлагаемое ниже краткое обсуждение кривизны поможет вам уяснить смысл

уравнений Эйнштейна. Однако вам придется принять на веру, что из них могут быть получены уравнения Фридмана, описывающие динамику стандартных космологических моделей.

Можно рассматривать как пространственноподобные, так и времениподобные геодезические. В стационарном пространстве-времени пространственные геодезические описывают форму прямых стержней. Пожалуй, лучшее описание кривизны пространства-времени, которое можно предложить, основано на представлении о конструкции из прямых стержней, которую договоримся называть *геодезическим квадратом*.

Важно понимать, что существует способ операционального определения прямого стержня. Для этого в дополнение к стержню нужно подобрать калибр с точно подогнанным под стержень прямым глубоким отверстием (рис. 35.1). Тогда, если стержень плотно входит в отверстие при всех восьми возможных ориентациях, мы будем называть его прямым. Следует отметить, что поперечное сечение должно быть квадратным и должно отсутствовать продольное кручение стержня.

Геодезический квадрат

[В искривленном пространстве-времени такая конструкция не может быть настоящим квадратом — и в этом вся суть, — но мне не хочется называть его геодезическим псевдоквадратом. Не могу я согласиться и с теми, кто будет настаивать, что здесь в действительности обсуждался прямоугольник.]

[Для доказательства отсутствия продольного кручения достаточно приложить два стержня друг к другу боковыми сторонами.]

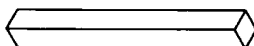
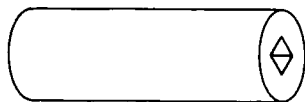


Рис. 35.1
Калибр для проверки прямого стержня.

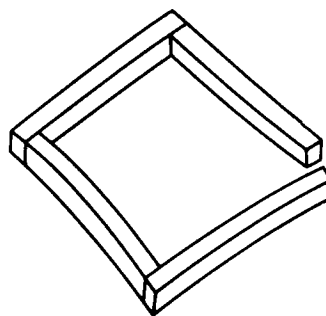


Рис. 35.2
Геодезический квадрат.

Перекошенный квадрат

Теперь возьмем четыре таких прямых стержня одинаковой длины и попытаемся построить из них квадрат, как на

рис. 35.2. Чуть выше уже отмечалось, что, вопреки надеждам, навешанным евклидовой геометрией, квадрат в общем случае будет разомкнутым. Кроме того, его концы будут повернуты по отношению друг к другу. Такое устройство очень удобно для измерения кривизны пространства-времени.

Пример 1

Построим геодезический квадрат на поверхности сферы, где геодезические — это знакомые нам большие круги. Для простоты разместим квадрат, как на рис. 35.3, расположив одну его сторону на экваторе, а две других на меридианах. Чтобы найти положение четвертой стороны, нужно проделать кое-какие расчеты.

Введем на сфере те же координаты (λ, ϕ) , что и в разд. 34. В таких координатах три заданные стороны квадрата представляют собой следующие параметризованные кривые:

$$\overline{AB}: s \mapsto (\lambda, \phi) = \left(0, \frac{s}{R}\right), \quad (35.1)$$

$$\overline{AC}: s \mapsto \left(\frac{s}{R}, 0\right), \quad (35.2)$$

$$\overline{BD}: s \mapsto \left(\frac{s}{R}, \frac{a}{R}\right), \quad (35.3)$$

где размеры квадрата (в действительности прямоугольника) определяются длинами отрезков a и b и величиной R , характеризующей размер сферы. Мы будем рассматривать только малые фигуры, для которых $a, b \ll R$. Параметр s , входящий в уравнения для всех трех кривых, представляет собой длину дуги.

Чтобы завершить построение квадрата, нужно найти четвертую сторону $\overline{CD'}$. Дифференциальные уравнения, описывающие эту кривую, получены в примере в разд. 34:

$$\frac{d\lambda}{ds} = \frac{1}{R^2}, \quad (35.4)$$

$$\frac{d\phi}{ds} = \frac{\cos \lambda_0}{R \cos^2 \lambda}, \quad (35.5)$$

$$\frac{dl}{ds} = -\frac{\cos^2 \lambda_0 \sin \lambda}{\cos^3 \lambda}. \quad (35.6)$$

[Поскольку мы собираемся показать, что никакого геодезического квадрата в полном смысле этого слова не существует, вам было бы полезно вернуться назад и поразмышлять о внутренней непротиворечивости этого шага.]

[Мы проведем этот расчет самым простым способом, что называется «в лоб». Более эффективные методы вычислений столь сложны, что могли бы привести неискушенного читателя в замешательство.]

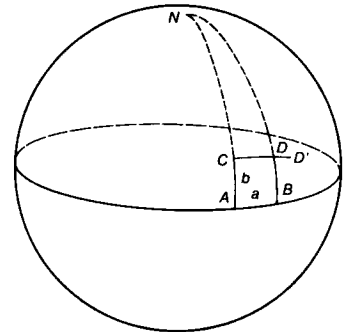


Рис. 35.3
Геодезический квадрат, построенный на поверхности сферы.

Начальные условия имеют вид

$$\begin{aligned}\lambda_0 &= \frac{b}{R}, \\ l_0 &= 0, \\ \phi_0 &= 0.\end{aligned}\tag{35.7}$$

Так как нам нужен лишь короткий отрезок кривой, будем искать решение в виде ряда по степеням параметра s :

$$\lambda = \frac{b}{R} + C_1 s^2, \tag{35.8}$$

$$\phi = C_2 s + C_3 s^2, \tag{35.9}$$

$$l = C_4 s, \tag{35.10}$$

где C_i — коэффициенты, подлежащие определению из предыдущих уравнений. Искомые отклонения углов λ и ϕ должны содержать члены типа ab^2 и a^2b , но не a^3 и b^3 , ибо последние не исчезали бы одновременно даже при сплющивании прямоугольника вдоль любой из сторон до нулевой площади.

Значения коэффициентов C_i получим, выполнив стандартную процедуру подстановки выражений (35.8) — (35.10) в уравнения (35.4) — (35.6) и приравняв члены одинакового порядка малости. Кривую $\overline{CD'}$ можно записать в следующей параметрической форме:

$$\overline{CD'}: s \mapsto \left(\frac{b}{R} - \frac{bs^2}{2R^3}, \frac{s}{R} + \frac{sb^2}{2R^3} \right) \tag{35.11}$$

Отсюда видно, что точка D' имеет координаты

$$(\lambda, \phi)_{D'} = \left(\frac{b}{R} - \frac{ba^2}{2R^3}, \frac{a}{R} + \frac{ab^2}{2R^3} \right). \tag{35.12}$$

Рассчитанный геодезический квадрат изображен на рис. 35.4. Концы D и D' «проскакивают» мимо друг друга на расстояние, пропорциональное кубу размера квадрата, а угловой перекос между этими концами пропорционален его площади.

Описание перекоса

Как описать кривизну пространства-времени математически? Величина перекоса квадрата является линейной функцией длин его сторон. Если задать некоторое направление и поставить ему в соответствие вектор, поворот которого можно было бы рассматривать как меру перекоса, то изменение этого вектора должно быть связано с его длиной тоже линейной за-

[Теперь видно, что мы правильно поступили, выбрав прямоугольник, а не квадрат. С прямоугольником легче работать.]

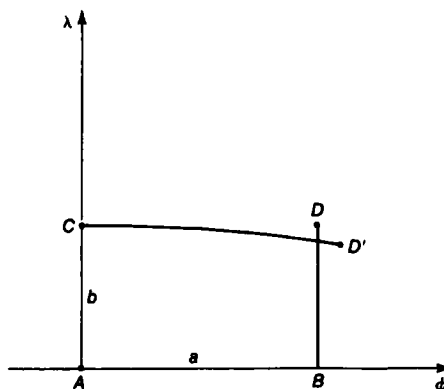


Рис. 35.4

Геодезический квадрат в координатах (λ, ϕ) .

висимостью. Следовательно, должен существовать линейный оператор, тензор, отображающий три вектора на четвертый. Доказательство существования такого оператора и того, что он представляет собой тензор, выходит за пределы наших возможностей, но тем не менее так оно и есть. Обозначим векторы, направленные вдоль сторон квадрата, A^μ и B^μ , а вектор, поворот которого мы собираемся измерять, η^μ . Тогда изменение δ^μ вектора η^μ при его параллельном переносе вдоль всех сторон квадрата будет определяться формулой

$$\delta^\mu = R^\mu_{\nu\sigma\tau} \eta^\nu A^\sigma B^\tau. \quad (35.13) \quad \text{Тензор кривизны}$$

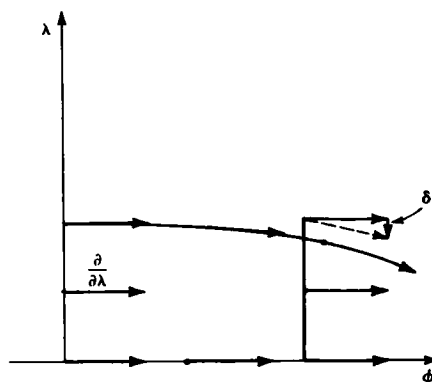


Рис. 35.5

Параллельный перенос касательного вектора вдоль всех сторон геодезического квадрата.

Коэффициенты $R^{\mu}_{\nu\sigma\tau}$ — это компоненты тензора кривизны или, как его называют, тензора Римана.

[Теперь становится понятно, почему мы сделали стороны прямыми, а не закругленными, — так легче сохранить информацию об ориентации.]

Пример 2

Вернемся к нашему геодезическому квадрату на поверхности сферы и проследим за изменением вектора $\partial/\partial\phi$. Этот вектор направлен вдоль верхней и нижней сторон квадрата и под прямыми углами к его боковым сторонам (рис. 35.5). Мерой поворота этого вектора в результате параллельного переноса вдоль всех сторон квадрата следует считать показанный на рисунке вектор разности δ . Угол между исходным вектором в точке D и результирующим вектором в точке D' равен ab/R^2 ; следовательно, вектор разности δ имеет вид

$$\delta = -\frac{ab}{R^2} \frac{\partial}{\partial\lambda}. \quad (35.14)$$

В соответствии с формулой (35.13) можно написать

$$-\frac{ab}{R^2} \frac{\partial}{\partial\lambda} = R^{\mu}_{\phi\phi\lambda} \left(\frac{\partial}{\partial\phi} \right) \left(\frac{a}{R} \frac{\partial}{\partial\phi} \right) \left(\frac{b}{R} \frac{\partial}{\partial\lambda} \right) \frac{\partial}{\partial x^{\mu}}, \quad (35.15)$$

откуда легко найти две компоненты тензора Римана на сфере:

$$\begin{aligned} R^{\phi}_{\phi\phi\lambda} &= 0, \\ R^{\lambda}_{\phi\phi\lambda} &= -1. \end{aligned} \quad (35.16)$$

Если теперь выполнить параллельный перенос вектора $\partial/\partial\lambda$, то получим еще две компоненты:

$$\begin{aligned} R^{\lambda}_{\lambda\phi\lambda} &= 0, \\ R^{\phi}_{\lambda\phi\lambda} &= +1. \end{aligned} \quad (35.17)$$

Перенос векторов при противоположном направлении обхода контура квадрата приводит лишь к изменению знака, т.е.

$$\begin{aligned} R^{\lambda}_{\phi\lambda\phi} &= +1, & R^{\phi}_{\phi\lambda\phi} &= 0, \\ R^{\phi}_{\lambda\lambda\phi} &= -1, & R^{\lambda}_{\lambda\lambda\phi} &= 0. \end{aligned} \quad (35.18)$$

Наконец, обход контура с площадью, равной нулю, должен давать нулевой результат; следовательно,

$$\begin{aligned} R^{\alpha}_{\beta\lambda\lambda} &= 0, \\ R^{\alpha}_{\beta\phi\phi} &= 0. \end{aligned} \quad (35.19)$$

где α и β — любые индексы. Таким образом мы нашли все шестнадцать компонент тензора Римана. Окончательно можно

[Этот способ вычисления компонент тензора кривизны годится только для симметрических пространств размерности симметрии больше двух. Общий случай рассматривается в книгах, рассчитанных на более подготовленного читателя.]

написать

Тензор Римана

$$R = \left(\frac{\partial}{\partial \lambda} \otimes d\phi - \frac{\partial}{\partial \phi} \otimes d\lambda \right) \otimes (d\lambda \otimes d\phi - d\phi \otimes d\lambda). \quad (35.20)$$

Итак, локальная кривизна многообразия описывается тензором Римана — линейным оператором, который при введении в него трех векторов на выходе дает четвертый. Так как, согласно определению, при умножении тензора Римана на длину в третьей степени в результате получается длина в первой степени, он должен иметь размерность $1/(\text{длина})^2$. Из соображений размерности следует ожидать, что вне тела массой M тензор Римана ведет себя как M/R^3 . Согласно другой интерпретации, тензор Римана описывает приливное воздействие гравитационного поля. В самом деле, приливный эффект пропорционален величине M/R^3 .

[Обратите внимание, насколько компактнее это выражение по сравнению с перечислением всех компонент тензора кривизны.]

Тензор Римана описывает локальную кривизну

Пример 3

Идеи, с которыми вы ознакомились в этом разделе, позволяют сделать грубую оценку угла отклонения света гравитационным полем массивного тела, подобного Солнцу. Один из источников неточности наших рассуждений связан с предположением, что луч света можно рассматривать как жесткий стержень. Корректный расчет показывает, что это может быть причиной погрешностей в оценке одного или даже двух коэффициентов, входящих в искомую формулу. К тому же, чтобы использовать выражение, полученное выше для малого прямоугольника, при расчете рассматриваемого в данной задаче (рис. 35.6) большого прямоугольника, нужно предположить, что последний разбит на множество малых, и затем найти результирующий эффект, просуммировав тензоры Римана по всему прямоугольнику. Но все не так просто, ибо здесь предполагается сложение тензо-

Отклонение светового луча



Рис. 35.6

Гигантский геодезический квадрат в космическом пространстве, предназначенный для расчета отклонения луча света гравитационным полем массивного тела. В действительности искривлены оба луча света.

ров, относящихся к разным точкам, а выполнение такой операции требует осторожности.

Если тензор Римана спадает как $1/r^3$, то наибольший вклад в сумму даст область, непосредственно примыкающая к массивному телу. Радиус этой области равен минимальному расстоянию, на котором луч света прошел бы мимо тела, если бы не было поля тяготения. Тогда можно получить следующую оценку угла отклонения θ :

$$\theta \approx R^2 \times \frac{M}{R^3} \approx \frac{M}{R}. \quad (35.21)$$

Разумеется, здесь следовало взять интеграл. Мы избежали этого, но в результате потеряли некоторый числовой множитель. Для Солнца наш упрощенный расчет приводит к углу отклонения света

$$\theta \approx \frac{M}{R} \approx \frac{5 \cdot 10^{-6}}{2} \approx 2 \cdot 10^{-6} \text{ радиан} \approx \frac{1}{2} \text{ дуг. сек.} \quad (35.22)$$

Аккуратные вычисления дают следующий точный результат:

$$\theta = \frac{4M}{R}. \quad (35.23)$$

Уравнения Эйнштейна

Важное значение тензора Римана заключается в том, что он является динамической величиной в уравнениях Эйнштейна. Эти уравнения описывают динамику пространства-времени почти так же, как уравнения Максвелла описывают динамику электромагнитного поля. Уравнения Эйнштейна имеют вид

$$R^\mu_{\cdot\alpha\nu\beta} g^{\alpha\beta} - \frac{1}{2} R^\gamma_{\cdot\alpha\delta\beta} g^{\alpha\beta} \delta^\delta_\gamma \delta^\mu_\nu = 8\pi T^\mu_\nu, \quad (35.24)$$

где $R^\mu_{\cdot\alpha\nu\beta}$ — тензор Римана, $g^{\alpha\beta}$ — тензор, обратный метрическому, δ^α_β — символ Кронекера, а T^μ_ν — тензор, описывающий плотность энергии, плотность импульса и плотность натяжений. Тензор T^μ_ν для краткости называют тензором энергии-импульса. Он играет роль источника гравитационного поля точно так же, как плотности заряда и тока играют роль источников электромагнитного поля. Методы обращения с этими уравнениями и их решение выходят далеко за пределы наших возможностей, но мы привели их здесь, чтобы вы хотя бы знали, как они выглядят.

ЗАДАЧИ

35.1. (14) Докажите с помощью геодезического квадрата, что кривизна (как мы ее определили в этом разделе) поверхности цилиндра равна нулю.

35.2. (16) Сделайте то же самое, что и в задаче 35.1, для конуса.

35.3. (25) Что можно сказать по поводу вершины конуса в задаче 35.2?

35.4. (12) Верно ли утверждение, что наибольшее значение тензора Римана для Солнечной системы достигается на Земле?

35.5. (33) Постройте на сфере геодезический квадрат с вершиной на Северном полюсе.

35.6. (36) Постройте геодезический квадрат на псевдосфере с метрикой

$$\mathcal{G} = R^2(dx^2 + \text{sh}^2 \chi d\phi^2).$$

Докажите, что кривизны сферы и псевдосферы имеют противоположные знаки.

35.7. (28) Постройте в пространстве с метрикой

$$\mathcal{G} = \frac{dx^2 + dy^2}{y^2}$$

(см. задачи 34.4 и 34.5) правильный пятиугольник с прямым углом при каждой вершине.

35.8. (29) Найдите компоненты тензора Римана на сфере. Постарайтесь решать задачу не «в лоб», а опираясь на симметрию.

35.9. (36) Распространите идею геодезического квадрата на случай, когда две его стороны — времениподобные кривые. Рассмотрите такой квадрат на поверхности Земли; пусть его стороны имеют радиально-временные направления. Времениподобные стороны этого квадрата будут мировыми линиями падающих частиц. Обнаруживается ли здесь кривизна пространства-времени? Какие компоненты тензора Римана можно вычислить с помощью данного геодезического квадрата и чему равны их значения?

35.10. (13) Почему в примере 1 на стр. 294 мы оставили только линейные члены разложения / в степенной ряд?

IV. Космология

[Прекрасное описание этого захватывающего времени дано в книге [18].]

Теория гравитации, развитая Эйнштейном в 1906 — 1915 гг., устранила противоречие между ньютоновой теорией тяготения и специальной теорией относительности. Описание гравитационных эффектов в теории Эйнштейна осуществляется в пространстве-времени, обладающем непостоянной метрикой, — искривленном пространстве-времени. Как и следовало ожидать, эта теория хорошо согласуется со специальной теорией относительности, поскольку одна из целей, которую ставил перед собой Эйнштейн в своих исследованиях, как раз и состояла в достижении такой согласованности. Вместе с тем теория Эйнштейна обладает еще двумя свойствами. Первое из них заключается в том, что мы получаем возможность вполне определенным и естественным образом описывать бесконечную Вселенную. Второе, гораздо более неожиданное свойство теории связано с тем фактом, что такая Вселенная может быть конечной по своим размерам и в то же время не иметь края или границы. Указанные теоретические представления являются одними из самых дерзновенных, волнующих и смелых идей об окружающем нас мире, высказывавшихся в точной физической науке. Это очень тонкие идеи, и, чтобы по-настоящему постичь их, нам потребуется почти все, что мы узнали в первых трех главах книги. Итак, за дело!

36. Из чего состоит Вселенная

Займемся теперь проблемой построения модели Вселенной на основе общей теории относительности. Для такого построения нам придется огромное число раз производить крупномасштабное усреднение. К примеру, рассмотрение «звезд как пылинок» в нашем случае слишком детально. Всю галактику в целом мы будем рассматривать в нашей модели как одну частицу, которой соответствует определенная мировая линия. Более того, мы будем приближенно считать, что все эти частицы, вместе

взятые, образуют непрерывную жидкость. Прежде чем приступить к изучению настоящей главы, желательно прочитать ту или иную хорошую популярную книгу по астрономии и космологии¹⁾.

Было бы излишним повторять здесь то, что прекрасно изложено в других книгах. Чтобы знать, на что именно обращать внимание при таком чтении, я привожу обзор важнейших составных частей Вселенной, которые должны найти свое отражение в нашей модели. При выборе значений свободных параметров, фигурирующих в модели, желательно использовать в качестве исходных данных лишь небольшое число известных фактов. С помощью такой модели можно было бы объяснить те известные факты, которые не были нами использованы, и тем самым получить подтверждение правильности выбранных теоретических построений. В хорошей модели отношение числа выходных к числу входных данных должно быть достаточно велико. Однако — увы! — наши знания о Вселенной столь скудны и несовершенны, что, даже используя их все, мы можем ответить лишь на очень ограниченное количество вопросов. Хорошо еще, что наблюдаемые свойства Вселенной по крайней мере не противоречат в корне нашим моделям. Самое большее, что мы можем сделать, — это очертить тот круг идей, в рамках которого вопросы типа «Какова конечная судьба нашего мира?» можно ставить и обсуждать с точки зрения физики, а не философии или религии.

Одно из важнейших космологических наблюдений абсолютно достоверно: ночью мы видим темное небо. Однако фоновая яркость точно не известна, поскольку при космологических наблюдениях мы вынуждены учитывать искажения, вносимые как земной атмосферой, так и нашей собственной Галактикой. Тем не менее с полной определенностью можно утверждать, что межзвездное пространство намного темнее, чем поверхность Солнца. Даже этот факт служит источником трудности при построении модели Вселенной.

[Полезен также сборник статей [12] из журнала Scientific American.]

Модели Вселенной

Темное ночное небо

¹⁾ Советскому читателю можно порекомендовать следующие популярные книги: *Агекян Т. А.* Звезды, галактики, метagalaktika. — М.: Наука, 1982; *Вайнберг С.* Первые три минуты. — М.: Энергоиздат, 1981; *Новиков И. Д.* Эволюция Вселенной. — М.: Наука, 1983; *Шкловский И. С.* Вселенная, жизнь, разум. — М.: Наука, 1980; *Шкловский И. С.* Звезды: их рождение, жизнь и смерть. — М.: Наука, 1977. — Прим. ред.

До переворота в научных взглядах, связанного с именем Коперника, темное ночное небо не вызывало недоумения. Ночью оно темное как раз потому, что вне нас ничего нет. У древних астрономов не было оснований полагать, что Вселенная не ограничена и что мы не находимся в ее центре. В настоящее время такие взгляды непопулярны. Но если мы не занимаем выделенного положения во Вселенной, то модель Вселенной, основанная на евклидовой геометрии, требует бесконечности нашего мира. Представление о подобной неограниченной Вселенной немедленно ставит на повестку дня ряд проблем. Так, например, с помощью ньютоновой теории тяготения оказывается невозможным описать бесконечное количество вещества. Хуже того, бесконечное число звезд в такой Вселенной заливало бы нас светом со всех направлений. Ночное небо было бы ярким, как поверхность Солнца. Модель бесконечной статической евклидовой Вселенной не позволяет найти выход из этого тупика, известного под названием *парадокса Ольбера*.

Звезды

Из более мелких объектов во Вселенной, которые представляют интерес для космологов, следует упомянуть звезды, планетные системы (или во всяком случае одну из них) и шаровые скопления звезд. Последние содержат около 10^5 звезд, которые динамически связаны друг с другом посредством взаимных сил притяжения и предположительно имеют общее происхождение. Для каждого из этих объектов мы можем оценить возраст, который зависит от наших представлений о физических процессах и начальных условиях. Если исключить из рассмотрения экзотические начальные условия, то можно утверждать, что собственное время существования Вселенной больше возраста каждого из упомянутых объектов. В Солнечной системе мы можем установить возраст Земли, Луны и метеоритов; при этом получается величина порядка $1,6 \cdot 10^{17}$ с (5 миллиардов лет). Детальные расчеты эволюции Солнца на ЭВМ дают основание оценить его возраст примерно той же величиной. Однако эта оценка содержит неопределенности, поскольку мы не знаем состава внутренних областей Солнца. Если, с другой стороны, мы наблюдаем звезды, входящие в некоторое шаровое скопление, то они в общем случае имеют различные массы, но сходный состав, обусловленный общим происхождением. Вид диаграммы Герцшпрунга — Рессела для рассматриваемых скоплений проливает свет на неизвестный химический состав. Указанное обстоятельство позволяет более точно определить возраст звезд в шаровом скоплении. В частности, такие скопления содержат очень старые, возможно самые старые звезды во

Звездные скопления

[Если вы не знаете, что такое диаграмма Герцшпрунга — Рессела, то вы недостаточно прочитали дополнительную литературы.]

Вселенной. Возраст таких шаровых скоплений оценивается величиной $(2 \div 4) \cdot 10^{17}$ с.

Наблюдения за перечисленными выше близлежащими объектами, с давних пор проводимые астрономами, дают ценные сведения. Многие космологические модели позволяют делать предсказания относительно распространенности различных химических элементов, а для большинства близких объектов эту распространенность можно изучать. В частности, происхождение Вселенной в результате Большого Взрыва неизбежно приводит к превращению значительного количества водорода в гелий с небольшой добавкой дейтерия. Относительное количество этих элементов зависит от детальной структуры ранней Вселенной на той стадии ее эволюции, когда характерное время изменения космологических параметров имело порядок времени жизни нейтрона. С указанной точки зрения распространенность дейтерия представляет особый интерес, поскольку его чрезвычайно трудно синтезировать каким-либо иным образом вне космологических условий, реализуемых в ранней Вселенной. Так, например, если дейтерий синтезируется во внутренних слоях звезды, то он выгорит при ядерных процессах, так и не достигнув ее поверхности¹⁾.

Из других объектов, представляющих космологический интерес, отметим галактики и скопления галактик. Галактики — это элементарные частицы в космологии. Как правило, из наблюдения галактики можно установить ее общее красное смещение, видимые светимость и угловой размер, а также размеры окружающих ее областей Н II. Можно изучать статистические свойства распределения галактик, степень их скученности, однородность распределения, возможную анизотропию. Исключительно важное значение для космологии имеет тот установленный наблюдениями факт, что спектр излучения более мелких и тусклых галактик (вероятно, находящихся на большем расстоянии) обнаруживает систематический сдвиг частот в красную сторону, как будто галактики от нас удаляются. Как станет ясно из дальнейшего, утверждение о том, что галактики движутся по направлению к нам (равно как и противоположное утверждение) бессодержательно. Точнее, оно безвредно, но не содержательно в смысле операциональном. Для удобства

Химические элементы

Гелий

Галактики

Красное смещение

¹⁾ Не исключено, однако, что дейтерий, синтезируемый во внутренних слоях звезд, может быть выброшен в межзвездное пространство при катастрофических процессах типа взрывов Сверхновых. — *Прим. ред.*

обсуждения принято считать — и мы сами будем придерживаться такой договоренности, — что удаленные галактики движутся по направлению от нас. Указанное систематическое «движение» галактик называется хаббловским разбеганием. Насколько можно судить, хаббловское разбегание однородно и изотропно, без каких-либо значительных систематических отклонений.

Радиоисточники

Вселенная гораздо прозрачнее для радиоволн, чем для света. Поэтому, используя радиоволны, мы можем проводить наблюдения вплоть до очень больших расстояний. На небе имеются источники радишума, которые, по-видимому, удалены от нас на поистине огромные расстояния. Будь эти объекты лучше изучены, они могли бы сыграть роль полезных космологических индикаторов. Например, если бы все они обладали одинаковой и неизменной во времени радиосветимостью, то подсчет числа таких источников как функции их светимости дал бы важную информацию о геометрии Вселенной. К сожалению, радиоисточники не образуют никакой однородной популяции, а светимость каждого из них не остается постоянной на протяжении времени его жизни. Мы наблюдаем также двойные радиоисточники. Если бы удалось каким-либо образом установить свойства таких пар, то их можно было бы использовать как эталоны длины, а их видимые угловые размеры опять же послужили бы источником информации о глобальной геометрии нашей Вселенной.

Кварзары

Наконец, среди мощных радиоисточников мы находим квазары, о которых уже так много написано. Эти объекты видны также и в оптическом диапазоне, причем характеризуются весьма большими красными смещениями. Если красные смещения действительно обусловлены хаббловским разбеганием, то квазары находятся от нас на гигантских расстояниях. Если бы мы знали, что из себя представляют квазары, то смогли бы извлечь из них большую пользу для космологической науки.

Микроволновое фоновое излучение

Последним «объектом» во Вселенной, представляющим интерес с космологической точки зрения, является универсальный фон чернотельного микроволнового излучения, которое приходит к нам со всех направлений и всегда (с точностью лучше 0,25%) имеет температуру около 3 К. Угловое распределение этого излучения характеризуется очень высокой степенью изотропии вплоть до масштабов порядка миллирадиан. Единственная жизнеспособная физическая интерпретация такого излучения заключается в том, что оно представляет собой реликт от некоторого сжатого и, следовательно, горячего состояния Все-

ленной. Среди прочих космологических свойств микроволнового фонового излучения особо следует отметить то обстоятельство, что оно выделяет абсолютную систему отсчета. Существует лишь одна скорость, при которой распределение температуры излучения кажется нам изотропным. Для прочих состояний движения излучение приобретает красное смещение и соответственно кажется холоднее в одном направлении, тогда как в противоположном направлении, наоборот, оно испытывает голубое смещение, т.е. кажется горячее. Современный уровень развития техники позволяет измерить пекулярную скорость Солнца относительно микроволнового фонового излучения. Число микроволновых фотонов, из которых состоит это излучение, значительно превосходит число прочих частиц во Вселенной¹⁾. Таким образом, большая часть энтропии Вселенной, связанная с фазами и координатами реликтовых фотонов, находится вне поля нашего зрения.

Можно попытаться непосредственно подсчитать массу вещества во Вселенной, содержащегося в известных нам объектах. Однако в каких бы масштабах ни производился подобный подсчет, всегда обнаруживается нехватка массы. Как сами галактики, так и скопления галактик в динамическом отношении ведут себя таким образом, будто они содержат больше вещества, чем имеется в их видимых компонентах. В принятой физической картине мира остается очень узкая лазейка для такой «скрытой массы». Похоже, что весьма кстати здесь оказывается тот дух академической сдержанности, который столь часто тормозит развитие астрофизики. Ведь мы, вероятно, и не знаем до сих пор всех физических возможностей.

Скрытая масса

Теперь мы готовы приступить к построению математической модели, отражающей перечисленные выше наблюдательные факты и совместимой с общей теорией относительности. На основе нашей модели мы сможем ответить на волнующие всех вопросы о том, каково было начало и какова конечная

¹⁾Имеется в виду, что средняя плотность числа реликтовых фотонов на девять порядков превосходит общую среднюю плотность других частиц (протонов и нейтронов, свободных и связанных в ядрах), входящих во все видимые формы материи. Если же в соответствии с обсуждаемой в книге моделью Большого Взрыва на достаточно ранней стадии космологического расширения действительно имело место почти полное термодинамическое равновесие и во Вселенной существуют такие нерегистрируемые в настоящее время реликтовые частицы, как нейтрино и гравитоны, то число последних должно быть того же порядка, что и число реликтовых фотонов. — *Прим. ред.*

судьба нашего мира. Ключевую роль при построении модели играют хаббловское разбегание и темное ночное небо. Мы заранее предполагаем, что Вселенная не является неизменной во времени, а эволюционирует, — это знаменует собой замечательный прогресс по сравнению с представлениями древних мыслителей. Нет сомнений в том, что в будущем нас ожидают столь же значительные успехи.

37. Модели пространства-времени Робертсона — Уокера

Невыделенность нашего положения во Вселенной

В наше время вряд ли кто-нибудь полагает, что Вселенная создана специально для нас и что мы находимся в ее центре. Следуя революционной идее Коперника, мы считаем, что занимаем некоторое типичное, а не единственно выделенное положение в мире. В настоящее время астрономическая книга с названием «В центре необъятности» вызвала бы улыбку. Конечно, в достаточно малом масштабе мы находимся в выделенном положении. К примеру, ни у кого не вызывает сомнений тот факт, что мы располагаемся именно на поверхности Земли. Расположение нашей планеты в Галактике также не типично: она значительно ближе к звезде, чем случайно выбранная точка. Однако где же еще можно было бы найти необходимые для жизни потоки энергии? Местоположение Солнца в Галактике менее выделено, но и здесь есть ограничение — ведь звездообразование происходит только в определенных местах. Кроме того, звезды должны содержать тяжелые элементы, необходимые для образования планет, т.е. должны представлять собой звезды второго поколения. Эта книга не могла быть написана на самой ранней стадии эволюции Вселенной: должно пройти достаточно много времени, чтобы звезды успели приобрести планеты, но не слишком много, чтобы они все еще продолжали светить. В упомянутых широких пределах, однако, наше положение в пространстве и времени ничем не выделено. Наша Галактика — самая обыкновенная. Она не является ни изолированной галактикой, ни членом одного из больших скоплений. Скопления тысяч галактик должны быть наиболее заметными ориентирами во Вселенной. Похоже, что на масштабе таких скоплений галактик Вселенная не имеет особенностей¹⁾, и это

¹⁾За время, прошедшее после выхода в свет американского издания этой книги, наблюдатели с определенностью установили, а теоретики объяснили существование во Вселенной сетчатой («сотовой») структуры как раз на масштабе скоплений галактик (см. обзор Шандарин

неудивительно. В эволюционирующей Вселенной с конечным временем жизни структуры столь огромных размеров могли не успеть образоваться. К сожалению, численные оценки показывают, что спонтанным путем к настоящему времени не успели бы образоваться даже такие сравнительно небольшие объекты, как галактики¹⁾. Мы остаемся в недоумении, откуда же они взялись и почему не возникли также структуры больших размеров.

Наша модель Вселенной имеет дело только с такими наибольшими масштабами. В ней нет места ни отдельным галактикам, ни даже крупным скоплениям. Мы рассматриваем лишь некоторую космическую жидкость, частицы которой суть галактики. На этом уровне крупномасштабного усреднения Вселенная особенностей не имеет. Перефразируя известный детектив, можно сказать, что имеется лишь одна характерная особенность, заключающаяся в отсутствии характерных особенностей. Такая крупномасштабно усредненная модель должна описывать усредненное поведение Вселенной. Произведя указанное упрощение, легко получить модели Вселенной. Такие модели называются моделями пространства-времени Робертсона — Уокера. Предположение о сферической симметрии вносит огромное упрощение. Без него вряд ли удалось бы достигнуть заметного прогресса, и в настоящее время почти все космологические исследования проводятся в рамках моделей пространства-времени Робертсона — Уокера. Насколько мы можем судить, в сверхбольших масштабах структура Вселенной действительно симметрична и не имеет особенностей. Вместе с тем в этом вопросе необходима некоторая осторожность. Локальная симметрия, наблюдаемая в современной Вселенной, отнюдь не

Сглаженная картина Вселенной

Симметрия

С. Ф., Дорошкевич А. Г., Зельдович Я. Б. — УФН, 1983, т.139, вып.1., с.83 — 134). Подчеркнем, однако, что наличие такой структуры отнюдь не противоречит описанному в тексте фундаментальному космологическому принципу, гласящему, что существует некоторый достаточно большой масштаб, вне которого Вселенная особенностей не имеет. Последний принцип получает лишь количественное уточнение. — *Прим. перев.*

¹⁾Проблема начальных возмущений, сформулированная в 1946 г. советским ученым Е. М. Лифшицем, за последние 5 лет сдвинулась, на конец, с мертвой точки. Так, выяснилось, что приемлемое решение упомянутой проблемы может быть достигнуто на основе сценария так называемой инфляционной Вселенной. Суть этого сценария заключается в экспоненциальном «раздувании» первоначально деситтеровской Вселенной с последующим переходом ее в плазменно-материальную (фридмановскую) стадию. — *Прим. ред.*

[Подробнее об однородности и изотропии реальной Вселенной будет рассказано в разд. 50.]

Однородность и изотропия

[В примере с электропроводностью в разд. 21 рассматривался однородный, но не изотропный материал.]

Сфера

означает симметрии в более ранние моменты времени. Точно так же симметричный огненный шар ядерного взрыва не означает симметрии конструкции ядерной бомбы.

Чем характеризуется максимально возможная степень симметрии пространства? Во-первых, одна точка должна быть неотличима от другой. Во-вторых, одно направление должно быть неотличимо от другого. Первое из этих свойств называется *однородностью*, второе — *изотропией*. Такой максимально возможной степенью симметрии обладает евклидова плоскость. Другим примером является поверхность сферы.

Пример 1

Чтобы доказать, что поверхность сферы однородна, найдем преобразования, которые оставляют метрику неизменной, но переводят одну произвольную точку в другую. Такие преобразования, сохраняющие метрику, называются *изометриями*. Одно из них — поворот вокруг Северного полюса, другое — поворот вокруг точки на экваторе. Комбинируя эти преобразования, можно перевести одну произвольную точку в другую. Таким образом, поверхность сферы представляет собой однородное пространство.

Коль скоро однородность уже доказана, т.е. все точки на поверхности сферы эквивалентны, достаточно показать изотропию в какой-либо одной точке. Посмотрим с этой целью на Северный полюс. Повороты вокруг Северного полюса являются изометриями и могут преобразовывать одно произвольно заданное направление в другое. Таким образом, поверхность сферы является также изотропной.

Полуплоскость

Пример 2

Что можно сказать относительно множества точек в верхней полуплоскости, $y > 0$, с метрикой

$$g = \frac{dx^2 + dy^2}{y^2} \quad (37.1)$$

Одним из примеров изометрии на этом многообразии может служить трансляция T_d на конечное расстояние d :

$$T_d: (x, y) \mapsto (x + d, y). \quad (37.2)$$

Ввиду явной зависимости от y трансляция вдоль y не обладает свойствами изометрии. Может показаться, что других изомет-

рий нет, но тем не менее они существуют. Рассмотрим «растяжения» E_α :

$$E_\alpha: (x, y) \mapsto (\alpha x, \alpha y). \quad (37.3)$$

Такие «растяжения» также оказываются изометриями. Чтобы обосновать это утверждение, покажем, что в новых координатах метрика $\mathcal{G}$ имеет такой же вид, как и в старых. Из координатных преобразований

$$\begin{aligned} x' &= \alpha x, \\ y' &= \alpha y \end{aligned} \quad (37.4)$$

получим

$$\begin{aligned} dx' &\mapsto \alpha dx, \\ dy' &\mapsto \alpha dy, \end{aligned} \quad (37.5)$$

вследствие чего

$$\mathcal{G} \mapsto \frac{dx'^2 + dy'^2}{(y')^2}. \quad (37.6)$$

Метрика в новых, штрихованных координатах имеет тот же функциональный вид, что и в старых, поэтому рассматриваемое преобразование представляет собой изометрию.

Действие этого преобразования показано на рис. 37.1. Кривая AB , являющаяся геодезической (задача 34.5), переходит в кривую $A'B'$. Поскольку метрика при таком преобразовании остается неизменной, длины упомянутых кривых должны совпадать.

Существуют всего три различных однородных и изотропных n -мерных многообразия. Одно из них — евклидово n -пространство, n -мерный вариант плоскости, другое — n -мерный вариант сферы и третье — пространство отрицательной кривизны, именуемое псевдосферой. Перечисленные три типа симметричных пространств отвечают соответственно евклидовой геометрии и двум неевклидовым геометриям, открытым Гауссом, Лобачевским и Бойаи. Одно из представлений двумерного варианта псевдосферы описано в приведенном выше примере 2. В указанном представлении изотропия такого пространства не очевидна, однако можно перейти к другим координатам, в которых изотропия выступает явно. Несколько следующих разделов мы посвятим изучению геометрии 3-сферы и 3-псевдосферы. Кстати сказать, сейчас как раз настало время привести метрику, описывающую каждое из этих пространств. При этом мы будем использовать такие координаты, которые

[Если необходимо, освежите в памяти рассмотрение этого вопроса в разд. 20.]

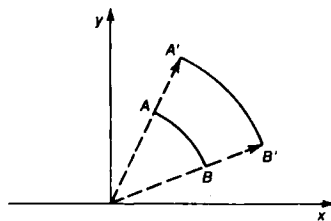


Рис. 37.1
Растяжение относительно начала координат. Оно представляет собой изометрию полуплоскости Пуанкаре.

Симметричные пространства

[Мы не собираемся здесь доказывать, что существуют только три таких симметричных пространства. Неформальное доказательство приведено в гл. 10 книги [25].]

Евклидовы пространства

максимально схожи со сферическими полярными координатами. Пространства $\mathcal{S}^3$ с евклидовой геометрией могут быть представлены следующим метрическим тензором:

$$\mathcal{G} = R^2[d\chi^2 + \chi^2(d\theta^2 + \sin^2 \theta d\phi^2)]. \quad (37.7)$$

Сферы

Мы обозначили безразмерную радиальную координату буквой χ , чтобы отличать ее от обычной радиальной координаты. Тензор

$$d\Omega^2 \equiv d\theta^2 + \sin^2 \theta d\phi^2 \quad (37.8)$$

описывает метрику на 2-сфере, т.е. на поверхности обычной сферы. В силу изотропии все наши симметричные пространства обладают сечениями, которые выглядят как 2-сферы.

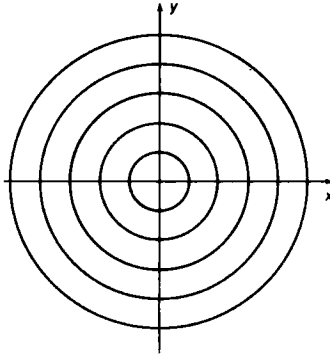


Рис. 37.2
Евклидова плоскость как объединение окружностей.

Сопутствующие координаты

Пример 3

Сечения как евклидовой плоскости, так и самой 2-сферы представляют собой окружности, т.е. 1-сферы. Эти окружности показаны на рис. 37.2 и 37.3.

Галактики движутся вдоль линий $(\chi, \theta, \phi) = \text{const}$. Сопутствующими называются координаты в такой системе отсчета, которая движется вместе с частицами вдоль линий тока. Во всех наших симметричных пространствах сечения $\chi = \text{const}$ представляют собой 2-сферы, размер которых зависит от χ , а метрика пропорциональна $d\Omega^2$.

Пространство, именуемое 3-сферой, обладает метрикой

$$\mathcal{G} = R^2(d\chi^2 + \sin^2 \chi d\Omega^2), \quad (37.9)$$

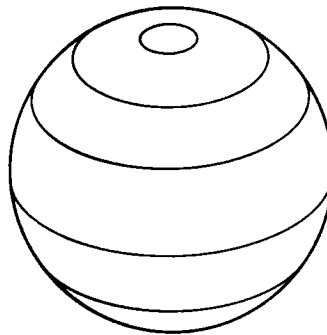


Рис. 37.3
2-сфера как объединение окружностей.

где координата χ меняется на сей раз лишь в пределах от 0, что соответствует одному полюсу, до π , что соответствует другому полюсу, а R — постоянная, имеющая размерность длины и называемая радиусом 3-сферы. Радиусы сечений, представляющих собой 2-сферы, выражаются теперь величиной $R \sin \chi$. Наибольшая 2-сфера в рассматриваемом пространстве имеет радиус R .

Пример 4

В случае когда 2-сфера разбита на 1-сферы, радиусы последних задаются величиной $R \sin \theta$, где R — радиус 2-сферы, а θ — обычная сферическая координата. Как и в рассмотренном выше случае, радиус 1-сферы имеет максимальную величину для точек на экваторе. Таким образом, 3-сфера содержит экваториальную 2-сферу, задаваемую в наших координатах соотношением $\chi = \pi/2$.

Наконец, 3-псевдосфера имеет метрику

$$\mathcal{G} = R^2 (d\chi^2 + \text{sh}^2 \chi d\Omega^2), \quad (37.10) \quad \text{Псевдосферы}$$

где χ изменяется от нуля до бесконечности. Чтобы избежать постоянных повторов в написании трех почти одинаковых соотношений, введем функцию $S(\chi)$, равную соответственно χ , $\sin \chi$ или $\text{sh} \chi$ для евклидова 3-пространства, 3-сферы и 3-псевдосферы. В случаях как 3-сферы, так и 3-псевдосферы координата χ имеет простой геометрический смысл. Например, расстояние «вокруг» 3-сферы, измеренное по координате χ , в точности равно 2π . В евклидовом же случае не существует никакого внутреннего определения χ , в расчеты входит лишь произведение $(R\chi)$. Этот случай характеризуется более высокой степенью симметрии, что находит свое отражение в геометрии: в евклидовом пространстве для каждой фигуры можно построить ей подобную, чего нельзя сделать ни в одном из двух других симметричных пространств.

Пример 5

Упомянутое выше отсутствие подобных фигур легко продемонстрировать на примере 2-сферы. На рис. 37.4 изображен сферический треугольник, стороны которого образованы большими окружностями, пересекающимися под прямыми углами в каждой из трех вершин. Треугольник, подобный данному, но вдвое меньшего размера, построить невозможно. При попытке сделать это углы получатся меньше прямых.

Подобные фигуры

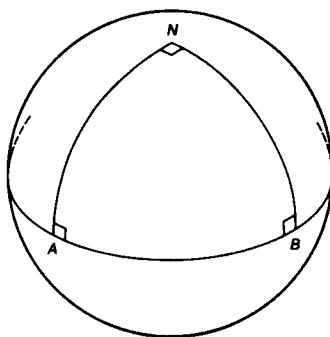


Рис. 37.4
Истинно прямоугольный треугольник.

Динамические модели

Мы будем применять модели Вселенной, каждая из которых содержит пространственные сечения, представляющие собой симметричные 3-пространства. Размеры таких пространственных сечений могут зависеть от времени, а поскольку мы собираемся описывать эволюционирующую Вселенную, они действительно будут зависеть от времени. Если теперь предпринять попытку распространить условие однородности пространства на все пространство-время, то мы получим космологические модели, в которых все моменты времени эквивалентны между собой. Такие модели не допускают эволюции, и ни одна из них не совместима с общей теорией относительности. Были предложены модели с такой симметрией, построенные на основе других теорий гравитации. Они называются моделями стационарной Вселенной. К сожалению, такие модели противоречат общей теории относительности, хотя само по себе это обстоятельство и не является для них роковым¹⁾. У нас все еще нет твердой уверенности в том, что общая теория относительности правильно описывает гравитационное поле столь больших объектов, как наша Вселенная. Тем не менее упомянутые стационарные модели не в состоянии объяснить эволюционных проявлений во Вселенной, и поэтому в настоящее время они не в ходу.

Космическое время

Вводя в рассмотрение эволюционирующую Вселенную, мы тем самым специальным образом выделяем временную координату. При этом любое изменяющееся свойство Вселенной играет роль «часов». Так, например, в качестве универсальной

¹⁾ Отметим, что исторически первая космологическая модель — статическая Вселенная Эйнштейна, построенная на основе уравнений общей теории относительности, дополненных космологическим членом, — была моделью такого типа. — *Прим. ред.*

временной координаты можно использовать температуру микроволнового фонового излучения. Такое время не измеряется часами в духе специальной теории относительности. Эти космические «часы» нелокальны. Регистрируемое в настоящее время микроволновое излучение приходит к нам из далекого прошлого и с огромных расстояний. Применение таких «часов» для определения универсального времени никоим образом не противоречит специальной теории относительности. Вместе с тем универсальное время весьма напоминает ньютоново абсолютное время.

Наиболее общее пространство-время, которое эволюционирует и вместе с тем обладает симметричными пространственными сечениями, имеет следующую метрику:

$$\mathcal{G} = -dt^2 + R^2(t) [d\chi^2 + S^2(\chi) d\Omega^2]. \quad (37.11)$$

Временной параметр t отвечает собственному времени вдоль мировой линии, где χ , θ и ϕ постоянны. Такие линии задают опорную космическую систему отсчета. Как показали недавние измерения температуры микроволнового фонового излучения, мировая линия нашей Галактики является примерно такой мировой линией. Скорость нашей Галактики относительно упомянутой опорной космической системы отсчета составляет величину меньше 10^{-3} от скорости света ¹⁾.

Функция $R(t)$ все еще подлежит определению. Представленная выше метрика содержит кинематические ограничения на однородность и изотропию. Динамика же должна задаваться теорией тяготения, скажем, общей теорией относительности. Все модели пространства-времени приведенного выше вида называются моделями пространства-времени Робертсона — Уокера. Прежде чем вводить в рассмотрение динамику, конкретизируемую общей теорией относительности, остановимся более подробно на изучении удивительной геометрии 3-сферы и 3-псевдосферы.

Модели пространства-времени
Робертсона — Уокера

¹⁾ Приведенная оценка нуждается в уточнении. В 1977 г. американские радиоастрономы Дж. Ф. Смут, М. В. Горенштейн и Р. А. Мюллер установили, что скорость движения Солнечной системы относительно введенной в тексте опорной космической системы отсчета составляет (390 ± 60) км/с. Используя данные о скорости вращения Галактики, удастся оценить скорость движения центра Галактики относительно микроволнового фонового излучения величиной порядка 600 км/с, т.е. $2 \cdot 10^{-3}$ от скорости света. — *Прим. ред.*

ЗАДАЧИ

37.1. (08) Во всех направлениях в нашей Вселенной вдалеке от нас наблюдается большая плотность квазаров, чем вблизи. Не ставит ли нас это обстоятельство в выделенное положение во Вселенной?

37.2. (12) Придумайте пространство-время, которое однородно, но не изотропно.

37.3. (18) Изотропно ли пространство-время с метрикой

$$\mathcal{G} = y^2(dx^2 + dy^2)?$$

37.4. (15) Почему в настоящем разделе не рассмотрено пространство-время с двумя неизвестными функциями:

$$\mathcal{G} = -A^2(t) dt^2 + R^2(t) [dx^2 + S^2(x) d\Omega^2]?$$

37.5. (24) Получите метрику полуплоскости Пуанкаре

$$\mathcal{G} = \frac{dx^2 + dy^2}{y^2}$$

в новых координатах (u, v) , определенных соотношениями

$$z \equiv x + iy,$$

$$w \equiv u + iv,$$

$$w = \frac{2}{z + i} + i.$$

Все отображения, задаваемые комплексными переменными, обладают свойством конформности, что показано в любой книге по теории функций комплексного переменного. Доказательство несложно. Используйте эту конформность, чтобы упростить расчет.

37.6. (19) Используя (u, v) -представление задачи 37.5, покажите, что полуплоскость Пуанкаре изотропна.

37.7. (31) Покажите, что модель пространства-времени Минковского может быть записана в виде модели пространства-времени Робертсона — Уокера с $S(\chi) = \chi$ или $S(\chi) = \sinh \chi$.

38. Глобальная структура 3-сферы

За исключением тривиальных случаев, разобраться в структуре 3-пространств совсем не просто. В то время как можно научиться зрительно представлять себе 2-сферы в достаточно малом числе измерений (по крайней мере в четырех или пяти), представить себе трехмерные объекты даже в случае евклидова 3-пространства весьма сложно. Архитекторы тратят годы, чтобы развить в себе это умение. Только обладая хорошо развитым воображением, можно зрительно представить себе 3-сферу. Чтобы обойти трудности, связанные с ограниченностью нашего воображения, мы должны придумать такие представления, которые заслуживали бы полного доверия. С целью приобретения соответствующих навыков остановимся сначала на нескольких представлениях для 2-сферы, а затем перейдем к 3-сфере.

Наиболее естественное представление 3-сферы, которую мы будем для краткости именовать S^3 , использует концепцию фиктивного евклидова 4-пространства. Пусть (x, y, z, w) — прямоугольные координаты в этом пространстве. Множество точек, расположенных на единичном расстоянии от начала координат,

$$x^2 + y^2 + z^2 + w^2 = 1, \quad (38.1)$$

имеет глобальную структуру 3-многообразия S^3 . Это естественное обобщение 2-сферы — множества точек, расположенных на единичном расстоянии от начала координат в евклидовом 3-пространстве. К этому утверждению следует относиться внимательно, но не понимать его слишком буквально! Многообразие S^3 не тождественно введенному выше множеству точек. Утверждается лишь, что S^3 обладает той же самой глобальной структурой, что и рассматриваемое множество точек.

Можно ввести такие координаты на S^3 (а фактически на всех S^n), которые аналогичны обычным сферическим полярным координатам. Однако прежде чем делать это, проследим внимательно соответствующую процедуру для случая S^2 . (Обычно сначала рассматривают аналоги, имеющие меньшее число измерений.)

Обычные сферические полярные координаты на S^2 определяются посредством соотношений

$$\begin{aligned} x &= \sin \theta \cos \phi, \\ y &= \sin \theta \sin \phi, \\ z &= \cos \theta. \end{aligned} \quad (38.2) \quad \text{2-сфера}$$

Проблемы зрительного представления

[См. небольшую заметку об искривленном пространстве в августовском выпуске журнала *Scientific American* за 1976 г. Она содержится также в сборнике [12].]

Погружение

[Читай « S три».]

[В настоящем разделе постепенно выяснится, что я понимаю под глобальной структурой.]

Сферические полярные координаты

Поскольку

$$x^2 + y^2 = \sin^2 \theta, \quad (38.3)$$

имеем

$$x^2 + y^2 + z^2 = 1 \quad (38.4)$$

для всех θ и ϕ . Каждая пара значений (θ, ϕ) задает одну точку на S^2 , причем различным значениям (θ, ϕ) может соответствовать одна и та же точка. Величины z и θ здесь связаны непосредственно: каждое значение z реализуется один и только один раз, когда θ пробегает область

$$0 \leq \theta \leq \pi. \quad (38.5)$$

При фиксированном z точки заматают окружность, задаваемую уравнением

$$x^2 + y^2 = 1 - z^2, \quad (38.6)$$

так что собственная область изменения ϕ должна иметь вид

$$0 \leq \phi \leq 2\pi. \quad (38.7)$$

Глобальная топология

[Этот удобный географический язык связан с геометрией только через оставшееся «гео».]

[Напомним, что формально диаграмма должна быть открытым множеством, и в этом случае вопрос о краях не возникает. Мы сочли более уместным включать в рассмотрение края и отождествлять повторяющиеся точки, чем использовать несколько диаграмм для описания множеств, являющихся многообразиями.]

[Теперь вы должны начинать понимать, что подразумевается здесь под «глобальной структурой». Она определяется характером отождествления краев.]

Таковы собственные области изменения наших координат. Что же происходит на краях этих областей? Для каждого значения θ_0 координаты $(\theta_0, 0)$ и $(\theta_0, 2\pi)$ задают одну и ту же точку на S^2 . На поверхности Земли получающийся таким образом «шов» называется международной линией смены дат. На полюсах координаты $(0, \phi_1)$ и $(0, \phi_2)$ задают одну и ту же точку для любых значений ϕ_1 и ϕ_2 . На рис. 38.1 представлена (θ, ϕ) -диаграмма с указанием надлежащих краевых условий. Легко

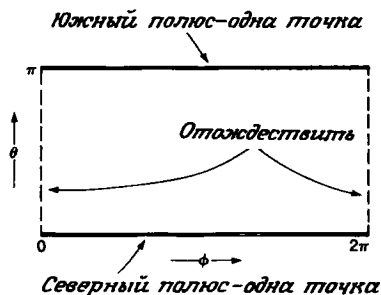


Рис. 38.1

Координатная диаграмма для 2-сферы. Двойные линии представляют собой в действительности отдельные точки, штриховые — одну и ту же линию.

убедиться, что эта диаграмма не содержит других составных точек. Договоримся изображать штриховой линией те швы, которые подлежат отождествлению, а двойной чертой — те линии, которые должны стягиваться в точки. Такая диаграмма дает одно из возможных полезных представлений множества точек, обладающего глобальной структурой S^2 .

Пример

Изображенный на рис. 37.4 треугольник с тремя прямыми углами представлен на нашей диаграмме на рис. 38.2.

Приведенная выше диаграмма и условия отождествления общих точек на ее границах служат удобным представлением S^2 . Действительно, она расположена в плоскости, и поэтому на ней можно производить графические построения. Любая книга по географии содержит массу примеров такого представления. Существуют и другие представления, которые имеют меньше странностей на краях, но в то же время все еще остаются в плоскости. Они играют роль промежуточных шагов при восстановлении S^2 , исходя из приведенной выше диаграммы.

Один из возможных первых шагов, модифицирующих приведенное выше представление, состоит в стягивании линии $\theta = 0$, которая на самом деле задает единственную точку на S^2 , в фактическую точку. В результате такого стягивания мы получаем сектор в форме кусочка торта, изображенный на рис. 38.3.

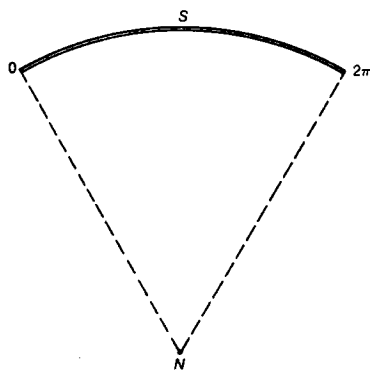


Рис. 38.3

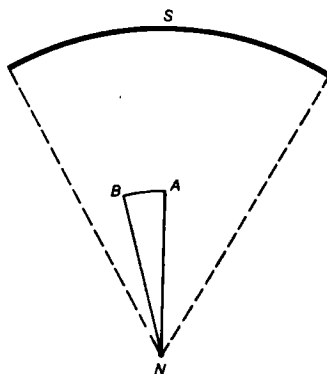


Рис. 38.4

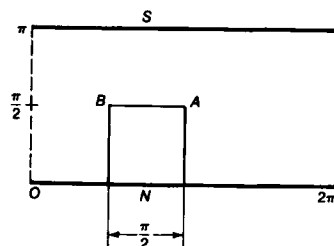


Рис. 38.2

Наш истинно прямоугольный треугольник, изображенный на рис. 37.4.

Модель двумерного диска

[Здесь не будут приводиться эти преобразования, так как они нам не нужны. Вы должны уметь записывать их сами.]

Пример

На рис. 38.4 в рассматриваемом представлении изображен наш истинно прямоугольный треугольник. Здесь он выглядит так, как и должен выглядеть треугольник. В соответствии с этим, однако, диаграмма не отображает структуры окрестности точки N .

Второй разумный шаг состоит в том, чтобы развернуть этот клин веером и затем склеить шов, как показано на рис. 38.5. В этом случае мы получаем представление, адекватное оригиналу всюду, за исключением точки S ; последняя оказывается развернутой во внешний обод. Такая диаграмма представляет собой полярную проекцию, столь популярную у военных стратегов. Это полезное представление, имеющее форму двумерного диска, оказывается наилучшим из всего того, что мы можем сделать, оставаясь в рамках двух измерений. Аналогичное представление S^3 будет наиболее полезным для нас представлением 3-сферы.

Рис. 38.6 демонстрирует заключительную стадию упомянутого выше восстановления S^2 . Вытянув диск в третье измерение, мы можем превратить линию, соответствующую Южному полюсу, в одну точку, одновременно сформировав вокруг нее надлежащую окрестность.

Таким образом, не существует принципиальных затруднений в графическом изображении описанной выше заключительной стадии восстановления S^2 , поскольку мы наделены интуитивной способностью к зрительному восприятию объектов в

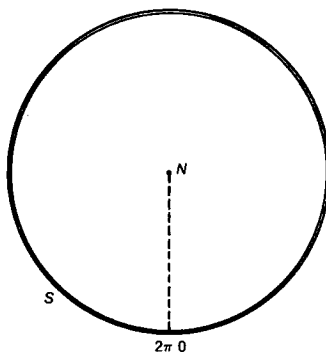


Рис. 38.5

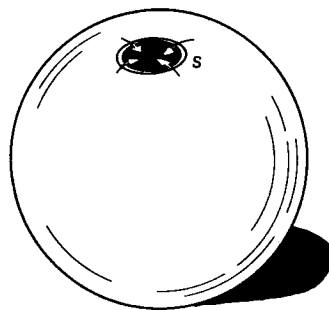


Рис. 38.6

трех измерениях. Однако для осуществления указанной стадии в случае S^3 требуется уже выход в четвертое измерение. Чтобы такой выход представить себе наглядным образом, необходимо прибегнуть к хитрости. Используем цвет для представления дополнительного измерения. Например, 2-поверхность, находящуюся в трех измерениях, можно представить в виде раскрашенной части плоскости. Расположение каждой точки в плоскости отвечает двум координатам, а цвет отвечает третьей координате. В качестве другой дополнительной переменной может выступать текстура. Ее труднее описывать словами, но легче изобразить.

Четвертое измерение

Используя цвет для представления третьего измерения, мы можем описать заключительную стадию восстановления S^2 , не выходя за пределы плоскости. Начнем с модели S^2 , имеющей форму двумерного диска. «Деформируем» его в третье измерение, раскрасив его синим цветом вблизи кромки с постепенным добавлением красного цвета по мере приближения к центру. В трех измерениях такая раскраска превращает диск в чашу. Теперь мы можем сжимать «окружность» Южного полюса, уменьшая размеры до тех пор, пока она не превратится в истинную точку, расположенную, если мы того захотим, прямо на вершине Северного полюса. Таким образом мы получим двумерный диск, каждая точка которого, за исключением кромки, раскрашена двумя цветами, указывающими на принадлежность к той или другой полусфере.

Воспользуемся теперь приведенными соображениями для реализации соответствующих представлений S^3 . Как и в случае сферических полярных координат, перейдем к параметрическому представлению, воспользовавшись следующими соотношениями:

3-сфера

$$\begin{aligned} w &= \cos \chi, \\ z &= \sin \chi \cos \theta, \\ x &= \sin \chi \sin \theta \cos \phi, \\ y &= \sin \chi \sin \theta \sin \phi. \end{aligned} \quad (38.8)$$

Как и выше, такая параметризация обеспечивает выполнение нашего определения:

$$x^2 + y^2 + z^2 + w^2 = 1. \quad (38.9)$$

Каковы пределы изменения координат? Каждое значение w соответствует единственному значению χ в области

$$0 \leq \chi \leq \pi. \quad (38.10)$$

Фиксируя χ , мы получаем множество точек, удовлетворяющих соотношению

$$x^2 + y^2 + z^2 = 1 - w^2, \quad (38.11)$$

которое показывает, что при $w \neq 1$ это множество представляет собой 2-сферу со сферическими полярными координатами (θ, ϕ) . Таким образом, для θ и ϕ должны иметь место обычные пределы изменения координат и аналогичные описанным выше отождествления. На рис. 38.7 изображена такая диаграмма и указаны правила отождествления.

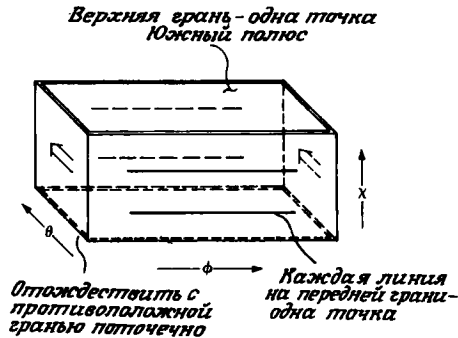


Рис. 38.7
3-сфера в координатах (χ, θ, ϕ) .

Модель трехмерного шара

[Отметим, что Северным полюсом (с заглавной буквы) мы будем называть только истинный, отвечающий условию $w = 1$.]

Опять мы можем получить более удобное представление, если начнем восстанавливать S^3 , исходя из диаграммы на рис. 38.7. Посмотрим с этой целью на грань $\theta = 0$. Для каждого χ эта грань соответствует не линии, а точке. Стянем эти линии так, чтобы они действительно стали точками. Указанные точки представляют собой северные полюсы «малых» 2-сфер, на которые можно разделить 3-сферу S^3 . Таким образом мы получим показанный на рис. 38.8 клин, напоминающий кусок сыра. Следуя двумерной аналогии, далее мы развернем его в полный круг сыра (рис. 38.9). Верхняя и нижняя поверхности этого круга отвечают единственной точке, а первая искривленная грань представляет собой одномерное семейство южных полюсов, $\theta = \pi$.

Можно повторить эту процедуру. Стянем нижнюю грань — Северный полюс — в одну точку. Мы получим шаровой сектор, изображенный на рис. 38.10. Наконец, постепенно увеличивая раствор телесного угла при вершине, развернем этот шаровой сектор и сомкнем линию южных полюсов в истинную

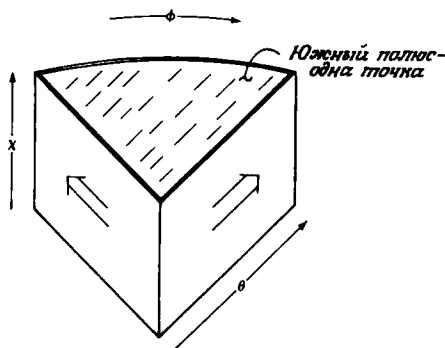


Рис. 38.8
Стянем линии, изображающие «малые» северные полюсы, в истинные точки.

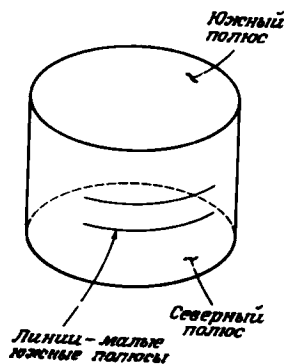


Рис. 38.9
Сожмем все международные линии смены дат 2-сфер.

линию, состоящую из точек. Полученное таким образом представление S^3 адекватно оригиналу в окрестности Северного полюса. Назовем это представление моделью трехмерного шара для S^3 ; оно изображено на рис. 38.11. В качестве радиальной координаты внутри такого шара выступает x , а угловые координаты (θ, ϕ) представляют собой обычные угловые координаты на сферах. В дальнейшем мы будем неоднократно обращаться к такому представлению S^3 .

Для окончательного завершения процесса восстановления S^3 вытянем Южный полюс в четвертое измерение. Окрасим трехмерный шар в синий цвет на поверхности с постепенным добавлением красного цвета по мере приближения к его центру.

В четыре измерения

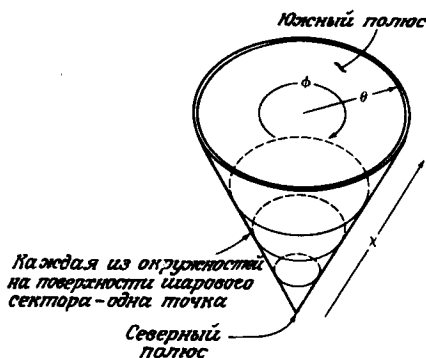


Рис. 38.10
Сожмем Северный полюс в одну точку.

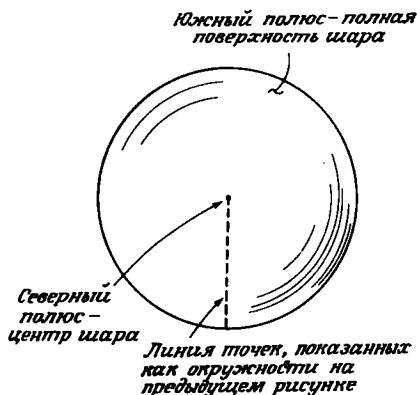


Рис. 38.11

Стянем все линии малых южных полюсов в точки.

Втянем затем Южный полюс внутрь и стянем его в линию, расположенную на вершине Северного полюса. Каждая точка трехмерного шара представляет теперь две точки S^3 . В последующих обсуждениях и расчетах вы будете неоднократно иметь возможность позабыться с S^3 . Используя полученное представление, можно убедиться в том, что S^3 имеет много общих свойств с S^2 . У нее конечный объем, но нет границы. На S^3 в общем случае не существует подобных геометрических фигур. Так, например, тетраэдры, размеры сторон которых отличаются в два раза, будут иметь различные двугранные углы. Плоскости, проходящие через начало координат фиктивного 4-пространства, образуют сечения, представляющие собой 2-сферы. В S^3 эти 2-сферы являются аналогами плоскостей точно так же, как большие окружности на S^2 аналогичны прямым линиям.

ЗАДАЧИ

38.1. (06) Что представляет собой острая кромка кусочка сыра, изображенного на рис. 38.8 (ближайшая к нам вертикальная линия): линию или точку? Найдите ее в модели трехмерного шара.

38.2. (16) Рассмотрите более подробно переход от рис. 38.10 к рис. 38.11. Сделайте несколько промежуточных рисунков.

38.3. (28) Начав с рис. 38.7, восстановите S^3 в следующей последовательности. Сначала соберите N в одну точку, затем склейте отождествляемые участки и сожмите окружности. Что получится в итоге? Нарисуйте.

38.4. (20) Соберите бутылку Клейна, представляющую собой квадратный участок плоскости с правилами отождествления краев, указанными на рис. 38.12. Используйте цвет для четвертого измерения. Что можно сказать относительно использования времени для представления четвертого измерения?

38.5. (22) На рис. 38.13 изображена дважды скрученная полоска. Полоска, скрученная один раз, была бы листом Мёбиуса. Покажите, как раскрутить эту полоску, не продевая ее сквозь себя, если допускается движение в четвертом измерении. Нарисуйте.

38.6. (10) Постройте диаграмму для окрестности Северного полюса на 2-сфере. Выпишите фактические отображения.

38.7. (12) Сделайте для S^3 то же самое, что и в задаче 38.6.

38.8. (15) Постройте диаграмму для окрестности международной линии смены дат на 2-сфере.

38.9 (20) Постройте такую же диаграмму, как в задаче 38.8, для поверхности отождествления $\phi = 0, 2\pi$ на S^3 .

38.10 (23) Геодезическими на S^2 служат большие окружности, получающиеся при пересечении сферы с плоскостями, проходящими через начало системы координат евклидова 3-пространства, с которого мы начинали наше рассмотрение. Что представляют собой эти кривые на (θ, ϕ) -диаграмме?

$$u \mapsto (\theta, \phi) = \left(\frac{\pi}{2} + \sin s, \sin 3s \right).$$

38.11. (30) Проведите для S^3 такое же рассмотрение, как и в задаче 38.10.

38.12. (13) Каждая из двух линий A и B , показанных на рис. 38.14, проходит через Северный полюс. Чему равен угол между ними? Представьте как графическое решение, так и аккуратное доказательство с использованием диаграммы, адекватной оригиналу в окрестности полюса. Где должно быть продолжение линии B ?

38.13. (33) Рассмотрите множество всех геодезических на S^2 . Это множество является многообразием. Покажите, что оно представляет собой часть S^2 , имеющую размерность два, но с некоторыми отождествлениями. Такое многообразие называется RP^2 , или действительным проективным 2-пространством.

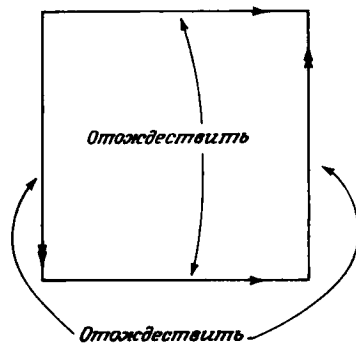


Рис. 38.12

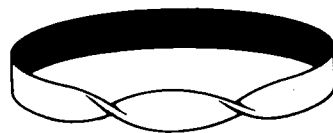


Рис. 38.13

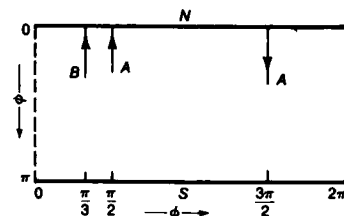


Рис. 38.14

38.14. (30) Как и лист Мёбиуса, RP^2 имеет только одну сторону. Внутреннее (т.е. в пределах поверхности) описание этого свойства может быть реализовано при помощи следующего приема. Попытаемся самосогласованным образом заполнить RP^2 подходящими для этой цели буквами (скажем, **B**, но не **S**). Покажите, что такое заполнение может быть осуществлено для S^2 , но не для RP^2 .

38.15. (28) Опишите свойства кривой, заданной на S^2 уравнением

$$u \mapsto (\theta, \phi) = \left(\frac{\pi}{2} + \sin s, \sin 3s \right).$$

39. Метрическая структура 3-сферы

Рассмотрев глобальную структуру S^3 , обратимся теперь к ее локальной метрической структуре. Однородность и изотропия S^3 относятся к ее метрической структуре. Мы можем получить эту метрическую структуру (появление которой в разд. 37 было подобно грому среди ясного неба), используя представление S^3 как подпространства евклидова 4-пространства. В следующем разделе мы сможем начать изучение космологических моделей, построенных на основе S^3 .

Метрика из погружения

Итак, рассмотрим представление S^3 как трехмерного подпространства евклидова 4-пространства. Мы можем воспользоваться методами разд. 20, чтобы ограничить базисные 1-формы этого 4-пространства на S^3 . Поскольку метрика на 4-пространстве может быть записана при помощи этих базисных 1-форм, она также может быть ограничена на S^3 . Таким образом мы можем получить естественную метрику на S^3 .

Пример 1

2-сфера

Рассмотрим сферу как подпространство евклидова 3-пространства и рассчитаем ограничение 1-формы dz . Координата z представляет собой функцию на этом подпространстве:

$$z^*: S^2 \rightarrow \mathbb{R}; (\theta, \phi) \mapsto \cos \theta. \quad (39.1)$$

Как функция на S^2 она обладает градиентом, который мы обозначим dz^* . Таким образом.

$$dz^* = -\sin \theta d\theta. \quad (39.2)$$

Полученная 1-форма на S^2 является ограничением dz как 1-формы на 3-пространстве. Как только мы определили ограничение базисных 1-форм, любая 1-форма может быть подвергнута этой операции.

Заметим, что наша теория развита как раз в форме, удобной для таких расчетов. Подпространства определены аналогично параметризованным кривым как отображения в пространство большего числа измерений; 1-формы всегда могут быть ограничены на подпространство, т.е. в направлении, противоположном направлению действия преобразований.

Мы можем также ограничить dx и dy на подпространство. Соответствующие формулы имеют вид

$$dx^* = \cos \theta \sin \phi \, d\theta + \sin \theta \cos \phi \, d\phi, \quad (39.3) \quad [\text{Продолжение примера 1.}]$$

$$dy^* = \cos \theta \cos \phi \, d\theta - \sin \theta \sin \phi \, d\phi. \quad (39.4)$$

С помощью этих формул мы ограничиваем метрику

$$\mathcal{E} = dx^2 + dy^2 + dz^2 \quad (39.5)$$

и находим метрику на S^2 :

$$\mathcal{E} = d\theta^2 + \sin^2 \theta \, d\phi^2. \quad (39.6)$$

[Мы продолжаем небрежно обращаться с обозначениями, используя один и тот же символ для $\mathcal{E}$ и его ограничения на подпространства. Будучи педантами, мы могли бы назвать эти ограничения ${}^3\mathcal{E}$ и ${}^2\mathcal{E}$.]

Поступая точно таким же образом, как и в приведенном выше примере, мы можем ограничить метрику 4-пространства

$$\mathcal{E} = dx^2 + dy^2 + dz^2 + dw^2. \quad (39.7)$$

В результате получаем метрику на S^3

$$\mathcal{E} = d\chi^2 + \sin^2 \chi \, d\Omega^2, \quad (39.8)$$

которая приводилась ранее.

Поскольку примененный нами подход полностью ковариантен, мы можем выбирать координаты, наиболее подходящие для решения тех или иных конкретных проблем. Иногда в качестве радиальной координаты вместо χ удобно использовать радиус 2-сфер. Такой выбор радиальной координаты, применяемый многими авторами, тесно связан с двойным покрытием представления S^3 в виде трехмерного шара. Определим новую радиальную координату посредством соотношения

Другие координаты

$$\rho \equiv \sin \chi; \quad (39.9)$$

тогда

$$d\rho = \cos \chi \, d\chi, \quad (39.10)$$

$$d\chi^2 = \frac{d\rho^2}{1 - \rho^2}. \quad (39.11)$$

Метрика $\mathcal{E}$ в этих координатах выглядит следующим образом:

$$\mathcal{E} = \frac{d\rho^2}{1 - \rho^2} + \rho^2 d\Omega^2. \quad (39.12)$$

Заметим, что решение уравнения (39.9) для χ в зависимости от ρ не единственно. Как и в случае диска, каждая тройка координат (ρ, θ, ϕ) , за исключением $\rho = 1$, соответствует на S^3 двум точкам.

Кривизна пространства
[Отличное рассмотрение 3-сферы содержится в книге [32]. На семи страницах автор описывает путешествие по 3-сфере, дополненное диаграммами с изображением того, что вы увидели бы в таком путешествии.]

Координаты (χ, θ, ϕ) фиксируют собственные расстояния между 2-сферами, на которых изменяются лишь θ и ϕ . Координаты же (ρ, θ, ϕ) фиксируют собственные размеры этих сфер. В силу того что S^3 представляет собой искривленное пространство, невозможно построить такие координаты, которые независимо задавали бы как собственные размеры упомянутых сфер, так и собственные расстояния между ними.

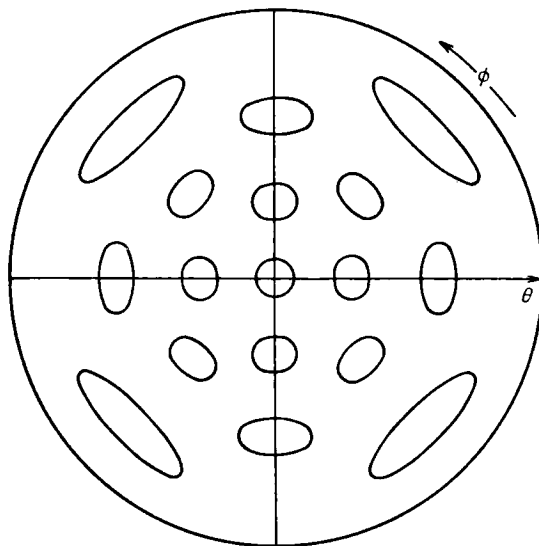


Рис. 39.1

Метрика на 2-сфере в представлении двумерного диска. Масштаб касательного пространства уменьшен в пять раз.

Мы можем описывать рассматриваемую метрическую структуру с помощью представления метрического тензора в касательных пространствах. Однако нарисовать такой трехмерный рисунок было бы сложно. В разрезе, проходящем через центр S^3 , получается S^2 . Используя введенное выше представление типа двумерного диска, мы изобразили метрическую структуру такой 2-сферы на рис. 39.1. Обратите внимание на то, как раздуваются метрические фигуры по мере приближения к ободу, который на самом деле представляет собой единственную точку. Ясно, что в этой точке происходит что-то странное. Метрические фигуры 3-сферы представляют собой эллипсоиды, которые получаются из изображенных на рис. 39.1 эллипсов посредством вращения вокруг радиального направления.

Метрические фигуры

Если представить себе, что рассматриваемые метрические фигуры соответствуют волновым диаграммам, то поведение геодезических можно объяснить с помощью построения Гюйгенса. На рис. 39.2 в представлении S^2 типа двумерного диска изображены геодезические с расположенными вдоль них поверхностями постоянной фазы. Указанные геодезические являются большими окружностями.

Геодезические

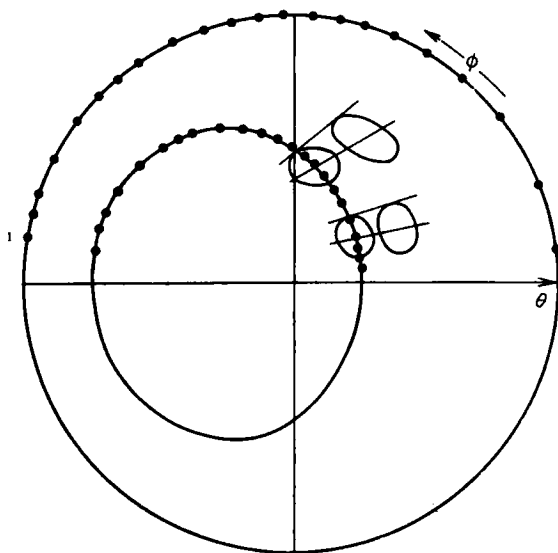


Рис. 39.2

Геодезические на 2-сфере в представлении двумерного диска.

О геометрии S^3 мы узнаем больше, когда действительно будем пользоваться ею. Обратимся теперь к космологическим моделям, основанным на S^3 . Соответствующее обсуждение свойств псевдосферы отложим до разд. 44. Геометрия псевдосферы еще более необычна. Нам будет легче заняться ее изучением, когда мы больше привыкнем к свойствам S^3 , а также приобретем некоторый опыт применения рассмотренных выше симметричных пространств.

ЗАДАЧИ

39.1. (17) Чему равна площадь поверхности большой сферы на S^3 ?

39.2. (16) Рассчитайте ограничение 1-формы

$$x dx + y^2 dy + z^3 dz$$

на подмногообразии $\mathbb{R}^3$, задаваемое следующим образом:

$$(u, v) \mapsto (x, y, z) = (u, v, v^2).$$

39.3. (16) Покажите, как вычислить размеры эллипсов, изображенных на рис. 39.1.

39.4. (30) Найдите для S^2 такую конформную систему координат, метрика в которой имеет вид

$$\mathcal{G} = f(u, v) (du^2 + dv^2).$$

39.5. (22) Перерисуйте рис. 39.2 в координатах задачи 39.4.

39.6. (28) Определите и вычислите объем S^3 .

40. Распространение света

*Чем клясть тьму,
Лучше зажечь свечу.*

Конфуций

Большая часть тех сведений о Вселенной, которыми мы располагаем, получена путем оптических наблюдений¹⁾. Чтобы опи-

¹⁾ Поразительные успехи космологии и астрофизики, достигнутые за последние несколько десятилетий, теснейшим образом связаны с освое-

сывать такие наблюдения, мы должны знать, как распространяется свет в моделях пространства-времени Робертсона — Уокера. Для ответа на этот вопрос достаточно рассмотреть движение в плоскости (t, χ) при постоянных θ и ϕ . При указанном условии интересующие нас световые сигналы выходят в радиальном направлении из Северного полюса. Такие световые сигналы в расширяющейся Вселенной испытывают сдвиг частот в красную сторону, что в действительности наблюдается в современной Вселенной. Вследствие однородности и изотропии Вселенных Робертсона — Уокера все световые сигналы эквивалентны, но радиальные сигналы распространяются вдоль одной из осей нашей системы координат, что облегчает их описание.

Радиальные световые сигналы

Изучение распространения света упрощается, если ввести новую временную координату, специально подобранную для этой цели. Эта новая временная координата, которую мы будем называть дуговым временем η , определяется посредством обыкновенного дифференциального уравнения

Дуговое время

$$\frac{d\eta}{dt} = \frac{1}{R(t)}, \quad (40.1)$$

где $R(t)$ — функция в моделях пространства-времени Робертсона — Уокера, описывающая эволюцию Вселенной. С помощью дугового времени метрику (37.11) можно записать в следующем виде:

$$\mathcal{G} = R^2(t) [-d\eta^2 + d\chi^2 + S^2(\chi) d\Omega^2]. \quad (40.2)$$

В координатах (η, χ) мировые линии радиальных световых сигналов наклонены, как и в специальной теории относительности, под углом 45° , поскольку их касательный вектор σ должен удовлетворять соотношению

$$\mathcal{G}(\sigma, \sigma) = 0; \quad (40.3)$$

записывая

$$\sigma = \dot{\eta} \frac{\partial}{\partial \eta} + \dot{\chi} \frac{\partial}{\partial \chi}, \quad (40.4)$$

нием новых, нетрадиционных диапазонов регистрации электромагнитных излучений. Прежде всего современные космология и астрофизика совершенно немыслимы без высокоразвитой радиоастрономии. Все большее развитие получает также инфракрасная, рентгеновская и γ -астрономия. Все перечисленные типы излучений распространяются, как и видимый свет, по изотропным геодезическим. — *Прим. ред.*

получим

$$\dot{\eta}^2 = \dot{\chi}^2. \quad (40.5)$$

Дуговое расстояние

На S^3 существует естественная мера расстояния — дуговое расстояние, соответствующее длине большой окружности на Земле. Дуговое расстояние вокруг экватора S^n по определению всегда должно быть равно 2π . В радиальном направлении дуговое расстояние для S^3 задается непосредственно координатой χ . Параметр η потому и назван дуговым временем, что он тесно связан с введенным дуговым расстоянием¹⁾, а именно за данный отрезок дугового времени световые сигналы проходят в точности такой же отрезок дугового расстояния. Эта важная связь неоднократно будет использоваться нами в дальнейшем.

Сдвиги частот

Рассмотрим два световых сигнала, испущенных в одном и том же радиальном направлении в разные моменты времени. Мы можем воспользоваться этими сигналами, чтобы вычислить сдвиг частот для источника и наблюдателя, покоящихся относительно симметричной опорной системы отсчета Вселенной. Как показано на рис. 40.1, интервал дугового времени между такими световыми сигналами должен быть постоянным. Как и при вычислении гравитационного красного смещения, это позволяет нам записать соотношение между $\Delta\tau_1$ (временем между двумя световыми сигналами, измеренным по часам источника) и $\Delta\tau_2$ (временем между ними, измеренным по часам наблюдателя) для источника и наблюдателя, движущихся в направлении $\partial/\partial t$. Для других состояний движения необходимо также включать эффект спецрелятивистского доплеровского сдвига. Полагая, что за один период волны размер Вселенной заметно не меняется, преобразуем дуговое время в собственное:

$$\Delta\tau = R \Delta\eta. \quad (40.6)$$

Поскольку длина световой волны пропорциональна ее периоду, имеем

$$\frac{\lambda_{\text{REC}}}{\lambda_{\text{EMIT}}} = \frac{R(t_{\text{REC}})}{R(t_{\text{EMIT}})}. \quad (40.7)$$

Если в момент приема света R больше, то принятая длина во-

[Здесь мы используем τ вместо t , чтобы подчеркнуть, что это именно собственное, а не координатное время.]

[Эту формулу легко запомнить. Как τ , так и R имеют размерность секунды, тогда как величина η безразмерна.]

¹⁾ В русской литературе вместо терминов «дуговое время» и «дуговое расстояние» употребляются другие — «конформное время» и «конформное расстояние», которые, впрочем, как и введенные здесь термины, не отражают в полной мере специфики соответствующих понятий. — *Прим. перев.*

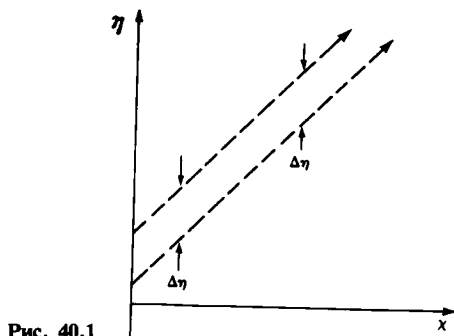


Рис. 40.1

лны λ_{REC} также больше, и мы говорим, что свет испытывает красное смещение. Общеупотребительный параметр красного смещения z определяется соотношением

$$(1 + z) = \frac{\lambda_{\text{REC}}}{\lambda_{\text{EMIT}}}. \quad (40.8)$$

При малых значениях параметра z он представляет собой просто скорость (отнесенную к скорости света), которая вызывала бы соответствующее красное смещение.

Такое смещение, пропорциональное длине волны, заставляет задуматься. Оно превращает излучение с чернотельным спектром в излучение с чернотельным же спектром, но другой температурой в соответствии с формулой

Чернотельное излучение

$$TR = \text{const}. \quad (40.9)$$

В процессе распространения света в расширяющейся Вселенной число фотонов сохраняется. Вместе с тем каждый фотон испытывает сдвиг в направлении больших длин и, таким образом, теряет свою энергию. Возникает вопрос: куда расходуется эта энергия? Пока никто не может на него ответить. Дело в том, что не существует способа определить полную энергию пространства-времени. Сказать, что энергия света идет на расширение Вселенной, — значит пробормотать нечто безвредное и даже приятное на слух, но вместе с тем не имеющее никакого смысла. Такая фраза не объясняет, она всего лишь успокаивает.

Сохранение фотонов

41. Фридмановские Вселенные

Чтобы идти дальше, мы должны использовать динамическую теорию тяготения для установления вида функции $R(t)$ в моде-

Динамика пространства-времени

[Вывод можно найти в гл. 10 книги [25].]

Опорная космическая система отсчета

Материя

Пыль

Излучение

лях пространства-времени Робертсона — Уокера. Здесь будет пробел в нашем рассмотрении. К сожалению, мы пока не умеем применять уравнения Эйнштейна, выписанные на с. 298, к моделям пространств-времени Робертсона — Уокера. Нужно потратить определенное время, чтобы овладеть в достаточной мере геометрическим формализмом. Лишь тогда вычисление тензоров кривизны не будет казаться весьма сложным делом. Уравнения Эйнштейна устанавливают взаимосвязь между динамикой пространства-времени и энергией, импульсом и натяжением содержащейся в нем материи. В результате вычислений получаются обыкновенные дифференциальные уравнения для масштабного фактора $R(t)$, плотности энергии и давления содержащегося во Вселенной вещества. На основе этих обыкновенных дифференциальных уравнений, называемых уравнениями Фридмана, мы сможем продолжить наше изучение космологии без дальнейших пробелов.

Задаваемая уравнениями Эйнштейна взаимосвязь между пространством-временем и содержащейся в нем материей такова, что симметричному пространству-времени должно соответствовать в той же степени симметричное распределение материи. Рассмотрим тела в опорной космической системе отсчета, ось времени которой ориентирована по направлению $\partial/\partial t$. Если бы в этой системе существовал какой-либо результирующий поток энергии, то он выделял бы некоторое особое направление. Это обстоятельство противоречило бы предполагаемой изотропии пространства. Аналогично из всех натяжений допустимо лишь изотропное давление. Другие натяжения, такие, как сдвиг, противоречили бы изотропии. Таким образом, набор возможных параметров, характеризующих материю и совместимых с предполагаемой нами симметрией, исчерпывается плотностью энергии $\rho(t)$ и давлением (плотностью потока импульса) $p(t)$. Несмотря на то что обе эти величины изменяются с течением времени, они не должны зависеть от χ , θ и ϕ .

Различные виды материи характеризуются различными соотношениями между давлением и плотностью. Такое соотношение называется уравнением состояния. Мы будем подробно изучать лишь два основных вида материи. Первый, называемый пылью, описывается уравнением состояния

$$p = 0. \quad (41.1)$$

Второй мы будем называть излучением; этот вид материи

характеризуется уравнением состояния

$$p = \frac{1}{3}\rho. \quad (41.2)$$

Для описания сглаженной жидкости, частицы которой представляют собой обычные галактики, воспользуемся уравнением состояния пыли. Чтобы показать, что давление таких частиц пренебрежимо мало, полезно пояснить понятие давления. Начнем с импульса. Импульс есть поток массы (энергии). Аналогично натяжение (давление) есть поток импульса. Поток импульса характеризует величину, которую обычно называют силой, приходящейся на единичную площадь. В жидкости, состоящей из частиц массой m , движущихся со скоростью v во всех направлениях, энергия, приходящаяся на одну частицу, ведет себя как γm . Импульс есть поток энергии. Для одной частицы он равен скорости, умноженной на плотность энергии, т.е. $v\gamma m$. Это векторная величина, и в газе, состоящем из частиц, которые движутся во всевозможных направлениях, общий поток импульса равен нулю. Аналогично давление представляет собой поток импульса, и для одной частицы оно ведет себя как $v^2\gamma m$. Оно не равно нулю даже для частиц, движущихся во всех направлениях, поскольку изменяется как v^2 . Для частиц типа галактик наблюдаемые случайные скорости имеют порядок 300 км/с, т.е. 10^{-3} от скорости света. Давление меньше плотности энергии благодаря множителю v^2 , имеющему порядок 10^{-6} . Таким образом, мы видим, что в настоящее время давление обычных галактик не играет существенной роли в эволюции Вселенной.

Галактики

Ситуация с излучением иная. Допустим, что частицы движутся со скоростью v , близкой к скорости света. Тогда величина, характеризующая среднее давление, будет содержать среднее от v^2 по всем направлениям. Поскольку независимых направлений три, среднее равно $1/3$; следовательно, уравнение состояния для частиц, движущихся со скоростью света, имеет вид

$$p = \frac{1}{3}\rho. \quad (41.3)$$

Это уравнение состояния описывает чернотельное излучение, представляющее собой хаотический фотонный газ.

Наша космологическая модель содержит три функции времени: масштабный фактор R , плотность энергии ρ и давление p . Из уравнений Эйнштейна получаются два обыкновенных дифференциальных уравнения для этих функций, а роль третьего уравнения играет уравнение состояния. Обыкновенные дифференциальные уравнения, получающиеся из собственно общере-

[«Следовательно» — сказано слишком сильно; приведенные соображения, однако, по крайней мере уменьшают степень неожиданности.]

Уравнения Фридмана

лятивистских расчетов, имеют следующий вид:

$$2\frac{R''}{R} + \left(\frac{R'}{R}\right)^2 + \frac{\kappa}{R^2} + 8\pi p = 0, \quad (41.4)$$

$$\left(\frac{R'}{R}\right)^2 + \frac{\kappa}{R^2} = \frac{8\pi\rho}{3}, \quad (41.5)$$

где

$$R' \equiv dR/dt \quad (41.6)$$

и κ определяется соотношением

$$S'' = -\kappa S, \quad (41.7)$$

т.е. $\kappa = +1$ для 3-сферы, $\kappa = 0$ для 3-пространства и $\kappa = -1$ для 3-псевдосферы. Модель пространства-времени Робертсона — Уокера, удовлетворяющая приведенным выше уравнениям и соответствующему уравнению состояния, называется Фридмановской Вселенной.

Экзотическая материя

Помимо уравнений для пыли, излучения и любой смеси этих двух видов материи иногда рассматриваются и другие уравнения состояния. Ученые, занимающиеся физикой элементарных частиц, предположили, что при огромных плотностях ядерная материя может подчиняться уравнению состояния

$$p = \rho. \quad (41.8)$$

Такая материя, называемая предельно жесткой, оказывает большее сопротивление сжатию, чем даже фотонный газ. Будь она еще жестче, скорость звука в ней превысила бы скорость света. Возможен и другой фантастический вид материи. Требование лоренц-инвариантности уравнения состояния приводит к следующему уравнению:

$$p = -\rho = -\frac{\Lambda}{8\pi}. \quad (41.9)$$

Материя, описываемая этим уравнением, имела бы всюду одно и то же постоянное отрицательное давление. Такому виду материи соответствует в уравнениях Эйнштейна так называемый космологический член, а Λ называется космологической постоянной. Эйнштейн сам выдумал этот вид материи именно для того, чтобы удержать от расширения свои модели Вселенной. Он разрабатывал такие модели еще до наблюдений Хаббла, установивших фактическое расширение Вселенной. Тем не менее с обоими этими видами материи интересно поупражняться.

42. Замкнутая пылевая Вселенная

Природа — это бесконечная сфера, центр которой повсюду, а периферия — нигде.

Паскаль

[В той или иной форме эта цитата приписывалась великому множеству авторов. Кое-что об этом рассказано в очерке «Огромная сфера Паскаля», содержащемся в «Лабиринтах» Джорджа Луиса Боргеса.]

Сначала рассмотрим следующую космологическую модель. Пусть ее пространственные сечения представляют собой 3-сферы, а в качестве материи она содержит пыль. Это самый простой и в то же время наиболее известный случай. Вы уже приобрели небольшой опыт в сферической тригонометрии, но пока еще совсем не имели дела с псевдосферической тригонометрией.

Для рассматриваемой пылевой Вселенной мы должны определить только две функции, R и ρ . Уравнения Фридмана имеют вид

$$2\frac{R''}{R} + \left(\frac{R'}{R}\right)^2 + \frac{1}{R^2} = 0, \quad (42.1) \quad \text{Уравнения Фридмана}$$

$$\left(\frac{R'}{R}\right)^2 + \frac{1}{R^2} = \frac{8\pi\rho}{3}. \quad (42.2)$$

Естественно предложить следующий способ исследования этой системы: сначала решить первое уравнение относительно $R(t)$, а затем вычислить ρ из второго. Однако эти уравнения открывают больше возможностей, чем может показаться с первого взгляда.

Для пыли плотность энергии ρ равна просто локальной плотности массы. Для жидкости, состоящей из частиц, плотность массы равна массе частиц, умноженной на локальную плотность числа частиц. Эти частицы (для нас — галактики) будут сохраняться, так что их полное число должно быть постоянным. Сказанное может быть справедливо только в том случае, если в процессе расширения Вселенной плотность числа частиц убывает как $1/R^3$. Что отсюда следует? Совместимо ли упомянутое условие с уравнениями Фридмана? Вычислим величину $(R^3\rho)'$ и посмотрим, равна ли она нулю. Запишем уравнение (42.2) в виде

Материя

$$R(R')^2 + R = \frac{8\pi\rho R^3}{3} \quad (42.3) \quad \text{Сохранение массы}$$

и продифференцируем обе части:

$$\frac{d}{dt} \left(\frac{8\pi\rho R^3}{3} \right) = (R')^3 + 2RR'R'' + R'. \quad (42.4)$$

Выражение в правой части (42.4) можно получить, умножив (42.1) на R^2R' ; поэтому мы заключаем, что

$$\frac{d}{dt} (\rho R^3) = 0. \quad (42.5)$$

Итак, условие сохранения массы содержится в уравнениях Фридмана. Воспользовавшись этим условием, определим постоянную a следующим образом:

$$a \equiv \frac{8\pi\rho R^3}{3}. \quad (42.6)$$

Величина a остается постоянной в процессе эволюции Вселенной. Различные значения a будут приводить к различным моделям Вселенной, т.е. моделям, характеризующимся различными размерами. Зная a , мы можем непосредственно связать ρ и R . Проще всего получить R из уравнения (42.3):

Приведенное уравнение
$$R(R')^2 + R = a. \quad (42.7)$$

Нахождение интеграла движения понижает порядок дифференциального уравнения со второго до первого.

Смысл постоянной a можно выяснить, рассмотрев уравнение (42.7). Если в некоторый момент времени

$$R' = 0, \quad (42.8)$$

то тогда, очевидно,

$$R = a. \quad (42.9)$$

Из уравнения (42.1) в этом случае имеем

$$R'' < 0; \quad (42.10)$$

иными словами, рассматриваемое состояние Вселенной будет состоянием ее максимального расширения, а постоянная a имеет смысл максимального радиуса. Функция $R(t)$ не может иметь никакой другой стационарной точки, следовательно, минимального отличного от нуля радиуса Вселенной не существует.

Дуговое время

Уравнение (42.7) можно также записать в дуговом времени:

$$\left(\frac{dR}{d\eta} \right)^2 = R(a - R), \quad (42.11)$$

откуда после интегрирования получаем:

$$R(\eta) = a \sin^2\left(\frac{\eta}{2}\right) = \frac{a}{2}(1 - \cos \eta) \quad (42.12)$$

Постоянная интегрирования выбрана так, что $t = 0$ при $R = 0$. Соответствующая зависимость представлена на рис. 42.1.

Для астрономических целей дуговое время — наиболее удобная временная переменная. Для геологических же целей, например для установления возраста горных пород и звезд, более удобно собственное время. Для преобразования одного времени в другое можно воспользоваться соотношением

$$\frac{dt}{d\eta} = R(\eta) = \frac{a}{2}(1 - \cos \eta), \quad (42.13)$$

из которого после интегрирования получаем

$$t(\eta) = \frac{a}{2}(\eta - \sin \eta). \quad (42.14)$$

Здесь мы также выбрали постоянную интегрирования с тем расчетом, чтобы время было равно нулю в момент, когда Вселенная имела нулевой радиус. Эволюция такой Вселенной в собственном времени показана на рис. 42.2.

Чтобы применять такую модель Вселенной, нам необходимо знать две величины. Одна из них — это параметр a — максимальный размер Вселенной, имеющий размерность длины. Другая — наше современное положение во Вселенной, современное дуговое или собственное время. Обе эти величины трудно поддаются измерению. Первая потому, что она вводит в рассмотрение масштаб длины, который делает необходимым относить наши лабораторные длины к астрономическим.

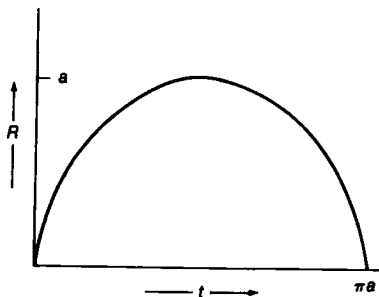


Рис. 42.2
Эволюция пылевой Вселенной в собственном времени.

Собственное время

[Надеюсь, мои коллеги не обидятся на меня за то, как я представляю себе более логичное разделение этих отраслей знания. Астрономия имеет дело со световым конусом прошлого, а геология — с прошлым времени подобных мировых линий.]

Параметры модели

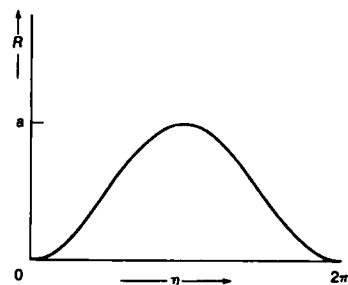


Рис. 42.1.
Эволюция пылевой Вселенной в дуговом времени.

Определение масштаба астрономических расстояний требует далекой и сложной экстраполяции. Измерение второй величины сложно потому, что оно требует учета эффектов второго порядка по расстоянию. Локально каждое пространство-время выглядит плоским. Цель наблюдательной космологии как раз и состоит в определении значений этих двух параметров. Такая задача весьма сложна, и к настоящему времени достигнуты лишь сравнительно скромные результаты. В следующем разделе мы обсудим наблюдения, которые в принципе позволяют определить указанные параметры.

Конечное время жизни

Эволюция рассматриваемой Вселенной происходит таким образом, что по истечении конечного дугового и собственного времени она коллапсирует опять в состояние $R = 0$. Вселенная существует в течение конечного интервала собственного времени πa и конечного интервала дугового времени 2π . Из последнего ясно, что за время существования Вселенной световой сигнал может обойти вокруг нее ровно один раз. Благодаря этому мы не можем, смотря вперед, увидеть собственный затылок. Теперь для практики построим в различных представлениях изображение такого светового сигнала, совершающего кругосветное путешествие.

Поскольку изменяются все четыре координаты, точное изображение соответствующей мировой линии получить трудно. Для начала не будем учитывать угловые координаты θ и ϕ . На рис. 42.3 показан вид такой мировой линии в плоскости (η, χ) :

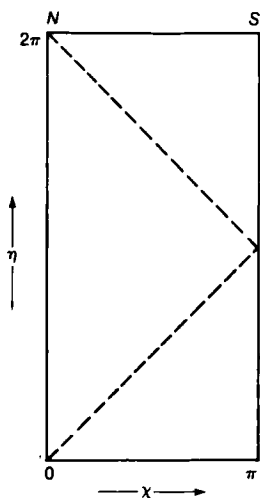


Рис. 42.3
Световой сигнал, распространяющийся вокруг Вселенной, в координатах (η, χ) .

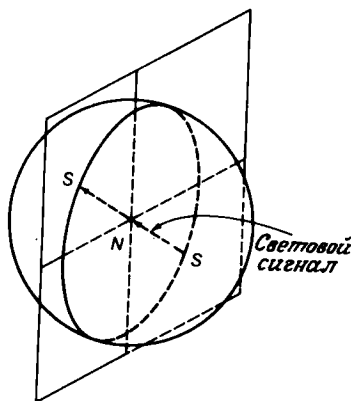


Рис. 42.4
Проекция светового сигнала в представлении трехмерного шара для 3-сферы.

она начинается на Северном полюсе, т.е. в точке $(\eta, \chi) = (0, 0)$, и со скоростью света подходит к Южному полюсу, удаленному на дуговое расстояние π . Далее она пересекает Южный полюс, где $(\eta, \chi) = (\pi, \pi)$, и, обойдя вокруг Вселенной «сзади», возвращается на Северный полюс. Чтобы понять, что мы имеем в виду, употребляя термин «сзади», посмотрим на рис. 42.4, где координата η опущена, а χ , θ и ϕ представлены в модели трехмерного шара. Изображенный на этом рисунке луч выходит из Северного полюса в направлении $(\theta, \phi) = (\pi/4, 0)$, выбранном так для наглядности, и возвращается на Северный полюс с противоположного направления $(\theta, \phi) = (3\pi/4, \pi)$.

Задавшись целью изобразить мировую линию вблизи Южного полюса непрерывным образом, обратим внимание на следующее обстоятельство. Нам кажется, что луч на рис. 42.4 полностью лежит в одной плоскости. Вспомним, однако, что обод соответствует единственной точке. На рис. 42.5 мы «собрали» упомянутую видимую плоскость, придав ей форму, которой она должна обладать в действительности, — форму сферы. По аналогии с большой окружностью на S^2 будем называть такую сферу большой. На этой большой сфере мы получаем, наконец, непрерывное представление рассматриваемого светового сигнала, совершающего кругосветное путешествие. В дальнейшем мы будем часто ее применять.

Мы неоднократно будем использовать описанные выше три частных представления. Все вместе они позволяют нам в какой-то мере наглядно представить себе ситуации в искривленном четырехмерном пространстве-времени.

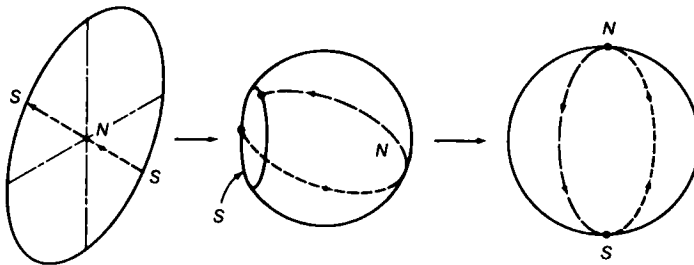


Рис. 42.5

Траектория светового сигнала на большой сфере, которая восстановлена исходя из представления типа двумерного диска.

ЗАДАЧИ

42.1. (11) Предположим, что наблюдается хаббловское разбегание с квадратичной зависимостью красного смещения от расстояния. Исключает ли это фридмановскую Вселенную?

42.2. (30) Рассмотрите задачу изометрического погружения описанных пылевых Вселенных. Исследуйте как евклидово, так и лоренцево пространства.

42.3. (19) Проанализируйте следующие два утверждения:

а) огненный шар, от которого приходит излучение с $T = 3 \text{ К}$, расположен на дуговом расстоянии χ ;

б) огненный шар заполняет всю Вселенную. Объясните, почему высказанные утверждения не противоречат друг другу.

43. Наблюдения в замкнутой пылевой Вселенной

Наверное, ни к одному разделу элементарной математики студенты не питают такого отвращения, как к сферической тригонометрии

Р. Тейт, Британская энциклопедия, 11-е издание

Чтобы решить вопрос о приемлемости модели замкнутой пылевой Вселенной, мы должны посмотреть, соответствуют ли ее предсказания фактическим наблюдениям. Критический анализ космологических моделей проводится в разд. 50. Здесь же мы обсудим некоторые типичные наблюдения и покажем, как соответствующие величины рассчитываются в этой модели.

Одна из наблюдаемых величин, которую мы можем надеяться получить, — это современный возраст Вселенной, измеренный путем исследования горных пород или звезд. Какое собственное время прошло от момента возникновения Вселенной? Мы хотим измерять собственное время, поскольку именно оно отвечает темпу хода ядерных часов. Искомый возраст t_0 связан с параметрами η_0 и a нашей модели следующим образом:

$$t_0 = \frac{a}{2} (\eta_0 - \sin \eta_0). \quad (43.1)$$

Следуя установившейся традиции в космологии, мы применяем

нижний индекс нуль для обозначения настоящего момента времени.

Другой наблюдаемый факт состоит в наличии так называемого хаббловского разбегания, которое обнаруживается по систематическому красному смещению, зависящему от расстояния. Хаббловское разбегание ставит на повестку дня весьма тонкий вопрос о расстояниях. Для близлежащих объектов расстояние можно измерить обычными способами. Пока масштабный фактор $R(t)$ пренебрежимо мало меняется в течение того времени, которое требуется свету для прохождения измеряемого расстояния, собственное расстояние связано с дуговым расстоянием $\Delta\chi$ посредством соотношения

$$d = R(t) \Delta\chi, \quad (43.2)$$

которое вытекает из вида метрики $\mathcal{S}$. Последнее соотношение позволяет дать следующую физическую интерпретацию масштабного фактора R , которую можно считать наилучшей. Этот фактор, имеющий размерность длины, позволяет преобразовать дуговое расстояние в собственное. Дуговое расстояние между галактиками, покоящимися относительно опорной космической системы отсчета, не меняется с течением времени. Именно поэтому дуговое расстояние столь полезно. Собственное же расстояние между такими галактиками меняется, поскольку изменяется масштабный фактор $R(t)$.

Предположим, что мы наблюдаем свет от какой-либо близлежащей галактики ($\Delta\chi \ll 1$) и сравниваем его спектральные линии со спектральными линиями тех же элементов, полученных в лаборатории (рис. 43.1). Галактические спектральные линии смещены в красную сторону на величину

$$(1+z) = \frac{\lambda_{\text{REC}}}{\lambda_{\text{EMIT}}} = \frac{R(t_{\text{REC}})}{R(t_{\text{EMIT}})}. \quad (43.3)$$

Поскольку галактика по предположению расположена недалеко, мы можем применить приближенную формулу

$$\begin{aligned} R(t_{\text{EMIT}}) &\approx R(t_0 - d) \approx \\ &\approx R_0 - dR'_0, \end{aligned} \quad (43.4)$$

где R_0 обозначает современное значение R , так что

$$(1+z) \approx 1 + d \frac{R'_0}{R_0}. \quad (43.5)$$

Расширение

[Я буду использовать обозначение $\Delta\chi$ для дугового расстояния даже в том случае, когда оно не измеряется вдоль направления χ .]

[Красное смещение в расширяющейся Вселенной обсуждалось в разд. 40.]

[Штрих здесь обозначает производную по времени.]

Для небольших расстояний справедлив линейный закон красного смещения

$$z \approx \left(\frac{R'_0}{R_0} \right) d. \quad (43.6)$$

Постоянная Хаббла Коэффициент пропорциональности в последней формуле называется постоянной Хаббла H_0 :

$$H_0 \equiv \frac{R'_0}{R_0}. \quad (43.7)$$

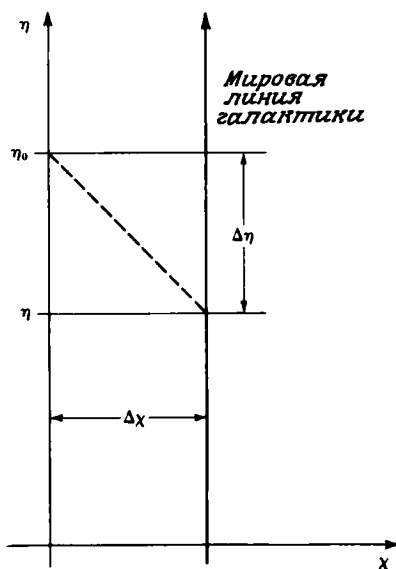


Рис. 43.1

Наблюдение света от близлежащей галактики.

На практике мы можем измерять этот коэффициент, регистрируя красные смещения для тех галактик, которые находятся от нас достаточно близко, чтобы можно было найти расстояния до них, и в то же время достаточно далеко, чтобы они не оказались частью местного скопления галактик. Дело в том, что галактики в одном скоплении не являются свободными частицами, а само скопление не расширяется вместе со Вселенной. Постоянная Хаббла имеет размерность обратной секунды. Однако из-за того способа, с помощью которого измеряется по-

стоянная Хаббла, астрономы используют для нее смешанную единицу измерения километр/(секунда · мегапарсек). Указанные единицы связаны следующим соотношением:

$$50 \text{ км/с} \cdot \text{Мпс} = \frac{1}{6,18 \cdot 10^{17} \text{ с}} = \frac{1}{19,6 \text{ млрд. лет}}. \quad (43.8)$$

Единицы измерения постоянной Хаббла

Напомним, что парсек почти точно равен 10^8 световых секунд. Используя параметры нашей модели η_0 и a , запишем

$$aH_0 = \frac{2 \sin \eta_0}{(1 - \cos \eta_0)^2}. \quad (43.9)$$

Для объектов, находящихся на значительном удалении, масштабный фактор $R(t)$ заметно изменяется за то время, которое требуется свету для прохождения соответствующего расстояния; поэтому для таких объектов не существует никакого практически разумного способа измерения собственного расстояния. Наоборот, дуговое расстояние определяется вполне корректно, но оно не относится к непосредственно измеряемым величинам.

Для многих таких удаленных объектов красное смещение можно точно измерить, и оно представляет собой наиболее приемлемую меру расстояния. Постоянную Хаббла можно рассматривать как параметр, который преобразует малые расстояния, измеренные при помощи безразмерной переменной красного смещения z , в собственные расстояния. При этом для больших расстояний никакого особого смысла величинам $H_0 z$ или $R_0 \Delta x$ приписать нельзя.

Красное смещение

Другая величина, которую, вероятно, можно наблюдать, — это современная плотность материи во Вселенной. Вопрос о том, можем ли мы видеть всю материю, остается открытым, но мы заведомо можем установить нижние пределы плотности массы. Выразим плотность ρ через параметры нашей модели, начав с определения

Плотность материи

$$a \equiv \frac{8\rho\pi R^3}{3}. \quad (43.10)$$

Используя выражение (42.12) для $R(t)$, получим

$$a^2 \rho_0 = \frac{3}{\pi(1 - \cos \eta_0)^3}. \quad (43.11)$$

Напомним, что ρ имеет размерность (секунда) $^{-2}$, так что величина $a^2 \rho$ безразмерна. Для перехода от одних единиц к другим удобно использовать плотность воды, равную $6,67 \cdot 10^{-8} \text{ с}^{-2}$.

Часто для плотности материи употребляют безразмерный параметр плотности Ω_0 , вводимый с помощью постоянной Хаббла:

$$\Omega_0 \equiv \frac{8\pi\rho_0}{3H_0^2}. \quad (43.12)$$

Этот параметр связан с нашим современным дуговым временем соотношением

$$\Omega_0 = \frac{2}{1 + \cos \eta_0}. \quad (43.13)$$

Будучи безразмерным, параметр Ω_0 может зависеть только от η_0 , но не от a . Для рассматриваемых замкнутых пылевых Вселенных значения параметра Ω_0 лежат в области

$$\Omega_0 > 1. \quad (43.14)$$

Имеется еще одна наблюдаемая величина, плодотворность использования которой выявится при изучении других моделей, — это параметр замедления q_0 , определяемый соотношением

$$q_0 \equiv -\frac{R_0 R_0''}{R_0' R_0'}. \quad (43.15)$$

Напомним, что

$$R' \equiv \frac{dR}{dt} \quad (43.16)$$

Поскольку параметр q_0 безразмерен, он должен быть связан с Ω_0 и η_0 . Для пылевых Вселенных такая связь имеет следующий вид:

$$\Omega_0 = 2q_0. \quad (43.17)$$

Мы имеем полное право пользоваться указанными формулами. Приведем несколько примеров постепенно увеличивающейся сложности.

Пример 1

Предположим, что свет от объекта с красным смещением z принимается в момент дугового времени η_0 . Как далеко от нас в пересчете на дуговое расстояние находится этот объект?

Пусть η — момент дугового времени, когда рассматриваемый объект испустил свет, как показано на рис. 43.1. Запишем

соотношение

$$(1+z) = \frac{R(\eta_0)}{R(\eta)} = \frac{1 - \cos \eta_0}{1 - \cos \eta}, \quad (43.18)$$

откуда определяем η . Искомое дуговое расстояние $\Delta\chi$ вычисляется теперь по формуле

$$\Delta\chi = \eta_0 - \eta, \quad (43.19)$$

поскольку в координатах (η, χ) скорость света равна единице.

Пример 2

Мы наблюдаем удаленный квазар с красным смещением $z = 2$. Допустим, что используется модель Вселенной с $q_0 = 1,9$. Чему был равен возраст Вселенной (в предположении, что она описывается замкнутой пылевой моделью) в момент испускания принятого нами света?

Зная q_0 , можно найти современное дуговое время, используя соотношение

$$q_0 = \frac{1}{1 + \cos \eta_0}. \quad (43.20)$$

Таким образом, получаем

$$\eta_0 = 2,06 \text{ рад } (118^\circ). \quad (43.21)$$

Световой сигнал, который распространяется в такой Вселенной с момента ее образования, мог бы пройти около трети всего пути вокруг нее. Зная η_0 , мы можем из соотношения (43.9) найти следующую комбинацию:

$$(aH_0) = 0,81. \quad (43.22)$$

Однако, не располагая какой-либо размерной величиной, мы не можем вычислить время, поэтому оставим в нашем ответе зависимость от неизвестной постоянной Хаббла.

Итак, квазар испускает свет в момент дугового времени, задаваемый соотношением

$$(1+z) = \frac{1 - \cos \eta_0}{1 - \cos \eta}, \quad (43.23)$$

откуда имеем

$$\eta = 1,03 \text{ рад } (59^\circ). \quad (43.24)$$

Собственный возраст Вселенной в момент испускания света

квazarом определяется из формулы

$$TH_0 = \frac{(aH_0)}{2} (\eta - \sin \eta), \quad (43.25)$$

откуда получаем

$$H_0 T = 0,0714, \quad (43.26)$$

или в общепринятых единицах

$$(TH_0) = (1,3 \text{ млрд.лет}) \cdot [50 \text{ км}/(\text{с} \cdot \text{Мпс})]. \quad (43.27)$$

Современный возраст Вселенной в рассматриваемой модели равен

$$T_0 H_0 = (16 \text{ млрд.лет}) \cdot [50 \text{ км}/(\text{с} \cdot \text{Мпс})]. \quad (43.28)$$

Если бы мы взяли для постоянной Хаббла значение порядка 100 км/(с · Мпс), то современный возраст Вселенной получил бы равным лишь 8 млрд.лет. Это слишком короткий промежуток времени, поскольку, как мы полагаем, мы видим звезды, возраст которых достигает 15 млрд.лет.

Пример 3

[Мне не хотелось бы создавать впечатление, что у меня существует предвзятое мнение о квазарах и о том, что все их красное смещение имеет космологическое происхождение. Когда я говорю здесь о квазаре, то имею в виду лишь очень удаленный видимый объект с измеримым космологическим красным смещением.]

Рассмотрим теперь геометрически более сложную ситуацию. Допустим, что наблюдаются два квазара, видимое угловое расстояние между которыми мало, $\epsilon \ll 1$. Предположим далее, что эти квазары удалены от нас в точности на одно и то же расстояние, соответствующее красному смещению z . Как далеко друг от друга находились они в момент испускания принятого нами света?

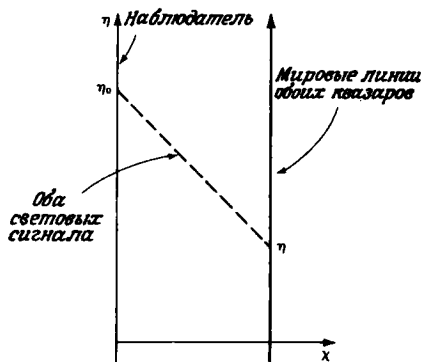


Рис. 43.2

Плоскость (η, χ) не выявляет пространственную геометрию.

Здесь применимы те же самые рассуждения, с помощью которых в предыдущем разделе была получена разумная картина кругосветного распространения светового сигнала. Предположим, что мы находимся на Северном полюсе, $\chi = 0$. Поскольку все точки эквивалентны, это допустимо. В координатах (η, χ) на рис. 43.2 мировые линии обоих квазаров, а также испущенных ими световых сигналов попарно совпадают. На рис. 43.3 мы опускаем координату η и изображаем события в представлении S^3 типа трехмерного шара. Наблюдатель на этом рисунке находится в центре шара. Оба квазара расположены на радиусе $\chi = \eta_0 - \eta$, где η — дуговое время, отвечающее моменту испускания ими света. Мы выбрали координаты таким образом, что оба квазара лежат на большой сфере $\phi = 0$. Эта большая сфера на рис. 43.3 кажется плоскостью, проходящей через центр шара; она изображена отдельно на рис. 43.4. Угол ϵ между лучами света хорошо виден как на рис. 43.3, так и на рис. 43.4. Таким образом, два рассматриваемых квазара имеют следующие координаты:

$$(\chi, \theta, \phi) = (\eta_0 - \eta, 0, 0), \quad (43.29)$$

$$(\chi, \theta, \phi) = (\eta_0 - \eta, \epsilon, 0). \quad (43.30)$$

Их координаты различаются как раз на величину

$$\Delta\theta = \epsilon. \quad (43.31)$$

Используя метрику

$$\mathcal{G} = R^2(\eta) [-d\eta^2 + d\chi^2 + \sin^2 \chi (d\theta^2 + \sin^2 \theta d\phi^2)], \quad (43.32)$$

можно преобразовать упомянутую разность координат в собственное расстояние

$$\delta = R(\eta) (\sin \chi) \epsilon. \quad (43.33)$$

Последнюю формулу можно записать в следующем виде:

$$(H_0 \delta) = \epsilon \frac{(H_0 a)}{2} (1 - \cos \eta) \sin (\eta_0 - \eta). \quad (43.34)$$

Для Вселенной с $q_0 = 1,9$ и разности между координатами порядка $2''$ формула (43.34) дает следующее значение длины:

$$(H_0 \delta) = (1,05 \cdot 10^{12} \text{ с}) \cdot [50 \text{ км/с} \cdot \text{Мпс}] \quad (43.35)$$

или

[Менее искусственная родственная задача заключалась бы в рассмотрении вопроса о том, насколько близко свет от одного квазара пройдет мимо расположенного перед ним другого квазара, если расстояние между ними велико.]

[Последовательность рисунков 43.2 — 43.4 соответствует последовательности рисунков 42.3 — 42.5.]

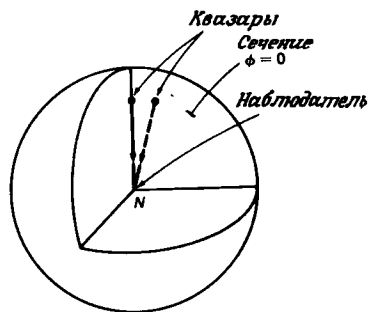


Рис. 43.3

Пространственная геометрия проявляется в представлении типа трехмерного шара.

[Угловая секунда соответствует самым мелким разрешимым оптическим деталям, видимым через атмосферу Земли, и потому является удобной единицей. Перевод ее в радианы: $1'' = 4,85 \cdot 10^{-6} \text{ рад.}$]

[Килопарсек, равный 10^{11} световых секунд, представляет собой удобную единицу длины, поскольку он соответствует размеру типичных особенностей в галактике. Радиус нашей Галактики составляет примерно 10 кпс.]

$$(H_0 \delta) = (10,5 \text{ кпс}) \cdot [50 \text{ км}/(\text{с} \cdot \text{Мпс})]. \quad (43.36)$$

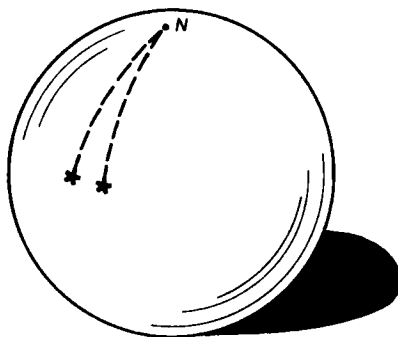


Рис. 43.4
Пространственная геометрия на большой сфере.

ЗАДАЧИ

43.1. (16) На каком расстоянии друг от друга расположены галактики в замкнутой пылевой Вселенной с $H_0 = 50 \text{ км}/(\text{с} \cdot \text{Мпс})$ и $\Omega = 2$?

43.2. (26) В замкнутой пылевой Вселенной плотность галактик задается величиной μ (галактик/с³). Выразите полное число галактик в такой Вселенной через H_0 и η_0 .

43.3. (22) Какова интерпретация величины $(R_0 H_0)$?

43.4. (12) Чему равна современная плотность в с⁻² и г/см³ для замкнутой пылевой Вселенной с $q_0 = 1,9$ и постоянной Хаббла порядка $50 \text{ км}/(\text{с} \cdot \text{Мпс})$?

43.5. (27) Предположим, что наблюдается некоторое скопление объектов, расположенных на одном и том же расстоянии, например совокупность галактик в одном скоплении. Для каждого объекта получено красное смещение z_i . Значения z_i различны, поскольку объекты в скоплении движутся относительно друг друга. Как оценить дисперсию скоростей таких объектов по наблюдаемой дисперсии их красных смещений? Предположите, что малой величиной является разность $z_i - \langle z_i \rangle$, но не сами z_i . (Угловые скобки обозначают среднюю величину.)

44. Псевдосфера

*Подозреваю, что если бы вы
шли до края света, то вы на-
шли бы там кого-нибудь,
кто идет дальше.*

Г. Д. Торо

Сделаем теперь перерыв и займемся изучением псевдосферы. Это симметричное пространство послужит основой для построения другой совокупности космологических моделей. Рассмотрение здесь будет идти параллельно рассмотрению 3-сферы в разд. 38 и 39. Глобальная структура псевдосферы проще, чем 3-сферы. Ее метрическая структура также не сложнее, хотя и более необычна.

Псевдосфера любого числа измерений может быть получена как подпространство с лоренцевой (а не евклидовой) метрикой, т.е. такое пространство, в котором одна из координат имеет времениподобный характер. Для 2-псевдосферы берем 3-пространство (w, x, y) с метрикой

$$\mathcal{G} = -dw^2 + dx^2 + dy^2 \quad (44.1)$$

и рассматриваем подмножество точек, расположенных на единичном времениподобном интервале от начала координат,

$$w^2 - x^2 - y^2 = 1 \quad (44.2)$$

в положительном направлении оси w . Мы полагаем, что такое подмножество должно быть однородным и изотропным, поскольку оно инвариантно относительно трех изометрий пространства (w, x, y) , а именно поворота

$$x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \quad (44.3)$$

и двух преобразований Лоренца

$$x \frac{\partial}{\partial w} + w \frac{\partial}{\partial x}, \quad (44.4)$$

$$y \frac{\partial}{\partial w} + w \frac{\partial}{\partial y}. \quad (44.5)$$

Каждая точка на этой 2-поверхности соответствует единичному вектору скорости, аналогичному нашим векторам 4-скорости в пространстве-времени. Можно предположить, что

Погружение

[Будьте внимательны! Рассматриваемая лоренцева метрика является фиктивной, так же как и евклидова метрика в разд. 38. Она не имеет ничего общего с пространством-временем.]

[Мы употребляем здесь w вместо t , чтобы избежать какой бы то ни было случайной связи с пространством-временем.]

Симметрии

[Было бы последовательно называть эти преобразования псевдоповоротами, но они уже названы преобразованиями Лоренца.]

Пространство скоростей

естественная мера скоростей — быстрота — будет в данном случае полезной координатой, и так оно и есть на самом деле. Естественная параметризация для этой поверхности задается в координатах (χ, ϕ) следующим образом:

$$\begin{aligned} w &= \operatorname{ch} \chi, \\ x &= \operatorname{sh} \chi \cos \phi, \\ y &= \operatorname{sh} \chi \sin \phi. \end{aligned} \quad (44.6)$$

Здесь χ представляет собой быстроту прямой мировой линии, выходящей из начала координат в точку, задаваемую приведенными выше соотношениями.

Топология

Глобальная структура в данном случае очевидна. Имеется шов, на котором точки с $\phi = 0$ и $\phi = 2\pi$ должны быть соединены. Быстрота χ может принимать любые положительные значения. Когда $\chi = 0$, мы имеем единственную точку, являющуюся аналогом Северного полюса на сфере. Глобальная структура здесь та же самая, что и для полярных координат. На рис. 44.1 изображено (χ, ϕ) -представление. Оно также может быть склеено таким образом, что получится аналог представления типа двумерного диска.

Метрика

Поскольку эта псевдосфера является подпространством большего многообразия, она наследует от него метрику

$$\mathcal{G} = d\chi^2 + \operatorname{sh}^2 \chi d\phi^2. \quad (44.7)$$

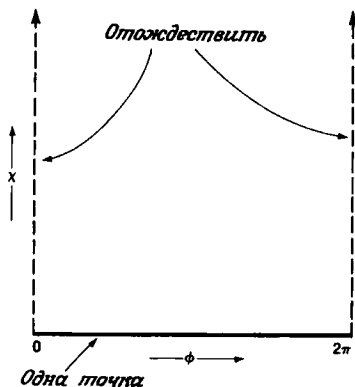


Рис. 44.1
2-псевдосфера.

Доказательство

Выписанная здесь метрика получена из параметрических соотношений (44.6). Ограничение на 2-пространство базисных 1-форм, заданных на 3-пространстве, записывается следующим образом:

$$\begin{aligned} dw^* &= \operatorname{sh} \chi \, d\chi, \\ dx^* &= \operatorname{ch} \chi \operatorname{ch} \phi \, d\chi - \operatorname{sh} \chi \sin \phi \, d\phi, \\ dy^* &= \operatorname{ch} \chi \sin \phi \, d\chi + \operatorname{sh} \chi \cos \phi \, d\phi. \end{aligned} \quad (44.8)$$

Метрика (44.7) получается после подстановки соотношений (44.8) в выражение (44.1).

Такая метрика является положительно определенной. Не существует векторов, имеющих нулевую или отрицательную длину.

Пример

Рассмотрим метрику в точке $(\chi, \phi) = (\alpha, 0)$, где

$$\begin{aligned} \operatorname{sh} \alpha &= 1, \\ \alpha &\approx 0,89. \end{aligned} \quad (44.9)$$

Метрика в этой точке имеет вид

$$d\chi^2 + d\phi^2, \quad (44.10)$$

а метрическая фигура представляет собой окружность. Очевидно, эта метрика евклидова, а не лоренцева. Поскольку пространство однородно, сказанное справедливо в каждой точке.

Для псевдосфер, как и для сфер, существуют стереографические представления. Одна из таких проекций, задающая карту для 2-псевдосферы в (x, y) -плоскости, схематически изображена на рис. 44.2. Одна такая диаграмма покрывает псевдосферу целиком, и ее достаточно для того, чтобы показать, что псевдосфера представляет собой многообразие. Такие стереографические представления могут иметь кое-какие интересные свойства. Точки, разнесенные бесконечно далеко на псевдосфере, могут оказаться на конечном расстоянии в проекции. Отображения могут быть конформными, т.е. сохраняющими углы. Кроме того, некоторые из них могут переводить конечные окружности в конечные же окружности.

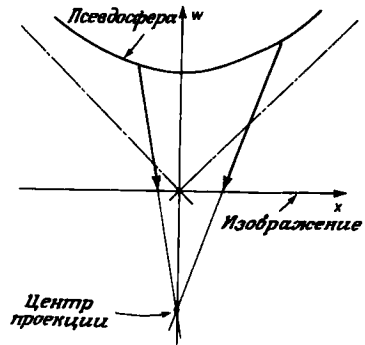


Рис. 44.2
Отображение псевдосферы на плоскость (x, y) .

Стереографические представления

[См. задачу 33.3.]

Пример

Проецирование на плоскость $w = 0$ из точки $(w, x, y) = (-1, 0, 0)$ обладает всеми перечисленными выше свойствами. Именно такое отображение представлено на рис. 44.2. Если во избежание путаницы координаты в (x, y) -плоскости назвать (u, v) , то получим

$$u = \frac{x}{w+1}, \quad (44.11)$$

$$v = \frac{y}{w+1}, \quad (44.12)$$

$$w^2 = 1 + x^2 + y^2. \quad (44.13)$$

Эти уравнения задают координаты (w, x, y) как функции от u и v . Решая их, получаем

$$x = \frac{2u}{\Delta}, \quad (44.14)$$

$$y = \frac{2v}{\Delta}, \quad (44.15)$$

$$w = \frac{1 + u^2 + v^2}{\Delta}, \quad (44.16)$$

где

$$\Delta \equiv 1 - u^2 - v^2. \quad (44.17)$$

[Операция ограничения на подпространство рассмотрена в разд. 20.]

С помощью выписанных соотношений мы можем выяснить, каким образом ограничиваются 1-формы dw , dx и dy . Например, имеем

$$dx^* = \frac{2}{\Delta^2} [(1 + u^2 - v^2) du + 2uv dv]. \quad (44.18)$$

В силу осевой симметрии достаточно вычислить dx^* , dy^* и dw^* вдоль линии $v = 0$. Для этой линии мы получаем

$$dx^* = \frac{2(1 + u^2)}{\Delta^2} du, \quad (44.19)$$

$$dy^* = \frac{2\Delta}{\Delta^2} dv, \quad (44.20)$$

$$dw^* = \frac{4u}{\Delta^2} du. \quad (44.21)$$

Метрику теперь можно записать в следующем виде:

$$-dw^2 + dx^2 + dy^2 = \frac{4(du^2 + dv^2)}{\Delta^2}. \quad (44.22)$$

В этом виде метрика заведомо симметрична; таким образом, мы получили правильное выражение для метрики, справедливое всюду.

Полная псевдосфера отображается на внутренность единичной окружности на (u, v) -плоскости. Такое отображение отличается от модели двумерного диска для S^2 тем, что оно не включает обод. В самом деле, расстояние от любой точки на псевдосфере до обода было бы бесконечным. Как и в случае сфер, плоскости, проходящие через начало координат в (w, x, y) -пространстве, пересекают псевдосферу по геодезическим. Осознав этот факт, мы можем без труда получить, что в описанном выше (u, v) -представлении геодезические являются дугами, пересекающими единичную окружность под прямыми углами. В рассматриваемом представлении легко обнаруживается неевклидов характер геометрии псевдосферы. На рис. 44.3 изображены две различные линии A и B , ни одна из которых не пересекает третью линию L . В евклидовой геометрии существует лишь одна линия, проходящая через данную точку и не пересекающая другую заданную линию. На рис. 44.4 изображен треугольник ABC ненулевой площади, все углы при вершинах которого равны нулю. Обсуждаемое представление позволяет оперировать с псевдосферой, пользуясь методами евклидовой

[Если вам не нравится такой кратчайший путь рассуждений, то проделайте соответствующие длинные выкладки.]

[Другим представлением 2-псевдосферы является полуплоскость Пуанкаре, рассмотренная в задачах 34.4 и 34.5. Дополнительно об этих представлениях можно прочитать в книге [34, с. 184] и [5] (приложение 20).]

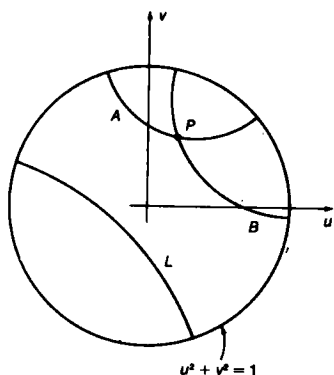


Рис. 44.3.
Две проходящие через P различные линии, которые не пересекают линию L .

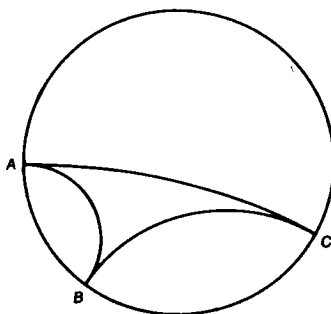


Рис. 44.4

геометрии. Поработайте с ним, чтобы развить свою интуицию. Рисунки Эшера под названием «Предел на круге» — это не что иное, как псевдосферы, полученные с помощью такого представления.

Непосредственным обобщением рассмотренной 2-псевдосферы является 3-псевдосфера. В 4-пространстве с лоренцевой метрикой она может быть представлена следующим множеством точек:

$$w^2 - x^2 - y^2 - z^2 = 1. \quad (44.23)$$

В параметрическом виде это множество задается уравнениями

$$\begin{aligned} w &= \operatorname{ch} \chi, \\ z &= \operatorname{sh} \chi \cos \theta, \\ x &= \operatorname{sh} \chi \sin \theta \cos \phi, \\ y &= \operatorname{sh} \chi \sin \theta \sin \phi, \end{aligned} \quad (44.24)$$

а наследуемая им метрика имеет вид

$$\mathcal{G} = d\chi^2 + \operatorname{sh}^2 \chi d\Omega^2. \quad (44.25)$$

Глобальная структура в координатах (χ, θ, ϕ) имеет тот же вид, что и в обычных сферических полярных координатах (r, θ, ϕ) .

Пример

Множество всех 4-скоростей в пространстве-времени Минковского представляет собой 3-многообразие. Если расстояние в таком многообразии определяется по относительной скорости, то соответствующее пространство скоростей обладает геометрией псевдосферы. Координата χ отвечает быстрой ψ , введенной в разд. 9.

ЗАДАЧИ

44.1. (19) Конечна или бесконечна площадь треугольника, изображенного на рис. 44.4?

44.2. (30) Рассмотрите уравнения для отображения, представленного на рис. 44.2. Опишите получающееся изображение. Является ли оно конформным?

44.3. (25) В координатах (χ, ϕ) получите три описанных в тексте векторных поля, которые порождают преобразования изотропии и однородности на поверхности 2-псевдосферы.

44.4. (28) Покажите, что дуговое расстояние χ между двумя точками на псевдосфере задается соотношением

$$\operatorname{ch} \chi = -A \cdot B,$$

где A и B представляют собой 4-векторы в фиктивном лоренцевом 4-пространстве. Воспользуйтесь симметриями псевдосферы, чтобы упростить доказательство.

44.5. (35) Используя аргументы, аналогичные примененным в задаче 12.8, покажите, что угол θ с вершиной B и сторонами BA и BC задается соотношением

$$\cos \theta = \frac{(A \cdot C) + (B \cdot A)(B \cdot C)}{[(B \cdot A)^2 - 1]^{1/2} [(B \cdot C)^2 - 1]^{1/2}},$$

где

$$A \cdot A = B \cdot B = C \cdot C = -1.$$

44.6. (26) Используя приведенный выше результат, найдите аналог закона косинусов для псевдосферических треугольников

$$\operatorname{ch} a = \operatorname{ch} b \operatorname{ch} c - \operatorname{sh} b \operatorname{sh} c \cos A,$$

где a , b и c — три дуговые длины сторон, а A — величина угла, противоположного стороне a .

44.7. (28) Используя результаты предыдущих задач, найдите аналог закона синусов

$$\frac{\sin A}{\operatorname{sh} a} = \frac{\sin B}{\operatorname{sh} b} = \frac{\sin C}{\operatorname{sh} c}.$$

44.8. (24) Покажите, как угадать закон косинусов для углов

$$\cos A = -\cos B \cos C + \sin B \sin C \operatorname{ch} a$$

из соответствующего закона для сферических треугольников. Что можно сказать относительно знаков?

45. Пылевые и радиационные фридмановские Вселенные

Мы можем решать уравнения Фридмана для пылевой Вселенной также и для двух других случаев: критической Вселенной, пространственные сечения которой представляют собой евкли-

довы 3-пространства, и открытой Вселенной, пространственные сечения которой представляют собой 3-псевдосферы. Тогда как объем Вселенных, сечениями которых являются 3-сферы, в каждый момент времени ограничен (поэтому такие Вселенные и называются замкнутыми), другим двум упомянутым случаям всегда соответствуют бесконечные пространственные сечения. Уравнения Фридмана могут быть также решены для того случая, когда материя, содержащаяся во Вселенной, удовлетворяет уравнению состояния излучения. Мы перечислим здесь свойства всех этих решений, чтобы они также могли служить основой для построения космологических моделей.

Для других моделей пылевых Вселенных, как и раньше, уравнения Фридмана могут быть решены. Во всех этих случаях величина

$$a \equiv \frac{8\pi\rho R^3}{3} \quad (45.1)$$

является интегралом движения. В открытом случае решение содержит гиперболические функции, а критическом — степени. Критический случай иногда называют евклидовой моделью. Такое наименование приводит к недоразумениям, поскольку эта модель описывает все-таки искривленное пространство-время. Иногда ее называют также моделью Эйнштейна — де Ситтера. Рассматриваемые решения приведены в табл. 45.1. В

ТАБЛИЦА 45.1. СВОЙСТВА ПЫЛЕВЫХ ФРИДМАНОВСКИХ ВСЕЛЕННЫХ

Параметр	Замкнутая	Критическая	Открытая
R	$\frac{a}{2}(1 - \cos \eta)$	$\frac{a\eta^2}{4}$	$\frac{a}{2}(\operatorname{ch} \eta - 1)$
t	$\frac{a}{2}(\eta - \sin \eta)$	$\frac{a\eta^3}{12}$	$\frac{a}{2}(\operatorname{sh} \eta - \eta)$
Ha	$\frac{2 \sin \eta}{(1 - \cos \eta)^2}$	$\frac{8}{\eta^3}$	$\frac{2 \operatorname{sh} \eta}{(\operatorname{ch} \eta - 1)^2}$
$\frac{\pi\rho a^2}{3}$	$\frac{1}{(1 - \cos \eta)^3}$	$\frac{8}{\eta^6}$	$\frac{1}{(\operatorname{ch} \eta - 1)^3}$
Ω	$\frac{2}{1 + \cos \eta}$	1	$\frac{2}{1 + \operatorname{ch} \eta}$
q_0	$\frac{1}{1 + \cos \eta}$	$\frac{1}{2}$	$\frac{1}{1 + \operatorname{ch} \eta}$

открытой модели масштабный фактор больше не имеет отношения к такому наглядному понятию, как длина большой окружности, однако он все же задает масштаб, на котором становятся существенными эффекты кривизны.

Как евклидова, так и открытая Вселенные расширяются вечно. Прошлое таких Вселенных заключено лишь в конечном промежутке собственного времени, до которого они пребывали в состоянии, аналогичном состоянию исчезающе малого радиуса в замкнутых Вселенных; однако будущее для них неограниченно. Если мы сможем сделать выбор между открытой и замкнутой моделями, то сумеем разрешить и проблему будущего Вселенной.

Параметр плотности Ω служит естественной мерой эволюции в открытых и замкнутых моделях. Для открытых моделей Ω начинается с единицы и уменьшается до нуля. Для замкнутых он начинается с единицы и увеличивается до бесконечности. Если мы знаем Ω , то можем не только установить, какое сейчас «время», но и определить, в каком типе Вселенной мы находимся. Когда те или иные космологические формулы могут быть записаны через Ω , они обычно бывают справедливыми во всех трех типах моделей, как будет видно из последующих примеров.

Параметр плотности

Критические модели характеризуются условием $\Omega = 1$ в каждый момент времени. Наблюдения не позволяют отличить евклидову Вселенную от очень молодой открытой или замкнутой модели. Критические модели развиваются во времени подобно самим себе и не содержат никакой внутренней меры своей эволюции. Они в еще большей степени симметричны, нежели открытые и замкнутые модели. В табл. 45.1 перечислены важные свойства пылевых фридмановских Вселенных.

Критические модели

Пример

Пусть современный возраст Вселенной задается параметром плотности Ω_0 . Чему был равен параметр плотности в тот момент времени, который соответствует красному смещению z ?

Рассмотрим сначала замкнутые модели, в которых $\Omega > 1$. Для них мы имеем

$$(1+z) = \frac{R_0}{R} = \frac{1 - \cos \eta_0}{1 - \cos \eta}, \quad (45.2)$$

а также

$$\cos \eta = \frac{2}{\Omega} - 1. \quad (45.3)$$

Это приводит к следующему выражению:

$$\Omega = \frac{\Omega_0(1+z)}{1 + \Omega_0 z}. \quad (45.4)$$

Как это часто бывает, то же самое выражение справедливо и для моделей, в которых $\Omega \leq 1$.

Планковская Вселенная

В отличие от ситуации для замкнутых моделей в рассматриваемых случаях масштабный фактор $R(t)$ не остается примерно равным тому размеру, который задается постоянной a , характеризующей данную модель. Напротив, масштабный фактор R увеличивается до бесконечности. Это обстоятельство подсказывает нам решение следующей философской проблемы. Вселенная описывается постоянной a , имеющей размерность длины. Объяснение Вселенной должно включать объяснение величины этой постоянной. Если отдать предпочтение квантовомеханическому описанию, то естественно предположить, что длина, построенная из постоянной Планка, G и c должна играть существенную роль в объяснении происхождения Вселенной. К сожалению, эта длина плохо подходит для размера Вселенной. В наших единицах (при условии $G = c = 1$) мы имеем

$$(\hbar)^{1/2} = 1,35 \cdot 10^{-43} \text{ с}. \quad (45.5)$$

[Обсуждение, приведенное в настоящем абзаце, правильное называть философским, поскольку оно представляет собой попытку узнать что-либо о Вселенной только путем размышлений.]

Замкнутая Вселенная с такими размерами существовала бы лишь в течение указанного промежутка времени. Открытая Вселенная обходит это затруднение, поскольку даже Вселенная с параметром

$$a \approx \sqrt{\hbar} \quad (45.6)$$

в результате эволюции достигает такого состояния, которое не слишком сильно отличается от состояния нашей современной Вселенной¹⁾.

Радикационные Вселенные

Другое простое и разумное уравнение состояния, которое мы собираемся рассмотреть, — это уравнение состояния чи-

¹⁾ Описанный подход к проблеме возникновения Вселенной является общепринятым. Из рассуждений автора неясно, почему замкнутая Вселенная с планковскими размерами рождается именно в состоянии максимального расширения, а ведь только в этом случае справедлив вывод о том, что время жизни такой Вселенной порядка планковского. Чаше, напротив, высказывается точка зрения, согласно которой квантовое рождение Вселенной, если оно и получит когда-либо корректное описание, то, по-видимому, как рождение замкнутого мира (см., например, *Зельдович Я. Б.* — *Природа*, 1983, № 9, с. 11—24, Письма в астрономический журнал, 1981, т. 7, № 10, с. 579—581). — *Прим. перев.*

стого излучения

$$p = \frac{1}{3}\rho. \quad (45.7)$$

Если применять эту модель к современной Вселенной, мы должны предположить, что последняя содержит ничтожное количество вещества, включая галактики. При этом само излучение, для того чтобы доминировать в материальном содержимом Вселенной, должно было бы присутствовать в той или иной до сих пор незарегистрированной форме¹⁾. Излучение в ящике оказывает давление на стенки и тем самым совершает работу, если ящик расширяется. Таким образом, излучение в расширяющемся ящике теряет энергию. При медленном расширении сохраняется не энергия, а число фотонов. Поскольку длина волны фотонов увеличивается пропорционально $R(t)$, их энергия ($E = h\nu$) должна уменьшаться как $1/R$. В этом случае для излучения в ящике справедлив следующий закон сохранения:

$$\rho R^4 = \text{const}. \quad (45.8)$$

Такой закон сохранения справедлив также и в расширяющейся Вселенной. Он автоматически содержится в уравнениях Фридмана, как и закон сохранения вещества в пылевой Вселенной.

Доказательство

Уравнения Фридмана для радиационно-доминированной Вселенной имеют следующий вид:

$$2\frac{R''}{R} + \left(\frac{R'}{R}\right)^2 + \frac{\kappa}{R^2} + 8\pi p = 0, \quad (45.9)$$

$$\left(\frac{R'}{R}\right)^2 + \frac{\kappa}{R^2} = \frac{8\pi\rho}{3}. \quad (45.10)$$

Объединив первое и второе уравнения, можно исключить член, описывающий материю; в результате получается простое уравнение второго порядка

$$\frac{R''}{R} + \left(\frac{R'}{R}\right)^2 + \frac{\kappa}{R^2} = 0. \quad (45.11)$$

[Я не знаю элементарного доказательства этого важного результата.]

Закон сохранения для излучения

[В разд. 48 мы увидим, что уравнения Фридмана всегда содержат закон сохранения.]

¹⁾ Наличие значительного количества «дополнительных» безмассовых реликтовых частиц повлияло бы на темп эволюции Вселенной в период космологического нуклеосинтеза, что в свою очередь привело бы к противоречащему наблюдениям соотношению между содержанием водорода и гелия. По этой причине плотность энергии безмассовых частиц в современную эпоху (включая всевозможные незарегистрированные виды) должна быть весьма мала по сравнению с плотностью энергии нерелятивистского вещества. — *Прим. ред.*

Продифференцировав величину, являющуюся по нашему предположению интегралом движения,

$$\frac{d}{dt} \left(\frac{8\pi\rho R^4}{3} \right) = 2R(R')^3 + 2R^2R'R'' + 2\kappa RR', \quad (45.12)$$

убеждаемся, что результат дифференцирования равен нулю, поскольку он представляет собой левую часть уравнения (45.11), умноженную на $2R'R^4$.

Обозначим интеграл движения через a^2 , так что постоянная a снова имеет размерность длины:

$$a^2 \equiv \frac{8\pi\rho R^4}{3}. \quad (45.13)$$

Чтобы найти $R(t)$, достаточно рассмотреть уравнение

$$(R')^2 = \frac{a^2}{R^2} - \kappa. \quad (45.14)$$

Его можно проинтегрировать непосредственно, что дает следующие результаты:

$$R(t) = \begin{cases} \sqrt{t(2a-t)} \\ \sqrt{2at} \\ \sqrt{t(2a+t)} \end{cases} \quad \text{для} \quad \kappa = \begin{cases} +1 \\ 0 \\ -1 \end{cases}. \quad (45.15)$$

Отсюда могут быть выведены все употребляемые в теории параметры моделей. Они приведены в табл. 45.2. Различия в поведении Ω и различия между моделями, связанные с эволюцией Вселенной, имеют такой же вид, как и раньше. Замкнутая радиационная Вселенная существует в течение собственного времени $2a$, причем a имеет смысл максимального радиуса Вселенной. Она существует в течение дугового времени π . Поэтому за время жизни радиационной Вселенной световой сигнал успевает пройти вокруг нее лишь половину всего пути.

Излучение в ранней Вселенной

[Более подробно вопрос об энергии современного содержимого Вселенной рассматривается в гл. 10 книги [25].]

Даже если мы не делаем вынужденного предположения, что Вселенная заполнена необнаруженным излучением, нам все-таки необходимы рассматриваемые радиационно-доминированные модели. Уже из-за наличия микроволнового фонового излучения ранняя Вселенная должна быть радиационно-доминированной. Наблюдаемая сегодня плотность вещества в галактиках примерно в пятьсот раз больше плотности энергии микроволнового фонового излучения. Однако относительная плот-

ТАБЛИЦА 45.2. СВОЙСТВА РАДИАЦИОННЫХ ФРИДМАНОВСКИХ ВСЕЛЕННЫХ

Параметр	Замкнутая	Критическая	Открытая
R	$a \sin \eta$	$a\eta$	$a \cdot \text{sh } \eta$
t	$a(1 - \cos \eta)$	$\frac{a\eta^2}{2}$	$a(\text{ch } \eta - 1)$
Ha	$\frac{\cos \eta}{\sin^2 \eta}$	$\frac{1}{\eta^2}$	$\frac{\text{ch } \eta}{\text{sh}^2 \eta}$
$\frac{8\pi\rho a^2}{3}$	$\frac{1}{\sin^4 \eta}$	$\frac{1}{\eta^4}$	$-\frac{1}{\text{sh}^4 \eta}$
Ω	$\frac{1}{\cos^2 \eta}$	1	$\frac{1}{\text{ch}^2 \eta}$
q_0	$\frac{1}{\cos^2 \eta}$	1	$\frac{1}{\text{ch}^2 \eta}$

ность энергии излучения была большей в прошлом $^{\text{в}}$ $(1+z)$ раз. Таким образом, даже исходя лишь из известного в настоящее время состава материи, мы получаем радиационно-доминированную Вселенную до того момента времени, который отвечает красному смещению $z = 500$. Вдобавок галактики в настоящее время обладают хаотическими скоростями порядка 10^{-3} от скорости света. По мере расширения Вселенной такие хаотические скорости уменьшаются как $1/R$. Итак, до момента времени, отвечающего $z = 1000$, хаотические скорости галактик (если бы, конечно, галактики тогда существовали¹⁾) были бы релятивистскими и сами по себе. Мы снова имели бы дело с материей, подчиняющейся уравнению состояния $p = \frac{1}{3}\rho$.

Отдельные столбцы в табл. 45.1 и 45.2 в действительности не являются независимыми. Их можно свести к единой группе формул при помощи системы обозначений, использующей комплексные числа. Такая система обозначений удобнее в общих расчетах, тогда как формулы в табл. 41.1 и 41.2 более удобны в конкретных случаях. Введем другой указатель типа модели, δ , аналогичный κ , так что

$$\delta = \begin{cases} +1 \\ 0 \\ i \end{cases} \text{ для } \begin{cases} \text{открытой} \\ \text{критической} \\ \text{замкнутой} \end{cases} \text{ модели.} \quad (45.16)$$

[См. задачу 45.12.]

«Универсальные» формулы

[Единственная причина, по которой я ввел в рассмотрение величину κ , состоит в том, что она употребляется очень часто. Ясно, что с тем же успехом можно использовать и δ , поскольку $\delta^2 = -\kappa$.]

¹⁾ По современным представлениям галактики должны были образоваться при $z \leq 10$. — Прим. ред.

ТАБЛИЦА 45.3. СВОЙСТВА ПРОИЗВОЛЬНОЙ ФРИДМАНОВСКОЙ ВСЕЛЕННОЙ

Параметр	Пыль	Излучение
$\frac{R}{a}$	$\frac{1}{\delta^2} \operatorname{sh}^2\left(\frac{\delta\eta}{2}\right)$	$\frac{\operatorname{sh} \delta\eta}{\delta}$
$\frac{t}{a}$	$\frac{1}{2\delta^3} (\operatorname{sh} \delta\eta - \delta\eta)$	$\frac{2 \operatorname{sh}^2\left(\frac{\delta\eta}{2}\right)}{\delta^2}$
(aH)	$\frac{\delta^3 \operatorname{sh} \delta\eta}{2 \operatorname{sh}^4\left(\frac{\delta\eta}{2}\right)}$	$\frac{\delta^2 \operatorname{ch} \delta\eta}{\operatorname{sh}^2 \delta\eta}$
$\frac{8\pi\rho a^2}{3}$	$\frac{\delta^6}{\operatorname{sh}^6\left(\frac{\delta\eta}{2}\right)}$	$\frac{\delta^4}{\operatorname{sh}^4(\delta\eta)}$
Ω	$\frac{2}{1 + \operatorname{ch} \delta\eta}$	$\frac{1}{\operatorname{ch} \delta\eta}$
q_0	$\frac{1}{1 + \operatorname{ch} \delta\eta}$	$\frac{1}{\operatorname{ch} \delta\eta}$

Формулы в этих комплексных обозначениях приведены в табл. 45.3. Заметим, что множители δ входят в формулы таким образом, что предел $\delta \rightarrow 0$ всегда существует и дает правильное выражение для критических моделей.

ЗАДАЧИ

45.1. (20) Как в общем виде получить параметр плотности Ω для фридмановской Вселенной из масштабного фактора и плотности?

45.2. (12) Как получить η_0 для пылевой Вселенной, если заданы ее возраст и постоянная Хаббла? Что можно сказать относительно радиационной Вселенной?

45.3. (17) Рип ван Уинкл ложится спать и заводит свой будильник на то время, когда температура микроволнового фонового излучения упадет до 2 К. Вычислите, как долго он будет спать, для следующих трех моделей Вселенной с $H_0 = 50 \text{ км/с} \cdot \text{Мпс}$:

а) для пылевой модели, $\Omega = 0,05$;

- б) для пылевой модели, $\Omega = 2$;
 в) для радиационной модели, $\Omega = 1$.

45.4. (28) Во Вселенной существуют масштабы длины, отвечающие галактикам ($2 \cdot 10^{12}$ с), расстояниям между галактиками ($3 \cdot 10^{14}$ с) и скоплениям галактик (10^{15} с). Преобразуйте эти расстояния в дуговые расстояния для упомянутых выше трех моделей.

45.5. (17) Чему равны наблюдаемые угловые размеры, соответствующие приведенным выше масштабам длины, при $z = 0, 1$; 1 и 10 в пылевой Вселенной с $\Omega = 2$?

45.6. (28) Предположим, что Вселенная содержит некоторые постоянно существующие объекты типа галактик. Плотность числа таких объектов μ как функция красного смещения представляет собой число объектов Δn , обнаруженных на всем небе в интервале красных смещений от z до $z + \Delta z$, т.е.

$$\mu \Delta z = \Delta n.$$

Чему равна плотность числа рассматриваемых объектов в пылевых и радиационных моделях?

45.7. (33) Пространство-время

$$\mathcal{G} = -dt^2 + t^2(d\chi^2 + \text{sh}^2 \chi d\Omega^2)$$

получено из пространства-времени открытой пылевой Вселенной в результате предельного перехода

$$a \rightarrow 0,$$

$$\eta \rightarrow \infty.$$

Покажите, как совершить такой предельный переход, и докажете, что рассматриваемое пространство-время описывает пустую Вселенную. Для пустой Вселенной пространство-время должно быть пространством-временем Минковского. Найдите преобразование к обычным координатам.

45.8. (30) Сделайте то же, что и в задаче 45.7, для радиационной Вселенной.

45.9. (26) Галактика, лежащая на луче зрения к более удаленному квазару, создает в спектре квазара узкие линии поглощения. Предполагая, что красное смещение z галактики мало, рассчитайте скорость изменения z с течением нашего собственного времени и выразите ее через H_0 и q_0 .

45.10. (35) Используя дисперсионное уравнение для нерелятивистских частиц, а также то обстоятельство, что длина волнового пакета должна увеличиваться по мере расширения Вселенной, покажите, что свободная частица, движущаяся относительно Вселенной, замедляется по следующему закону:

$$v \propto \frac{1}{R}.$$

45.11. (20) Хаотические движения галактик, наблюдаемые в настоящее время, характеризуются скоростями порядка 10^{-3} от скорости света. Когда эти скорости были сравнимы со скоростью света? Дайте объяснение.

45.12. (35) Решите задачу 45.10, не ограничиваясь рассмотрением одних лишь нерелятивистских частиц. Проверьте, что случай излучения надлежащим образом входит в Ваши рассуждения.

45.13. (30) Какой бы вывод Вы сделали, если бы некоторое достоверное наблюдение привело бы к заключению, что величина q_0 лежит в пределах $1,001 \pm 0,002$?

46. Диаграмма Хаббла

Расширяющееся искривленное пространство-время изменяет не только частоту света, но и его интенсивность. С первых дней существования общей теории относительности астрономы надеются использовать этот эффект, чтобы выяснить, в каком типе Вселенной и на какой стадии ее эволюции мы живем. Другими словами, они хотят измерить q_0 . На практике можно наблюдать лишь кажущуюся яркость далеких объектов и их красные смещения. Соотношение между этими величинами называется диаграммой Хаббла. Каждая космологическая модель предсказывает тот или иной специфический вид диаграммы Хаббла, который мы сейчас рассчитаем.

Наблюдения

В нашем распоряжении имеются данные о красном смещении z и спектральном распределении энергии света, принятого от целого ряда объектов с разными красными смещениями. Спектральное распределение энергии F_{ν_0} представляет собой плотность потока энергии, зарегистрированную приемником и характеризующуюся обычной единицей измерения ватт/(метр² · герц). Полный поток энергии вычисляется как интеграл от плотности потока по частотам. В этом случае

мощность определялась бы последующим интегралом от этого потока по заданной совокупной площади. Наконец, энергия была бы интегралом от мощности по времени.

Сам источник характеризуется светимостью L_ν , которая описывает спектральное распределение энергии испущенного света. Узким местом в использовании соотношения светимости — красное смещение для изучения Вселенной является необходимость делать те или иные предположения о собственной светимости L_ν . Это требует глубокого знания рассматриваемых объектов и наличия моделей, описывающих их эволюцию. К примеру, чтобы использовать данные о светимости всей галактики, надо уметь рассчитывать L_ν , исходя из ее кажущейся оптической яркости, возраста и спектра. При этом нам не нужны абсолютные значения, если мы можем выделить группу объектов, характеризующихся одним и тем же значением L_ν . Вычисление собственной светимости L_ν требует использования сложных и еще не вполне устоявшихся астрофизических представлений. В дальнейшем мы не будем их обсуждать, а продолжим рассмотрение геометрических аспектов сформулированной проблемы так, будто мы располагаем данными о совокупности объектов с фиксированным значением L_ν и различными красными смещениями.

Свойства источника

Основной теоретический результат, необходимый для рассмотрения этой проблемы, состоит в том, что в процессе распространения света по расширяющейся Вселенной число фотонов остается постоянным. Этот результат можно получить, обобщая уравнения Максвелла на случай искривленного пространства-времени и анализируя их на основе высокочастотного приближения, использованного в разд. 27.

Рассмотрим совокупность фотонов, испущенных в узком спектральном диапазоне с частотами от ν до $\nu + \Delta\nu$, в пределах которого светимость источника L_ν остается постоянной. Пусть N_ν — число фотонов, испущенных за одну секунду в полосе частот один герц. Тогда имеем

Пучок фотонов

$$L_\nu = h\nu N_\nu, \quad (46.1)$$

где мы воспользовались планковским законом для энергии фотона

$$E = h\nu. \quad (46.2)$$

Почти всегда в космологических рассуждениях полагают, что этот закон справедлив в течение всего времени эволюции Вселенной. Зависимость постоянной Планка h от времени измени-

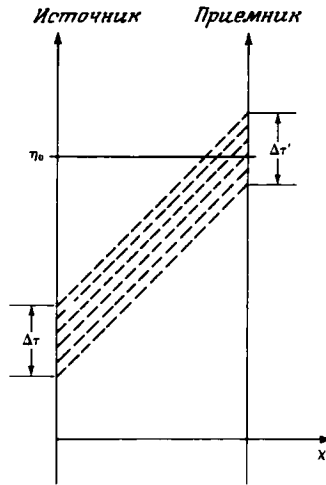


Рис. 46.1

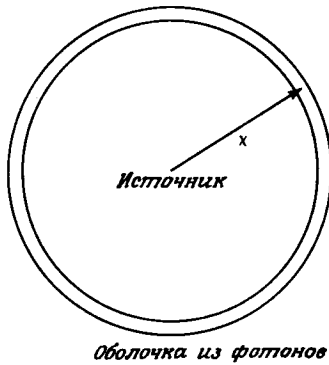


Рис. 46.2
Сечение нашего пространства-времени в момент η_0 , показанный на рис. 46.1. Рассматриваемые фотоны образуют в этом сечении оболочку с координатным радиусом χ .

ла бы некоторые тонкие особенности распределения энергии в спектрах далеких объектов. Поскольку у близких и далеких источников характер спектров один и тот же, за время существования Вселенной могло произойти лишь очень малое изменение h . Пусть, далее, n_{ν_0} означает наблюдаемую скорость поступления фотонов, отнесенную к единичной площади. Тогда справедливо следующее соотношение:

$$F_{\nu_0} = h\nu_0 n_{\nu_0}. \quad (46.3)$$

Из-за космологического красного смещения частота принятого света ν_0 меньше частоты испущенного:

$$\frac{\nu}{\nu_0} = (1 + z). \quad (46.4)$$

Фотоны, испущенные в спектральном диапазоне $\Delta\nu$, принимаются в более узком диапазоне $\Delta\nu_0$, а именно

$$\frac{\Delta\nu}{\Delta\nu_0} = (1 + z). \quad (46.5)$$

Это выражение справедливо при условии, что закон красного смещения линеен.

Рассмотрим теперь фотоны, испущенные за промежуток времени Δt . Они будут приняты за другой промежуток времени Δt_0 , который также подвергнется преобразованию за счет красного смещения:

$$\frac{\Delta t_0}{\Delta t} = (1 + z). \quad (46.6)$$

Таким образом, мы имеем дело с определенным числом фотонов, испущенных за промежуток времени Δt в конкретном спектральном диапазоне шириной $\Delta\nu$, как показано на рис. 46.1.

На рис. 46.2 те же самые фотоны изображены в модели трехмерного шара для нашего пространства-времени. Они распространились на дуговое расстояние χ от источника, который мы поместили на Северном полюсе, $\chi = 0$. Дуговое расстояние задается выражением

$$\chi = \eta_0 - \eta, \quad (46.7)$$

в котором дуговое время η определяется из соотношения

$$(1 + z) = \frac{R(\eta_0)}{R(\eta)}. \quad (46.8)$$

В последнем соотношении должна фигурировать соответствующая функция R . Все рассматриваемые фотоны пересекают 2-сферу, показанную на рис. 46.2. Площадь поверхности 2-сферы вычисляется по формуле

$$A = 4\pi R_0^2 S^2(\chi). \quad (46.9)$$

Число испущенных фотонов должно быть равно числу принятых на 2-сфере. Таким образом, должно выполняться равенство

$$L_\nu \frac{\Delta\nu}{\nu} \Delta t = F_{\nu_0} \frac{\Delta\nu_0}{\nu_0} \Delta t_0 A, \quad (46.10)$$

откуда получается следующее фундаментальное соотношение:

$$F_{\nu_0} = \frac{L_\nu}{4\pi R_0^2 (1+z) S^2(\chi)}. \quad (46.11)$$

Это соотношение устанавливает связь между наблюдаемым потоком и собственной светимостью для источника и наблюдателя, покоящихся в произвольной модели пространства-времени Робертсона — Уокера.

Соотношение (46.11) не очень удобно для практического использования. Действительно, масштабный фактор R_0 совершенно неизвестен; точно так же мы очень мало знаем о величине абсолютной светимости. На графиках обычно изображают логарифм от F_{ν_0} , чтобы превратить упомянутую неопределенность в численном коэффициенте в простой параллельный перенос. Далее, для малых z имеем

Новые переменные

$$S(\chi) \approx \chi \approx \frac{z}{(R_0 H_0)}, \quad (46.12)$$

так что для близких объектов с одной и той же абсолютной светимостью наблюдаемый поток будет отличаться как $1/z^2$. Это обычный закон обратных квадратов. В случае больших z мы анализируем раннюю Вселенную. Дуговое расстояние χ в пределе переходит в η_0 , а поток из-за красного смещения изменяется как $(1+z)^{-1}$. Чтобы отделить эффекты искривленного пространства-времени от указанных известных эффектов, рассмотрим величину

$$\varphi \equiv \frac{1}{2} \lg \left[\frac{z^2 F_{\nu_0}}{(1+z)} \right]. \quad (46.13)$$

При малых z выражение под знаком логарифма ведет себя как $z^2 F_{\nu_0}$, так что оно постоянно в пределе $z \rightarrow 0$. При больших z

оно ведет себя как zF_{ν_0} и тоже остается постоянным. Величина φ (46.13) определена только через наблюдаемые величины. Астроном может вычислить φ , не используя никакой космологической теории. Мы можем предсказать вид соотношения (χ, z) на основе той или иной конкретной космологической модели. Из выражения (46.11) вытекает

$$\varphi = \lg \frac{z}{(1+z)S(\chi)} + \frac{1}{2} \lg \frac{L_\nu}{4\pi R_0^2}, \quad (46.14)$$

причем неизвестные L_ν и R_0 появляются в этом выражении лишь в постоянном члене. Здесь χ в соответствии с соотношениями (46.7) и (46.8) является функцией от z . Зависимость ее от космологической модели заключена в данном случае в параметре η_0 .

Поскольку φ при больших z стремится к постоянному пределу, соответствующие кривые при таких z становятся малонаглядными. Более удобно характеризовать красное смещение естественной переменной

$$\mu \equiv \frac{z}{(1+z)} = \frac{\lambda_0 - \lambda}{\lambda_0}, \quad (46.15)$$

которая изменяется от нуля до единицы в пределах наблюдаемой Вселенной. Мы позволим себе называть график зависимости φ от μ диаграммой Хаббла, хотя сам Хаббл таких диаграмм не рисовал.

На рис. 46.3 представлены некоторые теоретические диаграммы Хаббла. Заметим, что открытая и замкнутая модели непрерывно переходят одна в другую вдоль линии, отвечающей критической модели. Форма этих кривых совсем не зависит от H_0 . Различие значений q_0 проявляется в различном наклоне кривых в каждой точке. Вертикальное расположение кривых произвольно, если мы не знаем L_ν и R_0 .

Использовать такие наблюдения для определения q_0 трудно, поскольку для этого требуется очень глубокое теоретическое понимание собственной светимости L_ν и ее возможной зависимости от z . К тому же свет от далеких объектов отклоняется всеми локальными гравитационными полями во Вселенной, так же как и гладким космологическим полем. Такое мелкомасштабное отклонение может внести систематические изменения в диаграмму Хаббла, а также увеличить ее разброс. Экспериментальная диаграмма Хаббла для квазаров в описанной

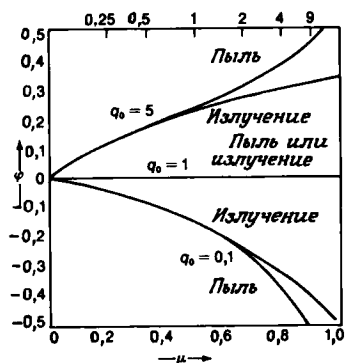


Рис. 46.3

[Поскольку φ - и (μ, φ) -диаграммы недавно введены новой, они еще не получили широкого распространения.]

Перспективы

здесь форме приведена в статье [6], где имеются такие ссылки на некоторые более ранние работы.

ЗАДАЧИ

46.1. (21) Покажите, что

$$\left. \frac{d\phi}{d\mu} \right|_{\mu=0} = \frac{\lg e}{2} (q_0 - 1).$$

46.2. (24) Как выглядит диаграмма Хаббла для пустой Вселенной, обсуждаемой в задачах 45.7 и 45.8?

46.3. (28) Покажите, что поверхностная яркость, определяемая как принятый поток, деленный на телесный угол, не зависит от z .

46.4. (29) Предположим, что для некоторого класса объектов с известными собственными размерами имеются наблюдательные данные относительно их видимых диаметров и красных смещений. Определите, какие из этих данных могут быть непосредственно нанесены на диаграмму Хаббла.

46.5. (28) Покажите, что данные относительно количества подсчитанных объектов, упомянутые в задаче 45.6, не могут быть нанесены на диаграмму Хаббла. Это другое измерение.

46.6. (39) Для случая

$$\Omega \ll 1$$

придумайте более удобную форму диаграммы Хаббла.

46.7. (31) Предположим, что все объекты в задаче 45.6 представляют собой источники радиоволн равной интенсивности. Выразите наблюдаемую плотность числа таких объектов через видимый радиопоток в различных моделях.

47. Ньютонова космология

Некоторые черты общерелятивистской космологии могут быть выявлены в квазиньютоновом подходе. Такой подход используется во многих вводных космологических учебниках, чтобы избежать громоздких рассуждений, необходимых для «честного» описания общерелятивистского пространства-времени. К сожалению, этот квазиньютонов подход во многих случаях приводит к неверным результатам; поэтому его следует использовать с осторожностью.

Поучительно проанализировать причины, по которым этот подход иногда работает. Это позволит не только глубже понять наше более тщательное рассмотрение, но и прольет некоторый свет на пределы применимости нашего приближения, в котором материя во Вселенной трактуется как непрерывная жидкость. Фактическое распределение масс во Вселенной очень далеко от однородного. Плотность массы внутри звезды может в 10^{36} и более раз превосходить плотность массы в межзвездном пространстве. Важный вопрос состоит в том, расширяются ли галактики вместе с Вселенной. Мы предполагаем, что не расширяются, но это сложный вопрос; к настоящему времени выполнено несколько хороших расчетов. В данном разделе мы будем исходить из того, что локальная динамика подчиняется законам Ньютона и не подвержена воздействию расширения Вселенной.

Пылевая Вселенная и пылевое облако

Сравним динамику фридмановской Вселенной с динамикой изолированного облака материи. В обоих случаях будем рассматривать материю как пыль, т.е. как вещество с пренебрежимо малым давлением. Если распределение плотности в облаке однородно и скорость каждой точки облака в любой момент времени пропорциональна радиус-вектору этой точки, то плотность остается однородной в процессе расширения облака или его сжатия под действием собственной силы тяготения.

Доказательство

При произвольном значении радиуса ускорение задается величиной m/r^2 , где m — масса, заключенная внутри сферы этого радиуса. Для однородной плотности мы имеем

$$m = \frac{4}{3}\pi\rho r^3, \quad (47.1)$$

следовательно, ускорение равно

$$\frac{d^2r}{dt^2} = -\left(\frac{4\pi\rho}{3}\right)r, \quad (47.2)$$

т.е. оно линейно возрастает с радиусом. Отсюда получается, что отношение расстояний от центра для любых двух частиц остается постоянным. Таким образом, сфера расширяется и сжимается однородно.

При этих предположениях облако полностью описывается своей полной массой M и радиусом $r(t)$.

Чтобы описать динамику такого пылевого облака, рассмотрим пробную частицу, находящуюся в точности на его поверхности, как показано на рис. 47.1. Рассматриваемая пробная частица ведет себя как небольшой участок поверхности и будет следовать за облаком по мере его сжатия. Таким образом, у нас есть уравнение

$$\frac{d^2 r}{dt^2} = -\frac{M}{r^2}, \quad (47.3)$$

которое имеет смысл закона тяготения Ньютона при $G = 1$. Интеграл движения (закон сохранения энергии) можно получить, умножив уравнение (47.3) на r' :

$$\frac{1}{2} \frac{d}{dt} (r'^2) = \frac{d}{dt} \left(\frac{M}{r} \right). \quad (47.4)$$

Интегрирование последнего уравнения дает

$$r(r')^2 - 2Er = 2M, \quad (47.5)$$

где величина E , представляющая собой постоянную интегрирования, имеет смысл полной энергии облака, т.е. суммы кинетической и отрицательной гравитационной энергии.

Затем мы уподобим рассматриваемое пылевое облако некоторой сферической части расширяющейся фридмановской пылевой Вселенной. Для этого будем сопутствовать веществу, заключенному внутри малой сферы с фиксированным дуговым радиусом $\chi_0 \ll 1$. Собственный радиус равен

$$r = R(t)\chi_0, \quad (47.6)$$

а его динамика задается уравнением, которое получается из уравнения Фридмана

$$R(R')^2 + \kappa R = a \quad (47.7)$$

посредством умножения на величину χ_0^3 :

$$r(r')^2 + (\kappa\chi_0^3)r = (a\chi_0^3). \quad (47.8)$$

Это уравнение имеет такой же вид, как и уравнение (47.5) для ньютоновой сферы, если принять для массы вещества внутри сферы следующее выражение:

$$M = \frac{a\chi_0^3}{2}. \quad (47.9)$$

Коллапс пылевого облака

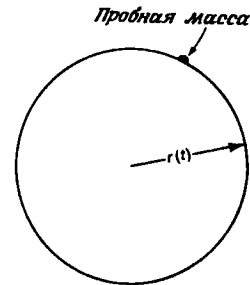


Рис. 47.1

Пылевая Вселенная

Используя определение a , получаем

$$M = \frac{4\pi}{3} \rho (R_{X_0})^3 = \frac{4\pi}{3} \rho r^3; \quad (47.10)$$

[Это верно только для маленькой сферы. Для большой сферы трудно определить полную массу, поскольку собственный радиус и собственная длина большой окружности не связаны евклидовым выражением.]

Влиянием Вселенной можно пренебречь

эта величина действительно представляет собой массу, заключенную внутри малой сферы во фридмановской Вселенной. Постоянная E имеет смысл полной энергии облака и зависит от его размера и начальных условий. Если $E < 0$, то сфера ограничена и может расширяться лишь до некоторого максимального размера, а затем коллапсирует, что соответствует поведению замкнутой Вселенной, т.е. случаю $\kappa = 1$.

Таким образом, мы видим, что динамика малого облака вещества не подвержена воздействию расширяющейся Вселенной, — результат, представляющий собой обобщение ньютонова результата, состоящего в том, что сферически-симметричное распределение масс не влияет на вещество, находящееся внутри. Тогда можно рассмотреть сферическую полость, вырезанную во фридмановской Вселенной. Сформируем из полученной описанным образом материи галактику, содержащую 10^{11} звезд. Динамика такой галактики должна быть ньютоновой, и эта галактика не будет расширяться по мере расширения Вселенной. Хотя приведенный вывод довольно нестрог, на деле его не удастся существенно уточнить. Между прочим, в радиационно-доминированной Вселенной нельзя надеяться вырезать такую дыру, поскольку в этом случае внутри галактики всегда должна быть определенная плотность энергии излучения. Утрата какого-то количества энергии излучения заставила бы галактику слегка расширяться по мере расширения Вселенной.

Чем больше мы размышляем о полученном результате, тем менее очевидным и более удивительным он становится для нас. Можно было бы предположить, что гравитационная потенциальная энергия всей Вселенной в том или ином смысле существует. Однако общая теория относительности часто действует так, будто гравитационное поле не имеет энергии. Это необходимо, чтобы дать конечное описание бесконечной Вселенной.

ЗАДАЧИ

47.1. (19) Завершите рассуждения в тексте и покажите, что однородное распределение плотности сохраняется.

47.2. (36) Предположим, что мы живем во Вселенной с $\Omega = 2$, заполненной излучением. По мере ее расширения масса излу-

ния, заключенного внутри орбиты Земли, изменяется. Насколько изменился бы орбитальный период Земли через миллион лет? Решение поставленной задачи требует более глубоких знаний по классической механике, чем многие из вас имеют. Нужно иметь представление об адиабатических инвариантах.

48. Космология «сделай сам»

В предыдущих космологических моделях использовались нереалистические уравнения состояния, благодаря чему можно было решить получающиеся уравнения. В общем случае нельзя надеяться, что нам удастся найти решение системы обыкновенных дифференциальных уравнений в элементарных или известных специальных функциях. Обычно такая система будет определять новые функции, так что приходится искать численное решение. Используя портативный калькулятор с программным управлением, вы можете без труда построить космологические модели для любого уравнения состояния. Сейчас мы покажем, как это делается. Вы можете сделать для себя лучшую модель Вселенной, чем любая из тех, которые были здесь предложены. К тому же проведение вычислений в организованной форме, необходимой для компьютера, может дать больше для понимания проблемы, чем точные аналитические решения в специальных случаях.

Система уравнений для произвольной фридмановской Вселенной имеет следующий вид:

$$2\frac{R''}{R} + \left(\frac{R'}{R}\right)^2 + \frac{\kappa}{R^2} + 8\pi p = 0, \quad (48.1)$$

$$\left(\frac{R'}{R}\right)^2 + \frac{\kappa}{R^2} = \frac{8\pi\rho}{3}. \quad (48.2)$$

Кроме того, нужно еще третье соотношение для давления p . Последнее может задаваться какой-либо аналитической функцией, таблицей или машинным алгоритмом. Уравнения (48.1) и (48.2) всегда содержат закон сохранения энергии, который, как и раньше, мы можем использовать для понижения порядка одного уравнения. Закон сохранения энергии имеет вид

$$\frac{d}{dt}(\rho R^3) + p \frac{d}{dt}(R^3) = 0. \quad (48.3)$$

[Тем, кто действительно этим заинтересуется, рекомендую книгу [1], содержащую много хорошего материала по численным методам.]

Уравнения Фридмана

Закон сохранения в общем виде

Это обычный закон, который говорит нам о том, что

$$\frac{d}{dt}(\text{Энергия}) = - \text{Давление} \cdot \frac{d}{dt}(\text{Объем}). \quad (48.4)$$

Доказательство

Умножим уравнение (48.2) на R^3 и продифференцируем; в результате получим

$$(R')^3 + 2RR'R'' + \kappa R' = \frac{d}{dt} \left(\frac{8\pi}{3} \rho R^3 \right). \quad (48.5)$$

После умножения уравнения (48.3) на $R^2 R'$ имеем

$$(R')^3 + 2RR'R'' + \kappa R' = -8\pi p R^2 R'. \quad (48.6)$$

Из этих соотношений легко получить закон сохранения.

Наша система уравнений состоит теперь из обыкновенных дифференциальных уравнений первого порядка

Система уравнений первого порядка

$$\frac{dR}{dt} = \left(\frac{8\pi\rho R^2}{3} - \kappa \right)^{1/2}, \quad (48.7)$$

$$\frac{d\rho}{dt} = (-) \frac{3(\rho + p)}{R} \frac{dR}{dt} \quad (48.8)$$

и уравнения для дугового времени

$$\frac{d\eta}{dt} = \frac{1}{R}. \quad (48.9)$$

В ранней Вселенной многое происходит в течение очень короткого интервала собственного времени. Если вместо собственного мы применим дуговое время, то мы сможем растянуть этот интервал. Сравните рис. 42.1 и 42.2. Наши уравнения в дуговом времени имеют следующий вид:

$$\frac{dR}{d\eta} = R \left(\frac{8\pi\rho R^2}{3} - \kappa \right)^{1/2}, \quad (48.10)$$

Уравнения в дуговом времени

$$\frac{d\rho}{d\eta} = \frac{-3(\rho + p)}{R} \frac{dR}{d\eta}, \quad (48.11)$$

$$\frac{dt}{d\eta} = R. \quad (48.12)$$

С помощью этих функций мы можем отвечать на общие космологические вопросы.

Если мы попытаемся интегрировать выписанные уравнения вперед, начиная с момента времени $t = 0$, то столкнемся с массой затруднений. Это связано не только с тем, что значение ρ бесконечно, а значение R равно нулю, но и с трудным вопросом о выборе начальных значений. Вместо этого будем вести расчет назад от современного момента времени. Что мы будем использовать в качестве начальных значений? Масштаб собственного времени сдвигово-инвариантен; следовательно, при желании мы можем современному моменту времени присвоить значение, равное нулю. Различные модели получаются из различных значений ρ_0 — существующей в настоящее время плотности энергии. Эта величина и будет одним из свободных параметров в наших моделях. Другим таким параметром является постоянная Хаббла H_0 . Из уравнения (48.7) имеем

Начальные значения

$$H_0^2 = \frac{8\pi\rho_0}{3} - \frac{\kappa}{R_0^2}, \quad (48.13)$$

поэтому постоянная Хаббла и плотность определяют начальное значение для R . Начальные условия могут быть также просто заданы с помощью параметра плотности

$$\Omega_0 = \frac{8\pi\rho_0}{3H_0^2}. \quad (48.14)$$

Коль скоро есть у нас уравнения и начальные значения, нужен какой-либо алгоритм для решения дифференциальных уравнений. В качестве алгоритма, приемлемого для микрокалькулятора, является схема Рунге — Кутты второго порядка. Для дифференциального уравнения

Численный алгоритм

$$\frac{dy}{dt} = f(y) \quad (48.15)$$

упомянутая схема задает переход от значений в момент t к значениям в момент $t + \Delta t$ в соответствии с формулой

$$y(t + \Delta t) = y(t) + \Delta t f\left\{y(t) + \frac{\Delta t}{2} f[y(t)]\right\}. \quad (48.16)$$

Новое грубое значение y на полшага вперед вычисляется в этой схеме исходя из выражения

$$y(t) + \frac{\Delta t}{2} f[y(t)], \quad (48.17)$$

которое представляет собой первый член ряда Тейлора. Затем в схеме Рунге — Кутты для вычисления u на полном шаге вперед используется производная в точке (48.17). Это эквивалентно взятию вторых производных, но с большим объемом вычислений и меньшими усилиями на программирование. Описанный алгоритм нетрудно распространить на произвольную систему уравнений.

Пример

Посмотрим, как ведет себя смесь пыли и излучения. Будем сравнивать две различные модели: одна из них — чисто пылевая, характеризующаяся современными значениями

$$\begin{aligned} H_0 &= 50 \text{ км/(с} \cdot \text{Мпс)}, \\ \Omega &= 0,043, \end{aligned} \quad (48.18)$$

а другая в настоящее время содержит столько же вещества, сколько излучения:

$$\Omega = 0,086. \quad (48.19)$$

Плотность энергии пылевой Вселенной равна

$$\begin{aligned} \rho &= \frac{3\Omega}{8\pi} H_0^2 = \\ &= 1,35 \cdot 10^{-38} \text{ с}^{-2}. \end{aligned} \quad (48.20)$$

Начальное значение для R получается из выражения (48.13):

$$(R_0 H_0)^2 = \frac{\kappa}{(\Omega - 1)}; \quad (48.21)$$

для чистой пыли

$$R_0 = 6,31 \cdot 10^{17} \text{ с}. \quad (48.22)$$

Чисто пылевая модель позволяет проконтролировать результаты нашего численного расчета, поскольку мы уже знаем соответствующее решение. Нам нужно как-то вычислить давление. Допустим, что пыль и излучение не взаимодействуют. Пусть e означает долю энергии, заключенной в излучении в момент t_0 . Тогда энергия излучения растет как R_0/R по отношению к энергии пыли. Давление будет равно $1/3$ от энергии излучения, т.е.

$$p = \frac{e R_0 \rho}{3(R + e R_0)}, \quad (48.23)$$

где ρ — полная плотность энергии.

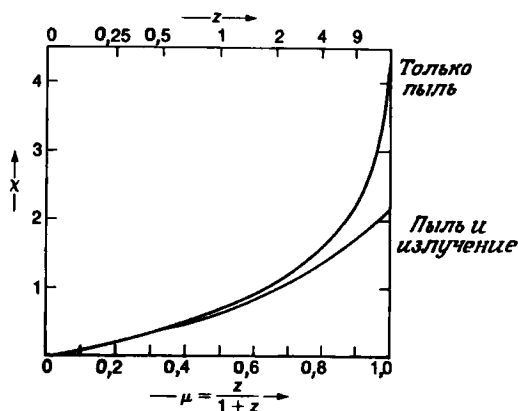


Рис. 48.1

Дуговое расстояние как функция красного смещения для двух моделей Вселенной: пылевой и радиационно-пылевой, в предположении, что пыль и излучение не взаимодействуют между собой.

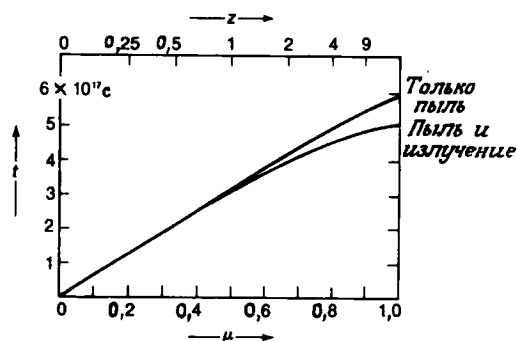


Рис. 48.2

Собственное время между моментом, отвечающим красному смещению z , и настоящим моментом времени для наших двух моделей.

На рис. 48.1 — 48.4 показана эволюция в прошлом обсуждаемых двух моделей. Приводимые зависимости $\chi(z)$, $t(z)$, $P(z)$ и $H(z)$ наиболее общепотребительны. Они определяют и все прочие зависимости. Мы снова использовали переменную $\mu = z/(1+z)$ для лучшего представления ранней Вселенной.

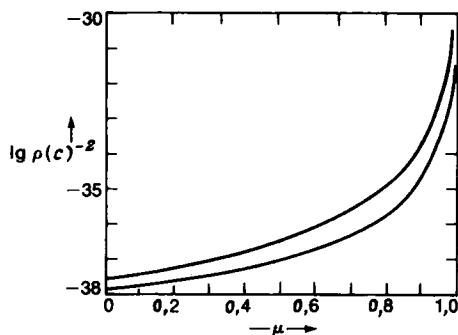


Рис. 48.3

Плотность как функция красного смещения для наших двух моделей.

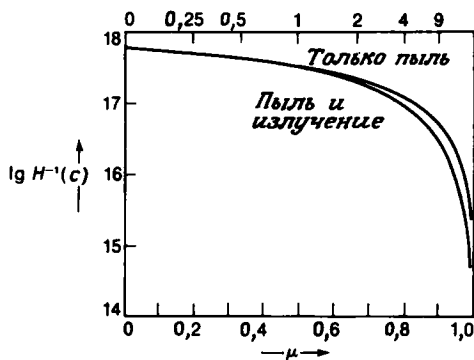


Рис. 48.4

Характерное время расширения (величина, обратная постоянной Хаббла) как функция красного смещения.

Усложнения

Если расширение происходит достаточно быстро, что не позволяет материи прийти в состояние термодинамического равновесия, то давление в реальной Вселенной зависит не только от плотности. В таком случае необходимо применять более детальную теорию материи. В частности, нужны дополнительные обыкновенные дифференциальные уравнения для скоростей реакций и распадов ядер и частиц. Поэтому на достаточно ранней стадии эволюции Вселенной уже нельзя продолжать описанные выше вычисления назад по времени. Чтобы вести рас-

чет из начальной сингулярности и найти соответствующие начальные условия, нужно применить кое-какие искусные приемы.

ЗАДАЧИ

48.1. (12) Почему две кривые на рис. 48.1 совпадают при малых z ? Совпадают ли они точно?

48.2. (10) Объект с известной светимостью наблюдается при $z = 9$ в обоих моделях, показанных на рис. 48.1. Насколько различаются видимые яркости?

48.3. (12) Почему две кривые на рис. 48.2 совпадают при малых z ? Совпадают ли они точно?

48.4. (18) Как нарисовать диаграммы Хаббла для моделей, заданных представленными здесь кривыми?

48.5. (15) Покажите, что параметр плотности Ω в произвольный момент времени может быть рассчитан по формуле

$$\Omega = \left[1 - \frac{3\kappa}{8\pi\rho R^2} \right]^{-1}.$$

48.6. (28) Покажите, что на ранних стадиях космологического расширения, когда материальное содержимое Вселенной ведет себя как чистое излучение, функция

$$\hat{\eta} \equiv \sqrt{2} \left[\frac{|1 - \Omega|}{\Omega} \right]^{1/2}$$

дает оценку дугового времени, прошедшего от Большого Взрыва.

48.7. (12) Покажите, что на ранних стадиях эволюции Вселенной мы имеем

$$z\hat{\eta} = \text{const},$$

где величина $\hat{\eta}$ определена в задаче 48.6.

48.8. (17) Установите связь между величиной $(z\hat{\eta})$ и размером критической радиационной Вселенной, описывающей раннюю эволюцию произвольной модели, содержащей излучение.

48.9. (29) Получите точное выражение $(z\hat{\eta})$ для моделей, содержащих невзаимодействующие пыль и излучение.

48.10. (24) Используйте величину $\hat{\eta}$ в вашей программе, с тем чтобы размер шага ваших вычислений уменьшался по мере приближения к Большому Взрыву.

48.11. (32) Сравните результаты численного интегрирования рассмотренного в тексте примера с результатами численного интегрирования космологической модели, которая получается путем сшивки чисто радиационной и чисто пылевой модели в момент времени, определяемый соотношением

$$\rho_{\text{RAD}} = \rho_{\text{DUST}}.$$

Потребуйте непрерывности R и R' при переходе через линию сшивки.

48.12. (33) Сравните результаты численных интегрирований с результатами для оптимально подогнанных моделей пустых Вселенных.

49. Большой Взрыв

Что можем мы сказать о ранних стадиях эволюции Вселенной? Что можно сказать, сделав предположение (возмутительное!) о полноте и адекватности наших сегодняшних представлений о составе Вселенной? Несмотря на то что было бы безумством воспринимать ответы на такие вопросы слишком серьезно, определенно имеется интересная проблема, достойная обсуждения. Получающийся сценарий называется «стандартной моделью Большого Взрыва».

В современной Вселенной материя, содержащаяся в галактиках, оценивается, по-видимому, величиной порядка

Известная материя

$$\rho_{\text{m0}} \approx 2 \cdot 10^{-31} \text{ г/см}^3. \quad (49.1)$$

Эта масса заключена в медленно движущихся частицах, поэтому она почти целиком обусловлена их энергией покоя. При постоянной Хаббла порядка 50 км/(с · Мпс) это дает параметр плотности

$$\Omega_0 = 0,043. \quad (49.2)$$

Последнее обстоятельство в свою очередь приводит к открытой Вселенной на поздней стадии ее эволюции

Открытая модель

$$\eta_0 \approx 5. \quad (49.3)$$

[Он известен только с точностью до множителя, равного двум. Параметр плотности всего лишь 0,01 также может оказаться совместимым с наблюдениями.]

В настоящее время имеется также небольшое количество материи в форме микроволнового излучения; современная плот-

ность его энергии примерно в 500 раз меньше. Попробуем теперь описать ранние стадии эволюции такой Вселенной в предположении, что фридмановские модели всюду верны.

На ранних этапах излучение доминирует, поскольку для него удовлетворяется соотношение

$$\rho_{\gamma} R^4 = \text{const}, \quad (49.4)$$

в отличие от закона сохранения вещества

$$\rho_m R^3 = \text{const}. \quad (49.5)$$

До момента $z = 500$ Вселенная была радиационно-доминированной. В еще более ранние моменты времени были релятивистскими и сами частицы. Поскольку, однако, на один нуклон сегодня приходится около 10^8 микроволновых фотонов и это соотношение приблизительно сохраняется в процессе эволюции Вселенной, частицы не играют существенной роли.

В окрестности точки $z = 1000$ во Вселенной произошла рекомбинация водорода первичной плазмы и в конце концов образовалась современная материя, состоящая в основном из нейтральных водорода и гелия, загрязненных атомами тяжелых элементов. Упомянутая рекомбинация ионов происходит при температуре около 3000 К. Микроволновое излучение приходит к нам из этой плазмы. Поскольку нейтральный водород прозрачен, излучение распространяется свободно, начиная с $z = 1000$, и сегодня имеет температуру 3 К.

Физика такой ранней Вселенной определяется двумя факторами. Первый — это высокая температура, равная приблизительно

$$T \approx T_0 z, \quad (49.6)$$

где T_0 — современная температура микроволнового фона. Вторым фактором — очень короткий интервал допустимого собственного времени. В критической радиационной Вселенной хаббловское время стремится к нулю по закону

$$\frac{1}{H} = a\eta^2. \quad (49.7)$$

На самой ранней стадии, которую мы можем реально обсуждать, Вселенная характеризовалась температурой около 10^{12} К и состояла из плазмы элементарных частиц: нейтронов, протонов, электронов, позитронов, мюонов, нейтрино, фотонов и т.д. Описание взаимодействий частиц в еще более ранние

Ранняя Вселенная

Первичный огненный шар

[Физика ранней Вселенной очень сложна. К сожалению, я вынужден просто сослаться на соответствующие результаты, поскольку нет возможности вывести их из какой-либо простой теории. Такой неприятный способ представления материала используется только в настоящем разделе.]

Равновесная фаза

моменты времени требует такой физики, которая слишком далеко выходит за пределы возможностей нашего лабораторного эксперимента, чтобы мы могли высказывать какие-либо предположения¹⁾. Даже несмотря на то что характерное время изменения обсуждаемой ранней Вселенной было порядка секунд, скорость ядерных реакций при столь высоких температурах была достаточной, чтобы поддерживать частицы в термодинамическом равновесии. Все частицы присутствовали в количестве, пропорциональном больцмановскому множителю

$$e^{-E/kT}. \quad (49.8)$$

[Энергия, эквивалентная массе протона, соответствует температуре порядка 10^{13} К.]

В ранние моменты времени энергия частиц очень мало зависела от их масс, поскольку все они двигались почти что со скоростью света. Эта была плазма, в которой все типы частиц были равноправны. По мере расширения и охлаждения Вселенной тяжелые частицы замедлялись; в конечном счете их энергия достигала того минимального уровня, который определяется их массой покоя.

При температуре ниже 10^{13} К существовало лишь такое количество тяжелых частиц, которое обеспечивало выполнение законов сохранения, в частности закона сохранения барионного заряда. Протоны и нейтроны при такой температуре продолжали существовать лишь постольку, поскольку материи было больше, чем антиматерии. Благодаря указанной асимметрии Вселенная в настоящее время и содержит материю. Мы можем считать, что эта асимметрия либо существовала изначально, либо возникла в результате неизвестной нам пока зарядово-несимметричной эволюции ранней Вселенной. Между 10^{13} и 10^{10} К почти вся энергия нейтронов и протонов представляет собой энергию покоя. При этом соотношение между числами нейтронов и протонов определяется различием их масс, так что мы имеем

[Разность масс нейтрона и протона составляет 1,29 МэВ.]

$$\frac{n_n}{n_p} = \frac{e^{-M_n/kT}}{e^{-M_p/kT}} = e^{-(M_n - M_p)/kT}. \quad (49.9)$$

При температуре около 10^{10} К взаимодействие между нейтронами и протонами уже не в состоянии обеспечить равновесие между ними, т.е. соотношение (49.9) перестает выполняться. При такой температуре в результате превращения нейтронов в

¹⁾ См. примечание на стр. 307 и предисловие редактора перевода. — *Прим. ред.*

протоны соотношение между ними становится следующим: один нейтрон приходится на каждые пять протонов. При более низких температурах число нейтронов уменьшается преимущественно за счет их распада с временем жизни порядка 1000 с.

В период от 10^{10} до 10^9 К протекало несколько конкурирующих между собой процессов. С одной стороны, из нейтронов и протонов образовывался дейтерий, с другой — дейтерий расщеплялся в результате бомбардировки его фотонами и, наконец, из дейтерия синтезировался гелий. Выше 10^9 К основную роль играет фоторасщепление, но при температуре 10^9 К баланс резко сдвигается и все оставшиеся нейтроны быстро превращаются в гелий. Этот гелий до сих пор должен присутствовать в нашей Вселенной, и очень важно знать, сколько гелия синтезировалось таким образом. Описанная выше картина определялась сложными ядерными реакциями, тогда как мы можем играть лишь роль беспомощных зрителей. Космология используется только для определения того, за какое время Вселенная охладилась от 10^{10} до 10^9 К, поскольку это время определяет количество распавшихся нейтронов. Вычисление этого собственного времени — прекрасное упражнение в космологии. При рассматриваемых температурах Вселенная была не только радиационно-доминированной, но и молодой ($\eta \ll 1$), так что она эволюционировала приблизительно как критическая радиационная Вселенная. Обратимся теперь к Вселенной, переходной между двумя рассматривавшимися до сих пор моделями. Чтобы описать критическую радиационную фазу модели, нам нужны два параметра. Масштабный фактор при красных смещениях $z \gg 1$ задается соотношением

$$R = \frac{R_0}{z}. \quad (49.10)$$

Другим параметром является энергия при таком красном смещении. Эта энергия определяется только сжатым излучением. Для современной плотности вещества имеем

$$\frac{8\pi\rho_{m0}R_0^2}{3} = (R_0H_0)^2\Omega_0 \approx \Omega_0, \quad (49.11)$$

поскольку для Вселенной с низкой плотностью

$$R_0H_0 \approx 1. \quad (49.12)$$

Сказанное относится к поздней фазе эволюции Вселенной. В настоящее время относительная доля излучения, которую мы

Образование гелия

[Вы можете облегченно вздохнуть, подойдя, наконец, к вычислениям, для проведения которых располагаете достаточными навыками.]

обозначим ϵ_0 , определяется соотношением

$$\frac{8\pi\rho_{\gamma 0}R_0^2}{3} = \epsilon_0\Omega_0. \quad (49.13)$$

Примем

$$\epsilon_0 \approx \frac{1}{500}. \quad (49.14)$$

В каждый момент времени для этого излучения справедливо следующее соотношение:

$$\rho_{\gamma}R^4 = \text{const}, \quad (49.15)$$

так что при красном смещении z имеем

$$\frac{8\pi\rho_{\gamma}R^2}{3} = \frac{8\pi\rho_{\gamma}R^4}{3R_0^2}z^2 = \epsilon_0\Omega_0z^2. \quad (49.16)$$

Зная плотность и радиус, мы можем определить параметры a и η критической фазы Вселенной.

Подробный расчет

Вычислим a и η из выражений в табл. 45.2. Мы имеем

$$R = a\eta, \quad (49.17)$$

$$\frac{8\pi\rho a^2}{3} = \frac{1}{\eta^4}. \quad (49.18)$$

Из соотношений (49.10) и (49.16) вытекает

$$a\eta = \frac{R_0}{z}, \quad (49.19)$$

$$\frac{8\pi\rho R^2}{3} = \frac{1}{\eta^2} = \epsilon_0\Omega_0z^2. \quad (49.20)$$

Второе соотношение дает

$$\eta = \frac{1}{z\sqrt{\epsilon_0\Omega_0}}, \quad (49.21)$$

а первое —

$$a = R_0\sqrt{\epsilon_0\Omega_0}. \quad (49.22)$$

Заметим, что a не зависит от z , как это и должно быть. Результат зависит от плотности излучения $\epsilon_0\Omega_0$, следовательно,

нам нужно знать лишь постоянную Хаббла, а не современную плотность материи. Это справедливо только во Вселенной с низкой плотностью

$$\Omega_0 \ll 1, \quad (49.23)$$

где мы можем положить

$$R_0 H_0 \approx 1. \quad (49.24)$$

Собственное время, отсчитанное от 10^{10} К, определяется следующей величиной:

$$\Delta t = \frac{R_0}{2\sqrt{\epsilon_0 \Omega_0}} \left(\frac{1}{z_9^2} - \frac{1}{z_{10}^2} \right), \quad (49.25)$$

[Здесь z_9 обозначает красное смещение z при температуре 10^9 К.]

которая примерно равна

$$\Delta t \approx \frac{R_0}{2\sqrt{\epsilon_0 \Omega_0}} \left(\frac{3}{10^9} \right)^2. \quad (49.26)$$

В принятой нами модели это выражение приводит к интервалу времени порядка 300 с.

Именно за такое время происходит распад нейтронов. Должно быть, грубое совпадение этого времени с временем жизни нейтрона (порядка 1000 с) следует рассматривать как космическую случайность. По истечении 300 с доля оставшихся нейтронов составляет

$$e^{-300/1000} = 0,75. \quad (49.27)$$

Итак, 25% (четвертая часть) нейтронов распадается, а оставшиеся три четверти соединяются с таким же числом протонов, образуя гелий. Таким образом, при Большом Взрыве 30% водорода (по массе) превращается в гелий. К настоящему времени проведены значительно более тщательные расчеты, и результат их состоит в том, что примерно 23% водорода превращается в гелий.

[Идею приведенного выше сценария мне подсказала отличная обзорная статья [14]. Более подробное рассмотрение можно найти в этой статье или книге [36].]

Замечательно, что, согласно результатам наблюдений, во Вселенной гелий всегда смешан с водородом приблизительно в указанной выше пропорции. Правда, иногда гелия бывает больше, поскольку водород выгорает в звездах, превращаясь в гелий; но в целом содержание гелия во Вселенной более или менее соответствует приведенному выше расчету. Это один из триумфальных результатов космологии Большого Взрыва. При его выводе использовалось не очень большое количество предположений сверх того, что даже при $z = 10^{10}$ Вселенная была достаточно однородной и изотропной, чтобы ее можно было описывать фридмановской моделью.

ЗАДАЧИ

49.1. (19) Можно было бы предположить, что учет материи, из которой в конечном счете образуются галактики, повлияет на распределение микроволнового фонового излучения при $z = 1000$. Оцените угловой размер такого эффекта для различных моделей.

49.2. (28) Сравните количества гелия, образующиеся в пылевой и радиационной моделях с $\Omega = 1$.

49.3. (35) Рассмотрите образование гелия во Вселенной, которая на ранней стадии своей эволюции подчиняется уравнению состояния

$$p = \rho.$$

50. Наблюдения реальной Вселенной

*Созерцание вещей, как они
есть, без ошибки или путаницы,
подтасовки или обмана,
ценнее всех плодов воображения.*

Фрэнсис Бэкон

*Принцип стратегии таков:
знание одной вещи должно
порождать знание тысячи вещей.*

Мусаши

Теперь мы можем смело ответить на вопрос о том, много ли мы знаем о Вселенной. Применимы ли рассмотренные фридмановские модели? О чем они нам могут рассказать? Вы, возможно, будете разочарованы, когда поймете, как мало мы знаем в действительности. Мы можем получать модели, совместимые с нашими наблюдениями. Бесконечная Вселенная — уже не проблема для нас, как и парадокс Ольберса. Микроволновое фоновое излучение и распространенность гелия в космосе прекрасно согласуются с нашими представлениями. Вместе с тем кое-что не укладывается естественным образом в наши построения. Больше всего затруднений мы испытываем из-за то-

го обстоятельства, что в рассматриваемых моделях, по-видимому, не образуются галактики. Чтобы ввести их в рассмотрение, мы должны заложить их с самого начала. Мы не можем также сделать выбор между открытыми моделями, расширяющимися вечно, и замкнутыми моделями, эволюция которых завершается обратным гравитационным сжатием. В настоящем разделе дается краткий обзор наблюдательных данных, имеющих отношение к этим моделям.

Краеугольным камнем наших моделей является симметрия, которую нам нужно тщательно исследовать. Вводим ли мы предположение о симметрии только для того, чтобы иметь возможность решить получающиеся уравнения? В какой степени Вселенная симметрична в настоящее время? Мы можем наблюдать изотропию окружающего нас мира, но вывод об однородности можно сделать только косвенным путем. Мы хотим знать также, насколько симметричной была Вселенная на ранней стадии своей эволюции. Малые отклонения от симметрии называются возмущениями, и их изучение ведется весьма активно. Растущие возмущения важны потому, что они образуют галактики. Хорошая теория должна объяснить, почему галактики имеют именно такие массы, которые наблюдаются в действительности. Затухающие возмущения также важны. Анализ поведения этих возмущений позволяет на основе отклонений от симметрии в настоящее время сделать заключение о таких отклонениях в прошлом. Изучение таких возмущений весьма затруднительно и пока окончательно не завершено. Физическая задача состоит в гидродинамическом рассмотрении релятивистской радиационно-доминированной самогравитирующей плазмы в искривленном расширяющемся пространстве-времени. Насколько нам сейчас известно, в радиационно-доминированную эру возмущения растут не очень быстро¹⁾. Указанное обстоятельство мешает образованию галактик, но дает нам основание предположить, что ранняя Вселенная могла и не быть слишком симметричной. Начальные возмущения, оставшиеся после радиационной эры, вырастают впоследствии в 10^2 или 10^3 раз. Если мы примем, что начальные возмущения имели порядок 10^{-3} , то они проскочат через радиационную эру и за-

Симметрия

Возмущения

¹⁾ Исследования последних лет показывают, что и в радиационно-доминированную эру во фридмановской Вселенной возможно существование быстрорастущих мод возмущений, происходящих с нарушением локального термодинамического равновесия (см., например, *Хлебников В. И.* — ЖЭТФ, 1984, т. 86, вып. 1, с. 13—24). Однако роль таких возмущений в эволюции реальной Вселенной пока еще не ясна. — *Прим. ред.*

Изотропия микроволнового фонового излучения

[Будьте осторожны — этот наблюдательный факт требует лишь существования поверхности со сферической симметрией. Чтобы распространить эти соображения на все пространство, мы еще должны предположить однородность. Занятный контрпример описан в статье [10].]

Горизонты

тем, начав увеличиваться после $z = 1000$, образуют галактики при значении z , близком к 10. Остается, однако, вопрос: откуда они взялись первоначально?

Основное наблюдательное свидетельство симметрии Вселенной — изотропия микроволнового фонового излучения. Исследования, проведенные с помощью высотной авиации, показали, что это излучение во всех направлениях с точностью 10^{-3} имеет одинаковую температуру. Это замечательный наблюдательный факт. Поскольку, как мы полагаем, начиная с $z = 1000$, такое излучение распространяется к нам свободно, мы можем сказать, что при $z = 1000$ Вселенная была симметричной с точностью 10^{-3} . Наблюдения микроволнового фонового излучения позволяют также определить скорость нашей Галактики относительно системы отсчета, в которой рассматриваемое излучение изотропно. Эта скорость оказывается величиной порядка 10^{-3} , т.е. как раз порядка наблюдаемых хаотических скоростей других галактик. Последнее обстоятельство согласуется с тем, что наша Галактика является типичной, как того и требует однородность. Если бы наша скорость оказалась равной 10^{-6} , мы были бы в недоумении.

При дальнейшем размышлении изотропия микроволнового фонового излучения кажется еще более удивительным наблюдательным фактом. Одно из следствий того обстоятельства, что Вселенная имеет конечный возраст по дуговому времени, заключается в том, что различные части Вселенной могут взаимодействовать лишь по истечении определенного интервала времени. Две точки, удаленные друг от друга на дуговое расстояние χ , не могут провзаимодействовать, пока не пройдет дуговое время χ . Когда мы смотрим на микроволновое фоновое излучение в двух различных направлениях, мы видим две различные точки в плазме при $z = 1000$. Хватило ли этим точкам времени, чтобы провзаимодействовать? На этот вопрос следует ответить отрицательно; ведь они находятся очень далеко в направлении своих лучей зрения.

Пример

Вычисление того расстояния, на котором должны находиться эти точки на небе, представляет собой забавную геометрическую задачу. Обратимся к моделям, использованным в вычислениях «своими руками». В табл. 50.1 представлено дуговое время χ , истекшее с момента $z = 1000$, а также дуговое время η_{BB} , прошедшее от начала расширения Вселенной до этого момента. Здесь рассматривается та же самая геометрическая си-

**ТАБЛИЦА 50.1. ВЕЛИЧИНЫ, ХАРАКТЕРИЗУЮЩИЕ ГОРИЗОНТ
ДЛЯ РАССМАТРИВАЕМЫХ МОДЕЛЕЙ
ВСЕЛЕННОЙ**

Тип	Дуговое расстояние до $z = 1000$	Дуговое время при $z = 1000$	Угловой размер горизонта, ϵ
Пыль	4,2 рад	0,3 рад	0,009 рад
Пыль с излучением	2,2	0,004	0,001

туация, что и в разд. 43 (с. 345 и далее), с той лишь разницей, что вместо указанного собственного расстояния между точками у нас фигурирует дуговое расстояние. Благодаря этому исчезает фактор $R(\eta)$. Кроме того, рассматриваемая модель представляет собой открытую Вселенную, так что $\sin \chi$ заменяется на $\text{sh} \chi$. Угол ϵ между двумя точками при $z = 1000$, которые только что вступили в физическую связь, в настоящее время имеет следующую величину:

$$\epsilon = \frac{\eta_{BB}}{\text{sh} \chi}; \quad (50.1)$$

соответствующие данные также приведены в табл. 50.1.

Угловое расстояние между несвязанными точками отчасти зависит от модели Вселенной, которую вы выберете. Модели с меньшим количеством излучения характеризуются более медленным ранним расширением, но даже в чисто пылевой модели точки, удаленные друг от друга более чем на $0,5^\circ$, все еще являются невзаимодействующими. Граница области, события в которой не могут быть связаны между собой физически, называется горизонтом. Для произвольного наблюдателя горизонт находится при $z = \infty$.

Пример

В замкнутой Вселенной в момент дугового времени 90° в принципе нельзя наблюдать больше половины Вселенной, сколь бы мощным телескопом мы ни пользовались.

Интересно отметить, что все части горячей плазмы (насколько мы можем судить о них по микроволновому фоновому излучению) с точностью 10^{-3} обладают одинаковой темпера-

турой, даже если они еще не имели возможности провзаимодействовать. Большой Взрыв должен был начаться в одно и то же время всюду. Спрашивается, как же тогда могло происходить согласование температур различных частей плазмы? Ни какой из известных нам физических процессов не мог бы выравнивать такие температурные флуктуации. Если рассматриваемые фридмановские модели верны, то соответствующая симметрия должна была присутствовать с самого начала.

Энтропия

[Я не уверен, что мои читатели достаточно хорошо знакомы с энтропией, чтобы по достоинству оценить приведенные соображения. По-настоящему я и сам не могу проследить их во всех деталях.]

Существует еще один косвенный аргумент в пользу того, что ранняя Вселенная должна быть почти столь же симметрична, как и фридмановская модель. Если бы флуктуации и структура в ранней Вселенной размывались в результате некоторых диссипативных процессов, то такая диссипация увеличила бы энтропию Вселенной. Мы можем оценить эту энтропию, поскольку наибольшая ее часть связана с микроволновым фоновым излучением (а также с соответствующим нейтринным фоном, который должен существовать, но ненаблюдаем). Полученная таким образом энтропия Вселенной не настолько велика, чтобы ее можно было согласовать с очень хаотической ранней Вселенной.

Теперь с некоторой осторожностью допустим, что фридмановская Вселенная представляет собой хорошую модель нашей Вселенной. Переходя от моделей Робертсона — Уокера, основанных лишь на симметрии, к моделям Фридмана, мы тем самым допускаем, что правильной теорией тяготения является общая теория относительности. Нельзя ожидать, что космология сама по себе даст ответ на интересующий нас вопрос. Если это фридмановская Вселенная, то какая именно? Их много в соответствии с выбором уравнения состояния (например, для пыли или для излучения), одного из трех возможных пространственных сечений, параметра a и нашего современного положения в пределах данной модели. В заключение настоящего раздела обсудим, насколько может быть ограничен произвол, обусловленный перечисленными возможностями.

Двухпараметрическое семейство

Выбор H_0 и Ω в качестве упомянутых двух параметров для нашей фридмановской модели наиболее удобен. Используя Ω как параметр возраста, мы автоматически учитываем все три возможности пространственных сечений. Считая Ω свободным параметром, мы тем самым допускаем возможность существования во Вселенной той или иной до сих пор не открытой материи. Для такой скрытой материи трудно или даже невозможно выбрать какое-либо одно из различных уравнений состояния. Вместе с тем мы знаем настолько мало, что любой дополнительный неопределяемый параметр убьет нас. Обычно

скрытую материю рассматривают как пыль, поскольку дополнительное давление излучения заставило бы Вселенную расширяться столь быстро, что соответствующее собственное время едва включало бы известный возраст шаровых скоплений. Скрытая материя могла бы представлять собой также предполагаемую лоренц-инвариантную материю, описываемую космологической постоянной. Такая материя обладает универсальным отрицательным давлением. Оно замедляет расширение и может подправить модель, которая иначе была бы слишком молодой. Примем модель пылевой Вселенной. Тогда каждая точка в пространстве (H_0, Ω) будет задавать единственную космологическую модель.

Величину H_0 можно измерить непосредственно, хотя и не очень точно. Такое измерение требует прямой оценки расстояний до объектов. Если расстояние до объектов невелико, то красные смещения, обусловленные расширением Вселенной, будут трудно отличимыми от доплеровских смещений, обусловленных хаотическим движением объектов. С другой стороны, для удаленных объектов возникает много новых проблем. Объекты, видимые с большого расстояния, должны быть яркими и, следовательно, редко встречаются локально. Далее, очень трудно измерять расстояние до удаленных объектов. Наконец, объекты, находящиеся на больших расстояниях, моложе. Чтобы использовать такие объекты, нужно построить модель, предсказывающую эволюцию их свойств. В результате многих десятилетий работы были получены следующие оценки:

$$30 \leq H_0 \leq 120 \text{ км/(с} \cdot \text{Мпс)}. \quad (50.2)$$

По-видимому, не таким уж плохим значением H_0 является 50 км/(с · Мпс). Мы использовали границы (50.2), чтобы отделить части плоскости (H_0, Ω) на рис. 50.1.

Следующая характеристика — возраст Вселенной. Осторожная оценка возраста звезд, звездных скоплений и самих элементов такова:

$$8 \cdot 10^9 \leq t_0 \leq 18 \cdot 10^9 \text{ лет}. \quad (50.3)$$

Оказывается, эти возрастные пределы исключают из рассмотрения на плоскости (H_0, Ω) почти те же самые области, что и пределы изменения постоянной Хаббла. Любая такая внутренняя согласованность весьма утешительна.

Нижний предел величины Ω можно получить из наблюдаемого содержания материи во Вселенной. Однако величины

[По-моему, использование такой материи без каких-либо намеков на ее существование равносильно мошенничеству при раскладывании пасьянса.]

[Здесь я следую обзорной статье [13]; как и авторы этой статьи, я буду обсуждать наблюдения в плоскости (H_0, Ω) . Несмотря на то что я следую доводам упомянутых авторов, я щедрее их при установлении соответствующих пределов. Автор книги должен учитывать, что для уточнения данных может потребоваться много времени, и быть более осторожным в своих высказываниях.]

Постоянная Хаббла

Возраст

Плотность массы

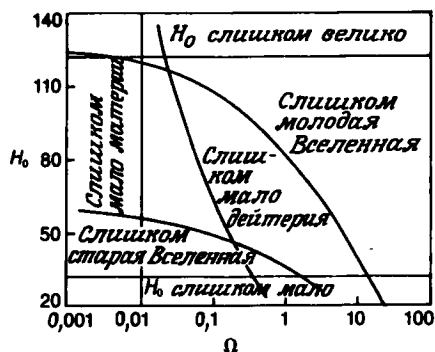


Рис. 50.1

Схематическое изображение (весьма осторожное) космологических представлений в конце 1970-х годов.

масс не наблюдаются непосредственно, и здесь всегда имеется некоторый произвол. Обычно величины масс выводятся из яркости в предположении, что галактики заполнены обычными объектами и все имеют сходное отношение массы к светимости. Можно также оценивать массы самогравитирующих скоплений галактик. Однако, как это следует из наблюдений, масса, необходимая для удержания такого скопления при заданных хаотических скоростях, должна быть примерно в 10 раз больше той, которая предсказывается первым методом. Большинство исследователей сошлись бы на нижнем пределе массы порядка

$$\Omega \geq 0,01. \quad (50.4)$$

При этом значение $\Omega \approx 0,05$ представляется вполне разумным для наблюдаемой материи. Оценка (50.4) ограничивает с одной стороны допустимую область значений (H_0, Ω) .

Диаграмма Хаббла

К сожалению, диаграмма Хаббла в действительности не дает хороших результатов. Яркость галактик изменяется со временем как вследствие старения и помутнения звезд, так и вследствие того обстоятельства, что издалека можно видеть лишь гигантские галактики, которые встречаются в скоплениях и могут становиться ярче за счет поглощения меньших галактик. Последний эффект почти невозможно смоделировать. Тем не менее интерпретацию современных данных трудно совместить с более высоким пределом, чем

$$\Omega < 5. \quad (50.5)$$

50. Наблюдения реальной Вселенной

Решающая информация для оценки величины параметра плотности может содержаться в данных по космической пространственности дейтерия. Известно, что в процессе образования гелия получается также некоторое количество дейтерия. Это количество непосредственно зависит от того, насколько быстро проскочила Вселенная через эру нуклеосинтеза. Оценка содержания дейтерия получается из детальных численных расчетов. Ясно, что космическая распространенность дейтерия может послужить источником важной информации. Однако образование дейтерия настолько чувствительно к условиям в первые несколько минут расширения Вселенной, что соответствующие данные могут оказаться ненадежными. Не исключено, что дейтерий, который мы видим, синтезировался в процессе эволюции некоторых малых изолированных неоднородностей. Это ослабило бы получающиеся ограничения. Реально мы можем очень мало что сказать относительно первых нескольких минут расширения Вселенной.

Дейтерий

На рис. 50.1 графически представлена сводка (очень осторожная) наших знаний о Вселенной. Если соображения, основанные на содержании дейтерия, правильны, то мы имеем $\Omega < 1$ (для пыли $q_0 < 1/2$); соответствующая Вселенная является открытой, она будет бесконечно расширяться и охлаждаться¹⁾. Однако кое-что может быть и неверным в «дейтериевой» аргументации, и Ω может оказаться больше единицы. Если это так, то наша Вселенная замкнута, и все мы завершим свое существование в горниле Большого Треска (big crunch).

Резюме

ЗАДАЧИ

50.1. (15) Как были проведены линии постоянного возраста на рис. 50.1?

50.2. (35) Какова современная температура нейтрино, если они стали свободными частицами раньше по красному смещению, чем микроволновые фотоны?

50.3. (15) Проверьте числа в первой строке табл. 50.1, используя соответствующие аналитические выражения.

50.4. (25) Проверьте числа во второй строке табл. 50.1 при условии, что последняя точка, вычисленная при

$$z \doteq 35,7,$$

¹⁾ В критической и открытой модели вещество проходит стадию образования черных дыр и последующего их испарения, в результате которого остается лишь неограниченно расширяющееся «разжиженное» поле излучения. — *Прим. ред.*

была удалена на дуговое расстояние

$$\chi = 2,04$$

и предсказанный возраст Вселенной (задача 48.6) был равен

$$\hat{\eta} \doteq 0,125.$$

Проверку сделайте в предположении, что в ранние моменты времени Вселенная описывалась критической радиационной моделью.

50.5. (31) Допустим, что микроволновой фон образовался при $z = 20$ во Вселенной с $\Omega = 2$. Какой угол на небе образуют сегодня две точки, каждая из которых при $z = 20$ в точности находилась на горизонте другой? Заметим, что в этой задаче вы не можете предполагать малость углов.

51. Сингулярности

Идеи приходят ко мне без всякой связи с графикой и так очаровывают меня, что у меня возникает неодолимое желание донести их и до других людей.

Сделать это с помощью слов было бы невозможно, поскольку идеи возникают как чисто мысленные, а не литературные образы, которые становятся понятны другим, только когда они оформлены в виде образов зрительных.

М.К. Эшер

Заинтересованный читатель спросит: что же происходило во Вселенной до Большого Взрыва? Современная физика не даст ответа на этот вопрос. Не существует ответа и на вопрос о том, что будет происходить в замкнутой Вселенной после завершения ею своей эволюции. Хотелось бы закончить книгу обсуждением некоторых аспектов затронутой проблемы. Такие проблемы, для решения которых физика просто не в состоянии дать каких бы то ни было рецептов, весьма привлекательны. В них должны скрываться новые физические законы, которые, вероятно, вскоре будут открыты.

В ранние моменты времени все фридмановские Вселенные выглядят как радиационные Вселенные, если только они не содержат какую-нибудь необычную форму материи с предельно жестким уравнением состояния. В ранние моменты времени мы имеем для радиационной Вселенной

$$R(\eta) \approx a\eta \quad (51.1)$$

и следующую метрику:

$$\mathcal{G} \approx a^2 \eta^2 [-d\eta^2 + d\chi^2 + S^2(\chi) d\Omega^2]. \quad (51.2)$$

На рис. 51.1 мы изобразили поведение такой метрики, нарисовав в нескольких местах метрические фигуры. Координата η была использована с тем расчетом, чтобы получилась простая структура световых конусов. Метрические фигуры расширяются по мере приближения к $\eta = 0$. Расстояние между произвольными двумя точками стремится к нулю по мере приближения к этой гиперповерхности.

Такое поведение не является неожиданным. Даже обычные полярные координаты ведут себя таким образом. Метрические фигуры для евклидовой метрики

$$\mathcal{E} = dr^2 + r^2 d\theta^2 \quad (51.3)$$

также беспредельно расширяются при стремлении r к нулю, как показано на рис. 51.2. Рассматриваемое представление метрики $\mathcal{E}$ плохо ведет себя при $r \rightarrow 0$, поскольку кажущаяся линия $r = 0$ в действительности представляет собой единственную точку. Аналогичное обстоятельство могло бы быть проблемой и для наших космологических моделей. По-видимому, изображением замкнутой Вселенной может служить рис. 51.3. Я не могу сказать ничего больше, чем «по-видимому», так как с позиции современных физических законов бессмысленно утверждать, что Вселенная начиналась с единственной точки. Точка

Ранняя Вселенная

[Множество $\eta = 0$ представляет собой трехмерное подпространство; такие подпространства обычно называют гиперповерхностями.]

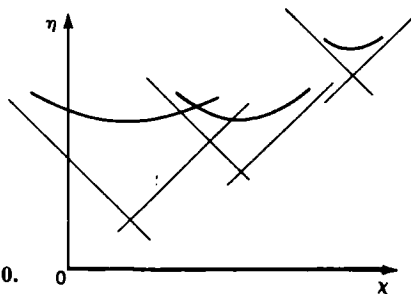


Рис. 51.1

Поведение метрики вблизи $R = 0$.

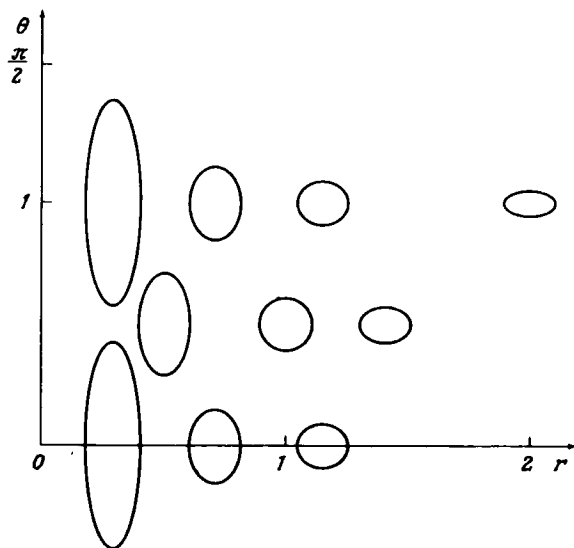


Рис. 51.2

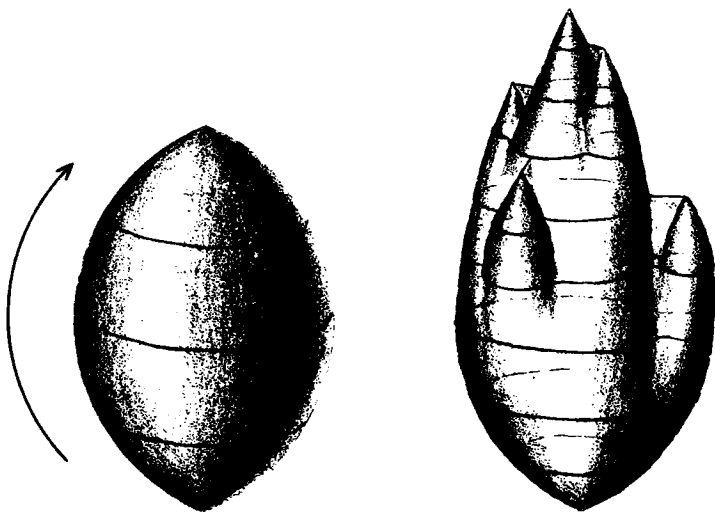
Поведение евклидовой метрики в полярных координатах.

Рис. 51.3

Представление замкнутой Вселенной, в котором Большой Взрыв и Большой Треск изображены как единственные точки.

Рис. 51.4

Попытка представить Большой Треск как совокупность многих Маленьких Тресков.



Начальная сингулярность

$R = 0$ в полярных координатах может быть правильно представлена как одна точка посредством перехода к некоторым другим координатам, например прямоугольным. В таких представлениях метрики $\mathcal{G}$ не происходит ничего странного в точке

$r = 0$. Вместе с тем, даже когда координаты для нашей космологической модели выбраны так, что $R = 0$ становится единственной точкой, в метрическом представлении все же сохраняются особенности. В этой точке расходится риманов тензор кривизны, что соответствует бесконечным приливным силам. Такая точка называется сингулярностью.

В действительности рис. 51.3 не может быть точным изображением будущего замкнутой Вселенной. Скучивание за счет самогравитации приводит к тому, что Вселенная становится все менее и менее симметричной. Хотя, как мы полагаем, Большой Взрыв мог быть гладким и упорядоченным, тем не менее конец эволюции замкнутой Вселенной, называемый Большим Треском, гладким не будет. Разные части Вселенной будут испытывать обратный коллапс с различными скоростями. Если в настоящее время существуют черные дыры, то они представляют собой части Вселенной, уже пришедшие в состояние Большого Треска. На рис. 51.4 мы попытались более детально отобразить хаотическую сущность Большого Треска. Отсутствие симметрии прошлого и будущего довольно интересно. Направление течения времени выглядит как направление от гладких сингулярностей к беспорядочным сингулярным структурам¹⁾.

Большой Треск

В предыдущем разделе была упомянута проблема конечного горизонта. Эта проблема наглядно поясняется на рис. 51.5. Во Вселенной не существует точек, сигнал из которых может быть послан одновременно в точки A и B , поскольку вторая находится за горизонтом первой, и наоборот. Несмотря на это, Вселенная в точках A и B настолько похожа, что, очевидно, эти точки должны как-то между собой «договориться».

Проблема горизонта

Одна из попыток разрешить сформулированную проблему состоит в продолжении Вселенной, что сделает возможным существование событий, происходящих до момента времени $\eta = 0$ ²⁾. Точка C на рис. 51.6 может взаимодействовать как с

До начала

¹⁾ Сказанное вплотную подводит нас к мысли о том, что энтропия классической фридмановской Вселенной возрастает в процессе ее эволюции. Как примирить этот вывод со сценарием вечно осциллирующей Вселенной, изображенной на рис. 51.7 и 51.9? Обсуждение проблемы энтропии вечно осциллирующей Вселенной и, в частности, возможной роли деситтеровской (инфляционной) стадии как квантового «чистилища» для накопленной энтропии см., например, в работе Марков М. А. Труды ФИАН СССР им. П. Н. Лебедева, 1983, т. 152, с. 3—11. — Прим. перев.

²⁾ Естественное решение проблемы горизонта дает сценарий инфляционной Вселенной; подробнее см. предисловие редактора перевода. — Прим. ред.

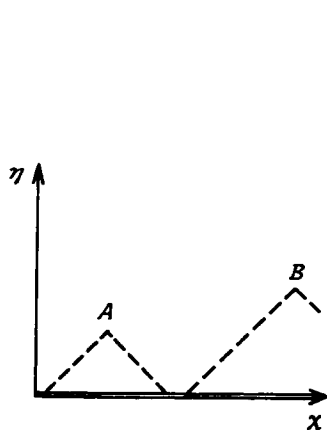


Рис. 51.5
Отсутствие причинного взаимодействия в ранней Вселенной.

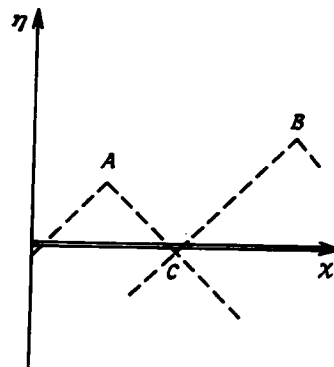


Рис. 51.6
События, происходящие до момента времени, отвечающего $R = 0$, могли бы взаимодействовать с разобщенными точками, изображенными на рис. 51.5.

[Это можно было бы назвать «конспиративной» теорией происхождения космоса.]

[Некоторые гипотезы более или менее подробно разработаны Хойлом и Нарликар. Популярное изложение этого круга вопросов содержится в книге [15].]

[По поводу таких причудливых множеств см. книгу о фракталах [23].]

[Эти рисунки настолько сложны, что наши математики не в состоянии их описать. В этой области ведутся активные исследования. Рисунки выполнены Кри-сом Шоу.]

точкой A , так и B . Идея заключается в том, что Вселенная может сколлапсировать, а затем испытать «отскок». Такую Вселенную называют *осциллирующей*. Концепция осциллирующей Вселенной позволяет перейти от картины на рис. 51.3 к картине на рис. 51.7. Беда здесь в том, что Большой Треск на рис. 51.3 представлен не совсем верно. Если же для представления Большого Треска используется рис. 51.4, то красивая упорядоченная осциллирующая Вселенная становится невозможной. Лучшее, что мы можем в этом случае предложить, — это кошмар, изображенный на рис. 51.8. На этой картине вся наша Вселенная умещалась бы внутри незначительной части большей Вселенной. Более того, сама наша Вселенная порождала бы меньшие дочерние Вселенные. Мы пришли бы к состоянию, которое лучше всего описывается как турбулентное пространство-время. Тогда рисунки 51.4 и 51.8 уже не могут служить хорошими представлениями Большого Треска. Они изображают Вселенную, распадающуюся в процессе эволюции на бесконечное число несвязанных кусков. Разумнее предположить, что Маленькие Трески собираются вместе и в результате образуют один Большой Треск. На рис. 51.9 сделана попытка наглядно изобразить эту идею.

Предыдущие причудливые сценарии представляют собой не что иное, как плоды нашей необузданной фантазии. Мы не знаем физических законов, которые позволили бы нам перейти че-

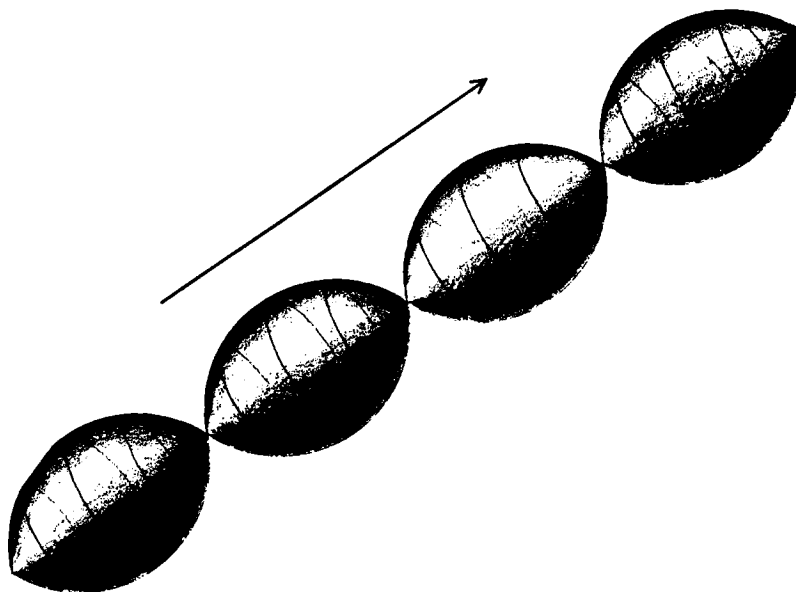


Рис. 51.7

Первое предположение о глобальной геометрии осциллирующей Вселенной. Время течет в направлении стрелки.

рез гиперповерхность $R = 0$. Такой переход недопустим в общей теории относительности. Никто, правда, и не считает, что вблизи $R = 0$ общая теория относительности строго справедлива. Нет сомнений в том, что это — всего лишь приближенная теория и она неприменима при экстремальных условиях вблизи гиперповерхности $R = 0$.

Мы высказали некоторые предварительные соображения о Большом Треске. Общая теория относительности может описывать коллапс облака вещества, давление которого не в состоянии сдержать его самогравитации. Конечное состояние такого коллапса называется черной дырой. В рассматриваемом случае объекты также коллапсируют к нулевому значению радиуса. На рис. 51.10 изображено пространство-время, описывающее такую ситуацию. Обратите внимание, в частности, на поведение световых конусов в этом пространстве-времени. Наблюдатель, падающий в черную дыру и пересекающий линию, помеченную как «горизонт», не в состоянии предотвра-

Гравитационный коллапс



Рис. 51.8

Схематическое представление рекуррентной Вселенной, которой соответствует рис. 51.4.

Черные дыры

тить попадания прямо в сингулярность, расположенную в начале координат. Чтобы избежать этого, потребовались бы скорости, превышающие скорость света. Язык метрических фигур и пространственно-временных диаграмм дает нам четкое представление о коллапсе. В действительности это единственный честный способ описания черной дыры.

Решающее свойство черной дыры заключается в настолько сильном тяготении, что даже свет не может избежать падения к центру. Коллапс в сингулярность мог бы показаться случайным результатом, обусловленным нашим предположением о точной сферической симметрии. Но это не так. В общей тео-

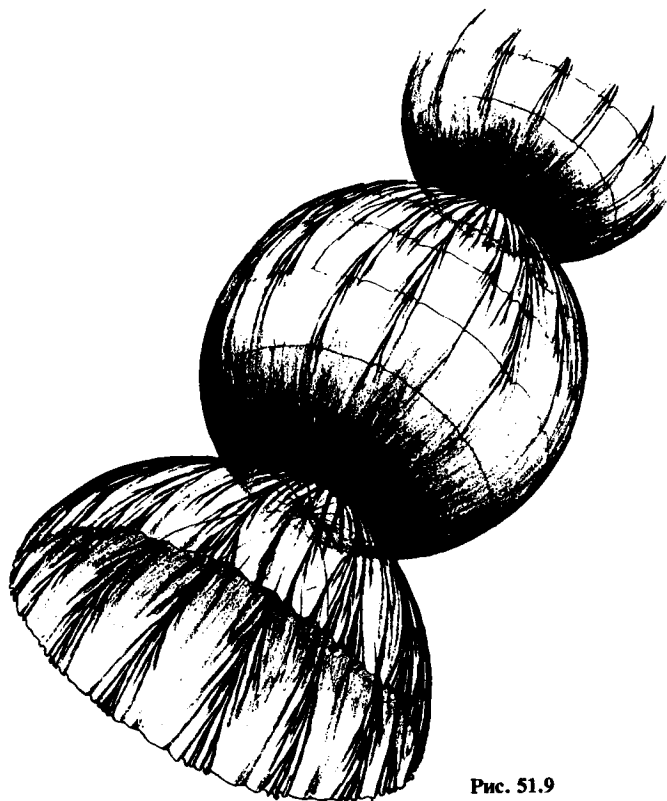


Рис. 51.9

рии относительности имеется несколько удивительных теорем, которые утверждают, что при довольно мягких условиях, даже без сферической симметрии, должны образовываться сингулярности некоторых типов.

Сингулярности в общей теории относительности, в частности сингулярность в момент возникновения фридмановской Вселенной, в настоящее время до конца не поняты. Один из путей экстраполяции наших физических законов за сингулярность Большого Взрыва состоит в интерпретации $\mathcal{G}$ не как метрики, а лишь как генератора дисперсионных уравнений. В этом случае нет необходимости сжимать поверхность $t = 0$ в одну точку. Она остается сингулярной поверхностью для дисперсионного уравнения. При таком подходе массы частиц, по-видимому, исчезают на рассматриваемой поверхности, но они могут быть легко образованы полем излучения. Однако такой подход, раз-

[С помощью тех методов, которыми вы овладели, можно изучить этот вопрос более подробно. Выражения для метрики, на основе которых был сделан данный рисунок, приводятся в разд. 31.4 книги [24]. Они называются координатами Эддингтона — Финкельштейна. В книге [24] рассматриваются также некоторые другие представления. Но вы не готовы к обсуждению вопроса о том, почему общая теория относительности приводит именно к такому пространству-времени.]



Рис. 51.10

Пространство-время вблизи черной дыры, образуемой коллапсирующим пылевым облаком. Метрические фигуры изображены в масштабе, приведенном к массе черной дыры.

витый Фредом Хойлом, не может дать полное описание сингулярностей. Проблема черной дыры заключается в том, что невозможно выйти за пределы горизонта, однажды попав внутрь. В настоящее время предполагают, что только полное квантовое рассмотрение может дать ответы на вопросы об этих сингулярностях, но к такому квантовому рассмотрению мы пока еще не готовы.

Литература

1. *Acton F. S.*, Numerical Methods That Work, Harper and Row, 1970.
2. *Allen C. W.*, Astrophysical Quantities, University of London Press, 1973. [Имеется перевод: *Аллен К. У.* Астрофизические величины. — М.: ИЛ, 1960.]
3. *Арнольд В. И.* Обыкновенные дифференциальные уравнения. — М.: Наука, 1975.
4. *Арнольд В. И.* Математические методы классической механики. — М.: Наука, 1974.
5. *Arnold V. I., Avez A.*, Ergodic Problems of Classical Mechanics, W. A. Benjamin, 1968.
6. *Baldwin J. A., Burke W. L., Martin Gaskell C., Wampler E. J.*, Nature, 273, 431 — 435 (1978).
7. *Bondi H.*, Cosmology, Cambridge University Press, 1961.
8. *Davies P. C. W.*, Space and Time in the Modern Universe, Cambridge University Press, 1977. [Имеется перевод: *Девус П.* Пространство и время в современной картине Вселенной. — М.: Мир, 1979.]
9. *Dodson C. T. J., Poston T.*, Tensor Geometry, Pitman, 1977.
10. *Ellis G. F. R., Maartens R., Nel S. D.*, Monthly Notices R.A.S., 184, 439 — 466 (1978).
11. *Feynman R. P., Leighton R. B., Sands M.*, The Feynman Lectures on Physics, Addison-Wesley, 1963. [Имеется перевод: *Фейнман Р., Лейтон Р., Сэндс М.* Фейнмановские лекции по физике. — М.: Мир, 1966.]
12. *Gingerich O.*, ed., Cosmology + 1, W. H. Freeman, 1977.
13. *Gott J. R., Gunn J. E., Schramm D. N., Tinsley B. M.*, Astrophys. J., 194, 543—553 (1974).
14. *Harrison E. R.*, Ann. Rev. Astron. Astrophys., 11, 155—186 (1973).
15. *Hoyle F.*, Astronomy and Cosmology, W.H. Freeman, 1975.
16. *Hoyle F., Hoyle G.*, Into Deepest Space, Harper and Row, 1974.
17. *Lanczos C.*, The Variational Principles of Mechanics, University of Toronto Press, 1957. [Имеется перевод: *Ланцош К.* Вариационные принципы механики. — М.: Мир, 1965.]
18. *Lanczos C.*, The Einstein Decade, Academic Press, 1974.
19. *LeGuin U. K.* "Direction of the Road", reprinted in Orbit 12, D. Knight, ed., Berkeley Medallion Books, 1973
20. *Lighthill J.*, Waves in Fluids, Cambridge University Press, 1978. [Имеется перевод: *Лайтхилл Дж.*, Волны в жидкостях. — М.: Мир, 1981.]
21. *Loomis L. H., Sternberg S.*, Advanced Calculus, Addison-Wesley, 1968.

22. *Lorain P., Carson D.*, Electromagnetic Fields and Waves, W. H. Freeman, 2nd ed., 1970.
23. *Mandelbrot B.*, Fractals, W.H. Freeman, 1977.
24. *Misner C. W., Thorne K. S., Wheeler J. A.*, Gravitation, W. H. Freeman, 1973. [Имеется перевод: *Мизнер Ч., Торн К., Уилер Дж.* Гравитация. В 3-х томах. — М.: Мир, 1977.]
25. *Ohanian H. C.*, Gravitation and Spacetime, W.W. Norton, 1976.
26. *Peebles P. J. E.*, Physical Cosmology, Princeton University Press, 1971. [Имеется перевод: *Пиблс П.* Физическая космология. — М.: Мир, 1975.]
27. *Pierce J. R.*, Almost All About Waves, MIT Press, 1974. [Имеется перевод: *Пирс Дж.* Почти все о волнах. — М.: Мир, 1976]
28. *Polya G.*, Induction and Analogy in Mathematics, Princeton University Press, 1954.
29. *Polya G.*, Patterns of Plausible Inference, Princeton University Press, 1954. [Имеется перевод: *Пойа Г.* Математика и правдоподобные рассуждения. — М.: Наука, 1975.]
30. *Porteous I. R.*, Topological Geometry, Van Nostrand-Reinhold, 1969.
31. *Poston T., Stewart I.*, Catastrophe Theory and Its Applications, Pitman, 1978. [Имеется перевод: *Постон Т., Стюарт И.* Теория катастроф. — М.: Мир, 1980.]
32. *Reichenbach H.*, The Philosophy of Space and Time, Dover, 1958.
33. *Rindler W.*, Essential Relativity, Springer, 2nd ed., 1977.
34. *Robertson H. P., Noonan T.*, Relativistic Cosmology, W.B. Saunders, 1968.
35. *Sciama D. W.*, Modern Cosmology, Cambridge University Press, 1971. [Имеется перевод: *Шама Д.* Современная космология. — М.: Мир, 1973.]
36. *Silk J.*, The Big Bang, W.H. Freeman, 1980. [Имеется перевод: *Силк Дж.* Большой взрыв. — М.: Мир, 1982.]
37. *Synge J. L., Schild A.*, Tensor Calculus, University of Toronto Press, 1956.
38. *Synge J. L.*, Proc. Roy. Irish Acad., **63**, 1—34 (1963).
39. *Rosen J.*, Symmetry Discovered, Cambridge University Press, 1975.
40. *Taylor E. F., Wheeler J. A.*, Spacetime Physics, W.H. Freeman, 1966. [Имеется перевод: *Тейлор Э., Уилер Дж.* Физика пространства-времени. — М.: Мир, 1969.]
41. *Weinberg G. M.*, An Introduction to General Systems Thinking, Wiley, 1975.
42. *Whitham G. B.*, Linear and Nonlinear Waves, Wiley, 1974. [Имеется перевод: *Уизем Дж.* Линейные и нелинейные волны. — М.: Мир, 1977.]

Предметный указатель

Аберрация 110
Абсолютное время 42

Базис координатный 173
Базисные касательные векторы 138
— 1-формы 137
Большой Взрыв 380
Большой Треск 397
Быстрота 81
 как координата на псевдо-
 сфере 350

Ван-дер-Поля осциллятор 194
Вектор
 базисный 97
 ковариантность по отношению к преобразованиям *Лоренца* 98
 свободный 96

Векторное поле 191
Векторное пространство 34
Великого объединения модель 7

Волновая диаграмма 216
Волновое уравнение 181, 220
Волновое число 150
Волновой пакет 149
 уравнение движения 213
Волны на воде
 групповая скорость на глубокой воде 214

 дисперсионное уравнение для глубокой воды 153
 — — — конечной глубины воды 236
 отражение от движущегося объекта 242
 при учете поверхностного натяжения 224
 симметрия 94, 234
Волны упругости (дисперсионное уравнение) 153
Время в расширяющейся Вселенной 312
Вселенная
 возраст 302, 340, 391
 евклидова 356
 замкнутая с пылью 335
 осциллирующая 398
 радиационная 361
 фридмановская критическая 356

Галактики (образование) 387
Галилеева симметрия 53
Гамильтонова механика 194
Гармонический осциллятор 194
Гелий (образование при Большом Взрыве) 303, 383
Геодезическая 281
Геодезический квадрат 292
Геометрические единицы 254
Глобальная структура пространства 316

- Горизонты 388
 Градиент 131, 137
 Групповая скорость 149, 154, 214
Гака закон 166
Гюйгенса принцип 270
- Движение быстрее света 95
 Действие 287
 Дейтерий 383, 393
 Декартово произведение 164
 Диск вращающийся 108
 Диспергирующие волны 256
 Дисперсионное уравнение 152, 213
 для волн на воде при учете поверхностного натяжения 224
 для различных волновых уравнений 153
 для частиц в искривленном пространстве-времени 282
 Дифракция рентгеновских лучей 241
 Диздрическое произведение 103
 Доплеровское смещение 104, 182
 Дуальность 126
 Дуговое время 329
 Дуговое расстояние 330
- Евклидов метрический тензор 162, 165, 174
 Евклидова Вселенная 356
 Евклидова геометрия 46
 Евклидово скалярное произведение 46
- Закон сохранения 4-импульса 115
 Замедление хода времени 74
- Замедления параметр 344
 Зеленый луч 273
- Излучение (уравнение состояния) 332
 Изометрия 308
 Изотропия 308
 4-импульс 114
 сохранение (из волновой диаграммы) 246
 Инвариантность 54
 Инерциальная система отсчета 40, 52
 Интегральная кривая 191
 Инфляционная стадия эволюции Вселенной 8
- Каноническая система отсчета 53, 65
 для волн на воде 228
 операциональное определение 83
- Карты 202
 для S^2 315
 для S^3 319
 неполные 203
- Касательное пространство 134
 Касательный вектор 128
 к искривленной мировой линии 101
 к прямой мировой линии 97
 на многообразии 204
- Катастроф теория 148
 Квазары 304
 Квазичастица 148
Кеплера закон 256
Клейна — Гордона уравнение 150
 дисперсионное 152
- Ковариантность 45, 49, 54
 относительно линейных преобразований 125

- Ковекторы 121
 - линейность пространства ковекторов 124
 - свободные 123
- Коллапс гравитационный 399
- Конфигурационное пространство 141
- Космологическая постоянная 334
- Красное смещение
 - гравитационное 267, 268, 280
 - космологическое 330
- Кривая параметризованная 128
- Кривизна 260, 291, 295
 - псевдосферы 299
 - 3-сферы 326
- Кронекера* символ 188
- Лагранжа* множители 127, 144
- Лауэ* уравнения 241
- Линейное векторное пространство 34
- Линейные преобразования 44
- Лоренца* преобразования 86, 99
- Лоренц-инвариантность 83
- Майкельсона — Морли* эксперимент 76
- Масса 253, 254
 - покоя 115
- Маятник 145
- Метрика вблизи Земли 257
- Метрическая фигура 48, 327
- Микроволновое фоновое излучение 304
 - в первичном огненном шаре 381
 - наблюдаемая изотропия 388
- Минковского* пространство-время 177, 183
- Мираж 273
- Мировая линия 33
- Многообразие 200
- Мюоны 58
- Неевклидова геометрия 92
- Немые индексы 186
- Ньютонова космология 369
- Обобщенная сила 143
- Обратный тензор 189
- Огибающая 184, 223
- Ограничение (pullback) 158
- Одновременность 68
 - световая 70
 - хронотранспортная 69
- Однородность 308
- Ольберса* парадокс 302
- Ортонормированный базис 97
- ОТО, геометрическое обоснование 252, 253
- Отображения 36
- Отражение 241
- Парадокс близнецов 95, 236
- Парадоксы 92
- Параметр
 - замедления 344
 - плотности 344
- Перевероты 209
- Планковская Вселенная 358
- Полилинейность 164
- Полуклассические модели 212
- Поток векторного поля 196
- Предельно жесткая материя 334
- Преобразования 44
 - инфинитезимальные 196
 - проективные 52
- Приливы 297
- Принцип
 - виртуальных перемещений 142
 - наименьшего действия 288
 - эквивалентности 279

- Продолжение (push-forward)
158
- Произведение
 декартово 164
 диздрическое 103
 скалярное 46, 98
 тензорное 170
- Производная по направлению
135, 139
- Пространство-время 32
 диаграмма 33
 черной дыры 253
- Пространство
 дуальное касательному 134
 импульсов 116
 поворотов 208
- Псевдосфера 311, 349
- Пуанкаре* полуплоскость 93, 308
- Пыль
 в замкнутой модели Вселенной 335
 уравнение состояния 332
- Пятиугольник (построение)
299
- Расстояние дуговое 330
- Римана* тензор кривизны 295
- Робертсона* — *Уокера* пространство-время 306
- Рунге* — *Кутты* численный алгоритм 375
- Рябь 224
- Свет
 «молодой», «старый» 62
 отклонение массивными объектами 297
- Световая секунда 57
- Световой конус 61
- Световой сигнал
 в ОТО 254
 групповая скорость 255
 постулаты 60
- Светоподобные координаты 85
- Свободные индексы 186
- Свободные частицы 37
 переопределение в ОТО 251
- Связи 143
- Сдвиги 44
- Сила обобщенная 143
- Симметрия 53, 196, 259, 264, 307, 387
- Сингулярность 394
- Системы отсчета
 привилегированные 65
 специализированные 63
- Скорость относительная 78
- 4-скорость 99
- Сложение скоростей 80, 101
- Снеллиуса* закон 237, 244
- Собственное время 102
- Событие 32
- Статика 141
- Стереографическая проекция 280, 351
- СТО
 «доказательства» 95
 непротиворечивость 92
- Суммирования правило 187
- 3-сфера 310, 319, 324
- Тахион 246
- Твердые стержни 76, 94
- Тензор 161
 проводимости 167
 симметричный 165
 упругости 166
- Тор 210
- Транспортир (в СТО) 81
- Ускорение 103
 влияние на скорость хода часов 57
 волнового пакета 279

Фазовая скорость 149

1-форма 132

Фотоны 116

Френеля уравнения 237

Фридмана уравнения 333, 373

Фридмановские модели Вселенной 331, 356, 362
радиационные 361

Хаббла

диаграмма 364

постоянная 342

Хаббловское разбегание 304

Характеристики 197

Хиггса поле 7

Частота 150

Часы 38, 57, 73

для волн на воде 226

для измерения абсолютного времени 42

парадокс 96

спецрелятивистские 65

ускоренные 57

Черная дыра 400

Шаровые скопления 303

Эйнштейна уравнения 298

Электромагнетизм 182, 195

Энтропия Вселенной 390

Этвеша эксперимент 279

Эшера предел на круге 354

Оглавление

Предисловие редактора перевода	5
Предисловие автора к русскому изданию	14
Предисловие автора	15
Благодарности	18
Введение	19
Как пользоваться книгой	26
Глава I. Специальная теория относительности	31
1. Структура пространства-времени	32
2. Часы	41
3. Евклидова геометрия	46
4. Инерциальные системы отсчета	50
5. Физические часы	55
6. Световые сигналы и канонические системы отсчета	60
7. Специальная теория относительности .	66
8. Одновременность	68
9. Относительная скорость	78
10. Лоренц-инвариантность	83
11. Непротиворечивость специальной теории относительности	92
12. 4-векторы	96
13. Доплеровское смещение	104
14. 4-импульс	114
Глава II. Геометрия	119
15. Векторы и ковекторы	121
16. Касательные векторы и 1-формы	128
17. Векторы координатного базиса	136
18. Пример: статическое равновесие	141
19. Пример: диспергирующие волны	148
20. Отображения	157
21. Тензоры	161

22.	Базисные векторы и тензорное произведение	170
23.	Пространство Минковского	177
24.	Индексные обозначения	184
25.	Векторные поля	191
26.	Многообразия	199
Глава III. Гравитация		212
27.	Распространение волновых пакетов	213
28.	Теория относительности и волны на воде	223
29.	Взаимодействие волновых пакетов	237
30.	Тяготение	250
31.	Пространство-время в окрестности Земли	257
32.	Гравитационное красное смещение	264
33.	Принцип Гюйгенса и падающее яблоко	270
34.	Геодезические	281
35.	Кривизна пространства-времени	291
Глава IV. Космология		300
36.	Из чего состоит Вселенная	300
37.	Модели пространства-времени Робертсона — Уокера	306
38.	Глобальная структура 3-сферы	315
39.	Метрическая структура 3-сферы	324
40.	Распространение света	328
41.	Фридмановские Вселенные	331
42.	Замкнутая пылевая Вселенная	335
43.	Наблюдения в замкнутой пылевой Вселенной	340
44.	Псевдосфера	349
45.	Пылевые и радиационные фридмановские Вселенные	355
46.	Диаграммы Хаббла	364
47.	Ньютонова космология	369
48.	Космология «сделай сам»	373
49.	Большой Взрыв	380
50.	Наблюдения реальной Вселенной	386
51.	Сингулярности	394
Литература		403
Предметный указатель		405

Уважаемый читатель!

**Ваши замечания о содержании книги,
ее оформлении, качестве перевода и другие просим
присылать по адресу: 129820, Москва, И-110, ГСП,
1-Рижский пер., д. 2, изд-во «Мир».**

У. Бёрке

**ПРОСТРАНСТВО — ВРЕМЯ, ГЕОМЕТРИЯ,
КОСМОЛОГИЯ**

Научный редактор И. Г. Нахимсон
Мл. научные редакторы Р. Х. Зацепина, В. Н. Цлаф
Художник Н. М. Иванов
Художественный редактор В. А. Захаров
Технический редактор Л. А. Тихомирова
Корректор И. Н. Холкина

ИБ № 5027

Подписано к печати 22.05.85 г.
Формат 70 × 90 $\frac{1}{16}$. Бумага офсетная № 1.
Гарнитура таймс. Печать офсетная.
Объем 13,00 бум.л. Усл.печ.л. 30,42.
Усл. кр.-отт. 60,84. Уч.-изд.л. 24,42.
Изд. № 2/3351. Тираж 11 000 экз, Зак. 649.
Цена 2 руб. 10 коп.

Набрано в издательстве «Мир» на фотонаборном комплексе
«Компьюграфик»
129820, ГСП Москва, 1-й Рижский пер., 2.

Можайский полиграфкомбинат Союзполиграфпрома при
Государственном комитете СССР по делам издательств, полиграфии и
книжной торговли.
143200, Можайск, ул. Мира, 93.