

Физический институт им. П.Н.Лебедева РАН

Учебно-Научный Комплекс

В.С.Бескин

Гравитация и астрофизика

Москва 2007

Содержание

1	Предисловие	5
2	Введение	7
2.1	Астрофизическое введение	7
2.2	Математическое введение	10
3	Общая теория относительности – одна из первых теорий поля	11
3.1	Ньютоновская гравитация – пример нерелятивистской теории	11
3.2	Почему потребовалось изменить теорию? Лоренц-инвариантность	12
3.3	Кинетическая энергия	20
3.4	Появляются тензоры	24
3.4.1	Трехмерное пространство	24
3.4.2	Четырехмерное пространство	29
3.5	Кривизна	34
3.5.1	Определение кривизны	34
3.5.2	Геодезические	37
3.5.3	Треугольники и окружности	38
3.5.4	Гауссова кривизна	42
3.6	Прямое усмотрение истины	44
3.6.1	Наводящие соображения	44
3.6.2	Уравнения Эйнштейна	53
3.7	Бескрайние возможности	62
4	Наблюдательные проявления общей теории относительности	67
4.1	Эффекты в слабых гравитационных полях	67
4.2	Классические постньютоновские эксперименты	72
4.2.1	Гравитационное красное смещение	72
4.2.2	Отклонение света в поле Солнца	75
4.2.3	Движение перигелия Меркурия	76
4.3	Новые эксперименты	77
4.3.1	Задержка Шапиро	77
4.3.2	Радиопульсары в тесных двойных системах	78
4.3.3	Гравитационные линзы	83
4.3.4	Микролинзирование	85
4.3.5	Прецессия гироскопов	86

5	Черные дыры	89
5.1	Метрика Шварцшильда	89
5.2	Вращающиеся черные дыры	96
5.3	Астрофизика черных дыр	101
5.3.1	Галактические черные дыры солнечных масс	102
5.3.2	Сверхмассивные черные дыры в центрах галактик	103
5.4	Электродинамическая модель "центральной машины" в активных галактических ядрах	106
5.4.1	Струйные выбросы	106
5.4.2	Источник активности – вращающаяся черная дыра	108
5.4.3	Униполярный индуктор	110
5.4.4	Черная дыра в магнитном поле	114
5.4.5	Процесс Блэндфорда-Знайека	116
6	Заключение	121
7	Приложения	123
7.1	Закон преобразования координат "из первых принципов"	123
7.2	Понятие о ковариантной производной	125
7.3	Плотность энергии электромагнитного и слабого гравитационного поля	126
7.4	Магнитодипольное излучение	128
7.5	Излучение гравитационных волн	130
7.6	Функция масс	131
7.7	Гравитационное красное смещение в двойных системах	134
7.8	Движение периастра	134
8	Упражнения	137

1 Предисловие

Еще со школы я обратил внимание на то, что в учебной литературе, посвященной общей теории относительности (ОТО), имеет место явный разрыв между популярным изложением, в котором практически отсутствуют формулы, и последовательным строгим изложением, требующим знания высшей математики и доступным поэтому лишь людям, имеющим соответствующую подготовку. Конечно, общую теорию относительности, окончательно сформулированную Альбертом Эйнштейном (1879-1955) в ноябре 1915 года, никак нельзя назвать простой теорией. Поэтому споры о ее справедливости ведутся до сих пор. Однако основные идеи, на которых была построена общая теория относительности, являются все же достаточно наглядными. Это связано, как мы увидим, с геометрической природой многих ключевых положений, лежащих в основе ОТО. Поэтому и количественные основы ОТО могут быть изложены на достаточно простом языке, доступным школьникам старших классов. При этом обсуждение математических основ общей теории относительности позволяет прикоснуться и к самой сердцевине современных теорий частиц и физических полей. Действительно, общая теория относительности была одной из первых дедуктивных теорий, построенной на основе некоторых основных постулатов, которым должна удовлетворять последовательная теория. Иными словами, наблюдения дали только первый толчок, а все дальнейшие математические построения подчинялись лишь внутренней математической логике. Именно таким способом строится большинство современных теорий физических полей и элементарных частиц.

Предлагаемая книга основана на годовом курсе "Введение в специальность", который читался студентам 1-2 курсов кафедры проблем физики и астрофизики МФТИ. Он задумывался именно как промежуточное звено между школьной подготовкой и серьезным изучением теоретической физики и астрофизики на старших курсах. К сожалению, в учебных программах всегда имеет место достаточно много фундаментальных вопросов, на которые обычно не хватает времени. Именно обсуждению этих вопросов при минимальном использовании математики и посвящена настоящая книга. Тем самым, по мнению автора, и закладывается прочный фундамент, на котором в последствии может быть построено прочное знание. Вместе с тем, хотелось бы надеяться, что излагаемый материал будет понятен и для школьников старших классов, а также для всех, кто интересуется проблемами современной астрофизики.

Автор благодарит И.Д. Новикова и Б.Л. Воронова, многочисленные беседы с которыми позволили не потерять строгость изложения, а также В.В. Лосякова, внимательно прочитавшего рукопись и сделавшего много полезных замечаний.

2 Введение

2.1 Астрофизическое введение

Как хорошо известно, основной областью применимости общей теории относительности является астрофизика, уже много лет играющая роль полигона для проверки новых идей. Это, впрочем, не удивительно, поскольку во Вселенной существуют физические условия, которые невозможно воссоздать в лаборатории. Напомним, что именно точные астрономические измерения планетных орбит привели в итоге и к построению ньютоновской теории гравитации (и, одновременно, к созданию дифференциального и интегрального исчисления). Поэтому не удивительно, что изложение основ общей теории относительности естественно приведет нас к астрофизическим примерам, позволяющим если не проверить, то по крайней мере проиллюстрировать основные положения теории.

Здесь нелишне было бы напомнить, что практически во всех случаях, касающихся конкретных астрономических объектов (звезды, туманности, астероиды, космические лучи), в том числе и обнаруженных в последние годы (радиопульсары, рентгеновские пульсары, источники гамма-всплесков), вначале происходило накопление наблюдательной информации, и лишь затем ее физическое осмысление. Так, например, в 1967 году случайно были открыты радиопульсары – одиночные быстровращающиеся нейтронные звезды (масса порядка солнечной, $M \approx 1.4M_{\odot} \sim 3 \cdot 10^{33}$ г, при радиусе R лишь 10 – 15 км, характерный период вращения $P \sim 1$ с), которые должны образовываться в результате катастрофического сжатия (коллапса) обычных массивных звезд на поздней стадии их эволюции или, например, белых карликов, превысивших в результате аккреции чандрасекаровский предел массы $M_{\text{Ch}} \approx 1.4M_{\odot}$.

Интересно, что о возможности существования таких объектов говорилось еще с тридцатых годов прошлого века. Более того, в начале шестидесятых годов активно обсуждалась возможность сверхтекучести и сверхпроводимости во внутренних областях нейтронных звезд. Тем не менее, считалось, что из-за малого размера (и, значит, малого энерговыделения) обнаружить нейтронные звезды практически невозможно. При этом не было осознано, что нейтронные звезды должны вращаться настолько быстро, что основным источником излучаемой энергии будет кинетическая энергия вращения. А этот вывод мог быть легко сделан, поскольку за счет сохранения углового момента $J_r = I_r \Omega$, где $I_r \sim MR^2$ – момент инерции, сжимающаяся звезда неизбежно должна была резко увеличить свою угловую скорость вращения $\Omega = 2\pi/P$.

В результате, не было предпринято и целенаправленных попыток обнаружить пульсирующее излучение у известных объектов. И это несмотря на то, что к тому времени в центре Крабовидной туманности уже была зарегистрирована необычная оптическая звезда, совпадающая с также необычным радиоисточником. Именно с активностью этой

звезды связывалось энерговыделение $W_{\text{tot}} \approx 5 \cdot 10^{38}$ эрг/с, необходимое для подпитки Крабовидной туманности релятивистскими электронами. В противном случае оптическое свечение Крабовидной туманности давно бы прекратилось.

Лишь после того, как стало ясно, что этот необычный источник действительно связан с вращающейся нейтронной звездой, был проведен анализ переменности его оптического потока. Оказалось, что оптическое излучение также приходит к нам в виде отдельных импульсов, период которых ($P \approx 0.033$ с) в точности совпадает с периодом, определенным по данным в радиодиапазоне. Момент же истины наступил после того, как была измерена скорость замедления вращения пульсара в Крабовидной туманности $\dot{\Omega} = \delta\Omega/\delta t$ и стало ясно, что

1. скорость потери энергии вращающейся нейтронной звезды, определенной по скорости замедления угловой скорости вращения $W = -I_r \Omega \dot{\Omega}$, совпадает с величиной $W_{\text{tot}} \sim 5 \cdot 10^{38}$ эрг/с,
2. характерный возраст радиопулсара $\tau_D = 0.5\Omega/|\dot{\Omega}| \sim 1000$ лет совпадает с возрастом Крабовидной туманности, возникшей, как известно, во время взрыва исторической сверхновой AD 1054.

В результате, было окончательно установлено, что пульсирующее излучение радиопулсаров связано с вращением нейтронных звезд. При этом на всех очень сильное впечатление произвел тот факт, что открытие, за которое, кстати, в 1974 году была присуждена Нобелевская премия по физике, фактически, лежало на поверхности. Если бы анализ на переменность оптического излучения необычной звезды в центре Крабовидной туманности был проведен ранее, нейтронные звезды могли быть открыты даже до появления радиоастрономии.

Приведенный выше пример (на наш взгляд, наиболее наглядный, но далеко не единственный) показывает, насколько прогресс в создании новых приемников космического излучения, приведший бурному развитию радиоастрономии, а потом рентгеновской и гамма-астрономии, опережал развитие теории. С другой стороны, в релятивистской гравитационной астрономии (и, в частности, в космологии) теория с самого начала заметно опережала наблюдения. Можно сказать больше – многие десятилетия не существовало вообще никаких доказательств справедливости общей теории относительности! Дело в том, что хорошо известные с начала прошлого века примеры экспериментальной проверки ОТО (движение перигелия Меркурия, гравитационное красное смещение, отклонение света в гравитационном поле Солнца) соответствуют случаю слабых гравитационных полей, причем все эти эффекты могут быть объяснены и в рамках альтернативных теорий гравитации. До сих пор мы не имеем прямых доказательств справедливости общей теории относительности в сильных гравитационных полях. В частности, нет прямого доказательства существования черных дыр – объектов, предсказанных еще в начале прошлого

века – хотя косвенно они обнаружены в тесных двойных системах, а также в центрах активных галактик, и их существование в настоящее время является общепризнанным. Однако, эти объекты проявляют себя лишь как невидимые компактные объекты, на которые, в частности, происходит перетекание вещества. При этом гравитационное поле гипотетических черных дыр проявляет себя лишь на больших расстояниях по сравнению с гравитационным радиусом $r_g = 2GM/c^2$, где его уже очень трудно отличить от ньютоновского. Прямых же указаний на необычные эффекты, связанные с существованием горизонта событий (таких, например, как отсутствие твердой поверхности, полное замедление хода времени) до сих пор получено не было.

Соответственно, нет прямых наблюдений и предсказанных ОТО гравитационных волн (достаточно чувствительные приемники стали строиться лишь в последнее десятилетие), хотя и здесь имеются надежные косвенные данные. Анализ движения нейтронных звезд в тесных (орбитальный период $P_b < 10$ часов) двойных системах показал, что такие системы действительно теряют энергию за счет излучения гравитационных волн, причем согласно с формулой Эйнштейна для мощности потерь, как мы увидим, в настоящее время проверено с точностью долей процента.

Таким образом, до сих пор в гравитационной астрономии теория заметно опережает наблюдения. Казалось бы, тем более удивительно, как Эйнштейну удалось сформулировать свою теорию почти сто лет тому назад, когда были известны лишь немногочисленные примеры отклонения от ньютоновской нерелятивистской теории гравитации, а точность наблюдений была значительно ниже. В этом смысле история создания общей теории относительности оказывается чрезвычайно поучительной. Она показывает, что иногда бывает полезно полностью отрешиться от наблюдений и строить теорию, основываясь на некоторых ключевых принципах, известных, например, из других областей физики. И лишь затем проводить сравнение предсказаний теории с данными наблюдений. Так, закон сохранения энергии, сформулированный вначале лишь для механических движений, затем был распространен на все физические явления. Собственно, именно так была сформулирована и эвклидова геометрия, хотя, конечно же, она в значительной степени опиралась на повседневный опыт. Как мы видим, идея аксиоматического подхода при построении физических теорий имеет многовековую традицию.

По этой причине при обсуждении основных положений, лежащих в основе общей теории относительности, мы будем обращать особое внимание именно на то, какие основополагающие принципы (аксиомы) были положены в основу общей теории относительности, и как из этих принципов возникло стройное здание математически строгой теории релятивистской гравитации.

2.2 Математическое введение

Сделаем еще несколько замечаний математического характера. Как уже говорилось, мы постараемся сформулировать все ключевые соотношения на элементарном языке, т.е. без использования высшей математики. С другой стороны, совсем обойтись без могучего аппарата дифференциального исчисления было бы неправильно. Как уже отмечалось, высшая математика создавалась И. Ньютоном (1643-1727) и Г.В. Лейбницем (1646-1716) в значительной степени именно для того, чтобы законы физики можно было сформулировать просто и компактно. Фактически же ими был создан новый математический язык, позволяющий адекватно описывать явления природы. До этого же основным языком был формализм конических сечений; именно на нем был написан и основной труд Ньютона "Математические основы натуральной философии". Любопытно, что еще многие годы эти два языка существовали параллельно. Например, многие декабристы, как образованные люди своего времени, изучали конические сечения в университете. Однако несомненное преимущество нового языка постепенно привело к тому, что дифференциальное исчисление стало основным языком научной литературы.

Поэтому в некоторых случаях мы будем приводить и дифференциальные соотношения, которые, как можно надеяться, должны прояснить обсуждаемые вопросы для читателя, знакомого с высшей математикой. Интегралы же будут появляться лишь в исключительных случаях. Для тех же, кто еще не знаком с понятием производной, следует рассматривать "иероглиф" dx как малое приращение величины x , т.е. $dx \approx \delta x$, а "иероглиф" dx/dt как скорость изменения величины x при изменении t

$$\frac{dx}{dt} \approx \frac{\delta x}{\delta t}. \quad (1)$$

Так, уже появившаяся выше "скорость изменения угловой скорости вращения" $\dot{\Omega} = \delta\Omega/\delta t$ есть ни что иное как производная угловой скорости Ω по времени t : $\dot{\Omega} = d\Omega/dt$. При этом использование термина "скорость", конечно же, не случайно. Ведь всем хорошо известная физическая величина – скорость частицы v_x – и есть производная координаты частицы x по времени t : $v_x = dx/dt$. "Иероглиф" же $\partial f(x, y)/\partial x$ соответствует т.н. частной производной функции $f(x, y)$ по x , при вычислении которой нужно считать, что переменная y остается постоянной.

Наконец, везде ниже будет использоваться система СГСЭ. Именно эта система единиц наиболее естественна при описании электромагнитных явлений, которые постоянно будут обсуждаться ниже. В частности, в этой системе электрические и магнитные поля имеют одинаковую размерность. В настоящее время именно эта система единиц повсеместно используется в астрофизической литературе.

3 Общая теория относительности – одна из первых теорий поля

3.1 Ньютонская гравитация – пример нерелятивистской теории

Вспомним прежде всего, что нам известно из школьного курса физики. Это, конечно же, закон Всемирного тяготения, связывающий напряженность гравитационного поля \mathbf{g} и массу тяготеющего тела M

$$\mathbf{g} = -G \frac{M}{r^2} \mathbf{n}. \quad (2)$$

Здесь $G = 6.67 \cdot 10^{-8} \text{ см}^3 \text{ г}^{-1} \text{ с}^{-2}$ – универсальная гравитационная постоянная, а \mathbf{n} – единичный вектор, направленный от тела, которое мы для удобства поместили в начало координат. Это означает, что сила $\mathbf{F} = m\mathbf{g}$, действующая на пробное тело массы m , будет равна

$$\mathbf{F} = -G \frac{Mm}{r^2} \mathbf{n}, \quad (3)$$

а согласно третьему – что та же сила, взятая с обратным знаком, будет действовать на тело массы M . Здесь имеет место полная аналогия с электростатикой, когда сила \mathbf{F} , действующая на заряд q , должна быть записана в виде $\mathbf{F} = q\mathbf{E}$, где \mathbf{E} – напряженность электрического поля. Воспользовавшись теперь вторым законом Ньютона $\mathbf{F} = m\mathbf{a}$, получаем, что в присутствии гравитационного поля все тела будут двигаться с одинаковым ускорением

$$\mathbf{a} = \mathbf{g}. \quad (4)$$

Поэтому-то напряженность гравитационного поля \mathbf{g} обычно и называют ускорением свободного падения, хотя этот термин скорее относится к пробным телам, а не к самому гравитационному полю.

Напряженности гравитационного поля \mathbf{g} (2) соответствует скалярный гравитационный потенциал

$$\varphi_g = -\frac{GM}{r} \quad (5)$$

(размерность – квадрат скорости). Удобство этой величины состоит в том, что она представляет собой потенциальную энергию тела единичной массы. В результате, полная энергия частицы массы m может быть записана как

$$\mathcal{E} = \frac{mv^2}{2} + m\varphi_g. \quad (6)$$

При этом в сферически симметричном случае ускорение свободного падения $g(r) = g_r(r)$ и гравитационный потенциал $\varphi_g(r)$ связаны простым соотношением

$$g(r) = -\frac{d\varphi_g}{dr}. \quad (7)$$

На первый взгляд, все сказанное выше достаточно элементарно. Однако, уже здесь можно сделать несколько важных замечаний более общего характера.

1. Из уравнения (4) видно, что ускорение тела не зависит от его массы. Это стало возможным благодаря предположению (вообще говоря, совершенно не очевидному) о равенстве гравитационной массы m , входящей в уравнение (3), и инертной массы m , входящей во второй закон Ньютона $\mathbf{F} = m\mathbf{a}$. Этот принцип, как мы увидим, является одним из краеугольных камней при построении общей теории относительности.
2. Закон движения $\mathbf{a} = \mathbf{g}$ (4) выглядит одинаковым в любой инерциальной системе отсчета. Это связано с тем, что в любой инерциальной системе ускорение \mathbf{a} , определяемое через вторые производные координат по времени, как и величина силы \mathbf{F} (3), остаются теми же самыми как в покоящейся, так и в равномерно движущейся относительно нее системе координат. Действительно, вторая производная $d^2\mathbf{r}(t)/dt^2$ от линейной функции $\mathbf{r}(t) = \mathbf{v}t + \mathbf{r}_0$ всегда равна нулю. Как мы увидим, принцип инвариантности (независимости математической записи от выбора системы координат) станет еще одним ключевым положением при построении общей теории относительности.
3. Уравнения можно читать как слева направо (тяготеющие массы создают гравитационное поле), так и справа налево (существование гравитационного поля приводит к ускорению пробных частиц). Именно поэтому мы сформулировали закон всемирного тяготения в несколько необычной форме (2), а не (3), поскольку для нас здесь важным было именно то, что материальные тела создают вокруг себя гравитационное поле.
4. Гравитационное поле может быть описано как векторной величиной – ускорением свободного падения \mathbf{g} (2), так и скалярным гравитационным потенциалом φ_g (5). Однако поскольку величина \mathbf{g} однозначно определяется через потенциал φ_g (см. соотношение (7)), то в действительности в нерелятивистской теории гравитации необходимо задать лишь одну скалярную функцию – гравитационный потенциал $\varphi_g(\mathbf{r})$.

3.2 Почему потребовалось изменить теорию? Лоренц-инвариантность

Теория тяготения Ньютона является замкнутой (т.е. внутренне непротиворечивой) теорией, позволяющей, в частности, однозначно предсказывать движение тел. Как хорошо

известно, именно на основе ньютоновской теории получили свое естественное объяснение законы И. Кеплера (1571-1630), а также было предсказано существование внешних планет. Тем не менее, к началу XX века стало ясно, что классическая теория тяготения должна быть пересмотрена.

Конечно же, основной причиной, приведшей к необходимости пересмотреть основные положения классической теории, были наблюдения. Теория тяготения Ньютона не могла объяснить ни дополнительного движения перигелия Меркурия, ни отклонения света в поле Солнца (в последнем случае предсказание нерелятивистской теории и наблюдения отличались в два раза). С другой стороны, классическая теория тяготения перестала удовлетворять и ряду принципиально важных свойств, которые уже были известны на примере других теорий, а именно теории электромагнитного поля, сформулированные Д.К. Максвеллом (1831-1879), а также специальной теории относительности Эйнштейна. Так, например, оказалось, что уравнения Ньютона не инвариантны по отношению к преобразованиям, открытыми Х.А. Лоренцом (1853-1928), которые изменяют не только пространственные, но и временные координаты.

Напомним, что в начале XX века лоренц-инвариантность представлялась как универсальное требование к фундаментальной теории. Во всяком случае, справедливость этих преобразований были экспериментально проверены для электромагнитных, а впоследствии для сильных и слабых взаимодействий. Однако, как мы увидим, для гравитационного взаимодействия от условия лоренц-инвариантности приходится отказаться. Тем не менее, поскольку условие лоренц-инвариантности сыграло ключевую роль при построении теории относительности, представляется целесообразным потратить некоторое время на то, чтобы сформулировать более точно, что мы в дальнейшем будем понимать под инвариантностью теории.

Фактически, мы уже сталкивались со свойством инвариантности на примере второго закона Ньютона. Независимо от системы отсчета (которая, однако, обязательно должна быть инерциальной), закон движения имеет вид $\mathbf{a} = \mathbf{g}$. При этом было выполнено даже более сильное условие – и ускорение свободного падения \mathbf{g} , и ускорение \mathbf{a} не изменялись при переходе от одной инерциальной системы к другой.

Прежде всего, напомним, что любая инвариантная величина (число спутников у планет, число нуклонов в ядре, число студентов, принятых в институт) вообще не должна зависеть от выбора системы координат. С другой стороны, скалярная величина, например, гравитационный потенциал $\varphi_{\mathbf{g}}(\mathbf{r})$, зависит лишь от положения точки, а не от того, как через эту точку проходят координатные линии. Если же мы имеем дело с вектором (направленным отрезком) \mathbf{A} , то его координаты A_x, A_y, A_z , т.е. проекции вектора \mathbf{A} на оси x, y, z , конечно же, будут зависеть от направления этих осей. Тем не менее, изменение координат должно происходить по вполне определенным законам. Так, например,

при поворотах вокруг оси z на угол φ (см. Рис. 1), когда новые и старые координаты некоторой точки будут связаны соотношениями

$$x' = x \cos \varphi + y \sin \varphi; \quad y' = -x \sin \varphi + y \cos \varphi; \quad z' = z, \quad (8)$$

координаты вектора A_x, A_y, A_z в одной системе координат и его координаты $A_{x'}, A_{y'}, A_{z'}$ в другой системе должны быть связаны соотношениями

$$A_{x'} = A_x \cos \varphi + A_y \sin \varphi; \quad A_{y'} = -A_x \sin \varphi + A_y \cos \varphi; \quad A_{z'} = A_z. \quad (9)$$

Именно должны, так как в противном случае объект \mathbf{A} , задаваемый его тремя координатами, вообще нельзя считать вектором. Иными словами, соотношения (9) есть определение, а не свойство вектора \mathbf{A} . При этом важно, что при преобразовании координат некоторые величины оказываются инвариантными. Прежде всего, это длина вектора $|\mathbf{A}| = \sqrt{\mathbf{A}^2}$, где

$$\mathbf{A}^2 = A_x^2 + A_y^2 + A_z^2, \quad (10)$$

которая остается постоянной при произвольном выборе ортогональных осей. К другим инвариантам можно отнести скалярное произведение двух векторов

$$(\mathbf{A} \cdot \mathbf{B}) = A_x B_x + A_y B_y + A_z B_z. \quad (11)$$

Действительно, эта величина может быть определена независимым от координатной сетки способом

$$(\mathbf{A} \cdot \mathbf{B}) = |\mathbf{A}| \cdot |\mathbf{B}| \cdot \cos \theta, \quad (12)$$

где θ – угол между векторами, который вообще не требует знания координат.

Если же одна система координат движется относительно другой со скоростью V вдоль оси x , так что координаты некоторой точки преобразуются как

$$x' = x - Vt; \quad y' = y; \quad z' = z, \quad (13)$$

то в этом случае координаты вектора вообще не будут изменяться (переменной будет лишь начало вектора). Напомним, что вместе с очевидным на первый взгляд предположением об инвариантности хода времени $t' = t$ преобразование (13) носит имя Г. Галилея (1564-1642).

Запишем теперь преобразования Лоренца, т.е. преобразования координат и времени для двух систем, движущихся относительно друг друга с постоянной скоростью V вдоль оси x . При этом мы сразу запишем эти преобразования для бесконечно малых промежутков времени dt и смещений $d\mathbf{r}$

$$dt' = \frac{dt - V dx/c^2}{\sqrt{1 - V^2/c^2}}; \quad dx' = \frac{dx - V dt}{\sqrt{1 - V^2/c^2}}; \quad dy' = dy; \quad dz' = dz. \quad (14)$$

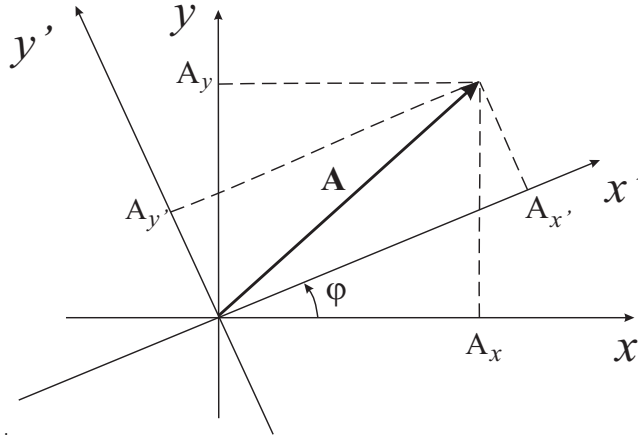


Рис. 1: Изменение координат вектора при повороте системы координат на угол φ

Сравнивая теперь соотношения (8) и (14), можно заметить, что преобразования Лоренца и преобразования координат при повороте координатных осей имеют много общего. Действительно, воспользовавшись известными тригонометрическими тождествами, можно записать преобразование (8) как

$$dx' = \frac{dx + \operatorname{tg}\varphi dy}{\sqrt{1 + \operatorname{tg}^2\varphi}}; \quad dy' = \frac{dy - \operatorname{tg}\varphi dx}{\sqrt{1 + \operatorname{tg}^2\varphi}}; \quad dz' = dz, \quad (15)$$

Как мы видим, коэффициенты преобразования переходят друг в друга при замене $-V^2/c^2$ на $\operatorname{tg}^2\varphi$. Поэтому преобразования Лоренца также называют поворотами. Но только не в обычном пространстве, где при поворотах остается инвариантной длина вектора (10), т.е. сумма трех квадратов его координат. А в четырехмерном пространстве-времени, в котором "квадрат длины", который носит название интервала, определяется как

$$(ds)^2 = -c^2(dt)^2 + (dr)^2. \quad (16)$$

Именно интервал ds между двумя точками в четырехмерном пространстве-времени остается инвариантным при преобразованиях Лоренца. Впервые геометрическая интерпретация плоского пространства-времени была сформулирована Г. Минковским (1864-1909).

Ясно, что изменение системы отсчета вновь должно привести и к изменению координат векторов, соответствующих различным физическим величинам. Однако, как и в случае обычных поворотов, преобразования Лоренца сохраняют инвариантными некоторые величины. В частности, оказывается, что при преобразованиях Лоренца остается постоянной величина $\mathcal{E}^2 - c^2\mathbf{p}^2 = m^2c^4$, где

$$\mathcal{E} = \frac{mc^2}{\sqrt{1 - v^2/c^2}} \quad (17)$$

есть энергия, а

$$\mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1 - v^2/c^2}} \quad (18)$$

– импульс частицы. Таким образом, энергия и импульс комбинируются также, как и промежуток времени dt и расстояние $d\mathbf{r}$. Следовательно, можно сделать и более сильное утверждение: хорошо известные нам величины – энергия (трехмерный скаляр) и импульс (вектор) в действительности являются четырьмя компонентами четырехмерного вектора энергии-импульса, ”длина” которого сохраняется при преобразованиях Лоренца. Поэтому не удивительно, что сами энергия и импульс при переходе от одной системы координат к другой преобразуются подобно тому, как преобразуются время t и координата \mathbf{r} :

$$\mathcal{E}' = \frac{\mathcal{E} - Vp_x}{\sqrt{1 - V^2/c^2}}, \quad p'_x = \frac{p_x - V\mathcal{E}/c^2}{\sqrt{1 - V^2/c^2}}, \quad p'_y = p_y, \quad p'_z = p_z. \quad (19)$$

Теперь мы можем более точно сформулировать, что означает лоренц-инвариантность теории. Прежде всего, отметим, что содержание любой теории не должно зависеть от того, ”под каким углом” мы на нее смотрим. Поэтому особую роль приобретают инвариантные величины, не зависящие от выбора системы координат. Даже если для описания какого-либо физического процесса нам приходится ввести векторы, всегда бывает полезно определить скалярные величины, имеющие ясный инвариантный физический смысл. По-видимому, это связано с тем, что глубинная природа физических законов имеет глобальную (интегральную), а не локальную (дифференциальную) природу.

Например, закон сохранения энергии означает, что изменение энергии в некотором объеме обязательно должно сопровождаться потоком энергии через поверхность, ограничивающий этот объем. В такой формулировке ни величина скорости изменения энергии в заданном объеме, ни полный поток энергии через поверхность, ограничивающий это объем, не зависят от выбора системы координат. Важно лишь, что это свойство должно быть справедливо для любого объема. Соответственно, в ньютоновской механике вектор скорости изменяется при движении в гравитационном поле, однако энергия тела остается постоянной. Иногда это свойство позволяет сразу получить ответ. Например, для определения высоты, на которую поднимется брошенный вертикально вверх камень, достаточно записать закон сохранения энергии, а не решать уравнения движения.

Важнейшим инвариантным свойством обладает и только что введенный нами интервал $ds^2 = -c^2dt^2 + d\mathbf{r}^2$. Действительно, пусть луч света был испущен в момент времени t_1 в точке \mathbf{r}_1 и поглощен в момент времени t_2 в точке \mathbf{r}_2 . Поскольку свет распространяется со скоростью света, то интервал (четырёхмерное расстояние между этими двумя пространственно-временными точками) будет равным нулю

$$(\delta s)^2 = -c^2(\delta t)^2 + (\delta \mathbf{r})^2 = 0. \quad (20)$$

Здесь $\delta t = t_2 - t_1$ и $\delta \mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$. Но, как мы видели, интервал не изменяется при преобразованиях Лоренца, и поэтому в любой другой инерциальной системе отсчета интервал между двумя событиями – излучением и поглощением света – будет равен нулю. А это и означает, что скорость света не зависит от выбора системы координат.

Наконец, в теории Максвелла при преобразованиях Лоренца инвариантными остаются скаляры $I_1 = \mathbf{B}^2 - \mathbf{E}^2$ и $I_2 = (\mathbf{E} \cdot \mathbf{B})$ (напоминаем, что здесь используется система СГСЭ). Поэтому если мы имеем дело с электромагнитной волной, в которой электрическое поле \mathbf{E} равно по величине магнитному полю \mathbf{B} , но перпендикулярно ему по направлению, так что $I_1 = I_2 = 0$, то это инварианты останутся нулевыми и в любой другой системе координат. А это значит, что волна остается волной в любой инерциальной системе. Таким образом, необходимым условием для справедливости утверждения – электромагнитное поле является волной – будут условия $I_1 = I_2 = 0$ (и, конечно же, $E \neq 0$, $B \neq 0$), одинаковые в любой инерциальной системе координат.

Различие интегрального и дифференциального подхода полезно проиллюстрировать и на примере теории Всемирного тяготения Ньютона. Действительно, гравитационный потенциал $\varphi_g = -GM/r$ (5) (и, соответственно, напряженность гравитационного поля $g_r = -GM/r^2$) на самом деле должны определяться как решение уравнения на гравитационное поле. В дифференциальной форме (и для центрально-симметричного распределения вещества, когда все величины зависят лишь от радиуса r), это уравнение имеет вид

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d\varphi_g}{dr} \right) = 4\pi G \rho_m, \quad (21)$$

или для ускорения свободного падения $g = g_r$

$$\frac{1}{r^2} \frac{d}{dr} (r^2 g) = -4\pi G \rho_m. \quad (22)$$

Здесь ρ_m – плотность вещества. Поэтому за пределами гравитирующего тела, где плотность вещества равна нулю, уравнение (22) сводится к уравнению

$$\frac{dg}{dr} + 2\frac{g}{r} = 0. \quad (23)$$

Умеющие дифференцировать могут легко проверить, что решением этого уравнения может быть любая функция вида C/r^2 , где C – постоянная. Однако коэффициент пропорциональности C , т.е. масса M , из уравнения (23) определена быть не может.

В интегральной же форме уравнение (22) гласит, что поток $\Phi = \sum \mathbf{g} \delta \mathbf{S}$ вектора \mathbf{g} через любую замкнутую поверхность равен суммарной массе вещества, заключенной в пределах этой поверхности, умноженной на $-4\pi G$

$$\Phi = -4\pi G M \quad (24)$$

(здесь $\delta\mathbf{S}$ – элемент площади; естественно, величина $\delta\mathbf{S}$ есть вектор, поскольку ориентация поверхности определяется вектором нормали). Это есть замечательная теорема Гаусса, справедливая в частности и для электромагнитных явлений; ее формулировка не зависит от системы координат. Для рассматриваемого здесь центрально-симметричного поля естественно определить поток Φ через сферу радиуса r . В этом случае он сводится просто к произведению площади сферы $4\pi r^2$ на величину вектора $g = |\mathbf{g}|$, причем взятый со знаком минус, поскольку вектор \mathbf{g} направлен по направлению к центру сферы. В итоге, мы мгновенно возвращаемся к выражению (2).

Здесь полезно сделать одно историческое замечание. Ньютон, сравнивая ускорение Луны при ее движении по орбите с ускорением свободного падения на Земле, легко оценил, что сила гравитации уменьшается обратно пропорционально квадрату расстояния. Однако он двенадцать лет не обнаружил свой результат, поскольку ему не было очевидно, что геометрическая сумма сил, действующая со стороны разных точек земного шара на тело, находящееся на ее поверхности, в точности равна силе тяжести, если бы вся масса Земли была бы сосредоточена в ее центре. Лишь после создания дифференциального исчисления, позволившего, в частности, доказать справедливость соотношения (24), этот результат стал очевидным.

Таким образом, в интегральной форме законы физики действительно можно сформулировать в виде, не зависящем от выбора системы отсчета. Однако изначально не очевидно, что математическая запись физического закона в его дифференциальной форме окажется одинаковой в любой системе координат. Условие же инвариантности как раз и состоит в том, что математическая запись физического закона в дифференциальной форме не должна зависеть от выбора координатной сетки (напомним, что в классической механике ускорение не зависело от выбора инерциальной системы координат). В частности, лоренц-инвариантность теории означает, что в любых двух системах, в которых пространственные и временные координаты связаны соотношениями Лоренца, буквенное выражение закона должно оставаться неизменным.

Так, например, если в одной системе координат электрическое поле равно нулю, а магнитное – нет, то в другой системе координат электрическое поле уже будет отличаться от нуля. Однако сами законы электромагнитного поля (а именно, их математическая запись) остаются одинаковыми в любой системе координат. В частности, остаются инвариантными утверждения, что источниками электрического поля являются заряды, а магнитное поле является вихревым, т.е. силовые линии либо замкнуты, либо уходят на бесконечность. Другой пример: в любой системе отсчета закон движения частицы в электромагнитном поле будет иметь вид

$$\mathbf{F} = e\mathbf{E} + e\frac{\mathbf{v}}{c} \times \mathbf{B}, \quad (25)$$

хотя все величины, входящие в это уравнение – импульс \mathbf{p} и время t , определяющие

силу $\mathbf{F} = \delta\mathbf{p}/\delta t$, а также электрическое и магнитное поле – будут различными в разных системах отсчета. Ясно, что такое удивительное свойство может быть справедливо далеко не при всяком преобразовании координат. Замечательный же факт состоит именно в том, что форма уравнения (25), определяющего движение заряженной частицы в теории электромагнетизма, остается неизменной тогда и только тогда, когда мы имеем дело с преобразованием Лоренца.

Теперь становится понятным, что имелось в виду, когда мы говорили о том, что ньютоновская теория гравитации не является лоренц-инвариантной. Сам внешний вид дифференциальных уравнений движения начинает зависеть от скорости инерциальной системы. Действительно, если мы запишем, например, x -компоненту уравнения движения в однородном гравитационном поле \mathbf{g} , направленном вдоль оси x

$$\frac{\delta p_x}{\delta t} = mg, \quad (26)$$

то при переходе к другой системе отсчета, движущейся со скоростью V вдоль оси x , оно приобретает вид

$$\frac{\delta p'_x + V\mathcal{E}/c^2}{\delta t' + V\delta x'/c^2} = mg, \quad (27)$$

что не совпадает с ожидаемым значением

$$\frac{\delta p'_x}{\delta t'} = mg. \quad (28)$$

Таким образом, ньютоновская теория гравитации требовала какого-то существенного изменения. Еще раз напомним, что в начале XX века существовали лишь несколько наблюдательных фактов, противоречащих классической ньютоновской теории, так что эксперимент не мог стать надежным фундаментом при построении новой теории. Поэтому Эйнштейн предпринял попытку подойти к проблеме построения теории с другой стороны, а именно со стороны некоторых общих принципов, которым должна удовлетворять последовательная теория. Именно в этом и заключался прорыв в физику нового столетия.

Действительно, удача, которая сопутствовала Эйнштейну, показала, что для понимания физических явлений (например, для интерпретации наблюдений в астрономии) не всегда необходимо отталкиваться от наблюдений. Иногда ключевые (и неизвестные заранее) физические свойства могут быть получены (выведены) из общих принципов, справедливость которых проверена в других областях науки. В настоящее время подобный дедуктивный метод, к обсуждению которого мы и переходим, и является основным при построении современных физических теорий.

3.3 Кинетическая энергия

Итак, общая теория относительности была, фактически, первой последовательной теорией, построенной исходя из некоторых общих принципов, а не на основе анализа наблюдений. Однако внутренняя непротиворечивость теории, конечно же, не может служить достаточным аргументом в пользу ее справедливости. Известные же примеры экспериментальной проверки ОТО, как мы увидим, могут быть объяснены и в рамках альтернативных теорий гравитации.

Ключевыми словами в предыдущем абзаце были "последовательная теория" и "альтернативные модели". Они наводят на мысль о том, что,

1. правильная теория должна удовлетворять некоторому набору фундаментальных свойств (аксиом),
2. общая теория относительности на самом деле не является единственно возможной теорией гравитации (обобщая – теорией поля).

Иными словами, возникает вопрос, можно ли определить вид теории (т.е. вид уравнений, описывающих ее основные законы), исходя лишь из общих принципов, т.е. полностью отвлекаясь от наблюдений.

Чтобы попытаться ответить на этот вопрос, мы сделаем несколько шагов назад, вернувшись к казалось бы очевидному вопросу о зависимости кинетической энергии \mathcal{E}_{kin} от скорости v и массы m в нерелятивистской теории. При этом мы на время должны забыть, что мы знаем релятивистское обобщение – известное соотношение $\mathcal{E}_{\text{kin}} = mc^2/\sqrt{1 - v^2/c^2}$ (17).

Как хорошо известно, в ньютоновской механике $\mathcal{E}_{\text{kin}} = mv^2/2$, так что для одномерного движения

$$\mathcal{E}_{\text{kin}} = \frac{mv_x^2}{2}. \quad (29)$$

Поставим теперь несколько неожиданный вопрос: является ли это выражение действительно единственно возможным?

Прежде всего, хотелось бы, чтобы правильная (точнее, более общая) теория при малых скоростях и массах совпадала с ньютоновским пределом, справедливость которого подтверждено многочисленными экспериментами. Это значит, что речь идет о поправках к выражению (29), которые начинают играть заметную роль при больших v и m . Иными словами, мы попытаемся понять, какие слагаемые можно добавить к классическому выражению (29) не нарушая при этом общих принципов.

Прежде всего, посмотрим, можно ли добавить к классическому выражению (29) слагаемые, пропорциональные m^2 . Конечно, при этом необходимо быть внимательным в отношении размерности, поскольку величина m^2v^2 не имеет размерности энергии. Эта

проблема, однако, может быть легко разрешена. Для этого (внимание, важное место!) необходимо ввести в теорию дополнительный параметр M размерности массы. В результате, можно записать

$$\mathcal{E}_{\text{kin}} = \frac{mv_x^2}{2} + \frac{m^2v_x^2}{M}. \quad (30)$$

Это выражение начинает существенно отличаться от классического лишь при достаточно больших массах частиц, так что при $m \ll M$ предсказания новой теории не будут существенно отличаться от предсказаний классической механики Ньютона. Однако, как легко проверить, выражение (30) нарушает другое важное свойство – сумма энергий двух тел массы m не будет равна энергии тела с удвоенной массой. А с этим уже никак нельзя согласиться, поскольку энергия не может зависеть от того, в какой последовательности мы считаем полную энергию тела. Как мы видим, этот путь оказывается тупиковым, по крайней мере в такой простой модели.

Попробуем теперь добавить в выражение для энергии величину, пропорциональную v_x^3 . Здесь нам потребуется уже константа C с размерностью скорости. В результате, получаем

$$\mathcal{E}_{\text{kin}} = \frac{mv_x^2}{2} + \frac{mv_x^3}{C}. \quad (31)$$

Однако эта формула для энергии тела также обладает существенным недостатком, поскольку добавка к классическому выражению меняет знак при повороте оси x на 180° , т.е. при замене v_x на $-v_x$. Ясно, что энергия не должна зависеть от выбора направления координатных осей, и поэтому энергия не может содержать слагаемые, пропорциональные v^3 .

С другой стороны, слагаемые, пропорциональные v_x^4/C^2 , могут быть безболезненно добавлены к нерелятивистскому выражению (29). При этом константа C (которая, вообще говоря, может быть произвольной) задает нам характерную скорость, при которой начинаются заметные отличия от нерелятивистской механики. Вспоминая теперь релятивистское выражение (17) для энергии, которое при достаточно малых скоростях может быть записано в виде

$$\mathcal{E}_{\text{kin}} = mc^2 + \frac{mv_x^2}{2} + \frac{3}{8} \frac{mv_x^4}{c^2} + \dots, \quad (32)$$

мы приходим к заключению, что в действительности $C = \sqrt{8/3}c$, где c – скорость света.

Обсудим теперь дополнительные возможности, связанные с трехмерностью нашего пространства. На первый взгляд, естественно было бы просто заменить квадрат скорости v_x^2 на полный квадрат $v^2 = v_x^2 + v_y^2 + v_z^2$, тем более что, как уже отмечалось, квадрат вектора не зависит от выбора системы отсчета. Однако оказывается, что на этом пути появляются гораздо более широкие возможности. Действительно, в общем случае даже для нерелятивистского движения, когда кинетическая энергия тела пропорциональна

квадрату скорости, энергия в общем случае может быть записана в виде

$$\begin{aligned} \mathcal{E}_{\text{kin}} = & \frac{m_{xx}v_x^2}{2} + \frac{m_{xy}v_xv_y}{2} + \frac{m_{xz}v_xv_z}{2} + \frac{m_{yx}v_xv_y}{2} + \frac{m_{yy}v_y^2}{2} \\ & + \frac{m_{yz}v_yv_z}{2} + \frac{m_{zx}v_xv_z}{2} + \frac{m_{zy}v_yv_z}{2} + \frac{m_{zz}v_z^2}{2}. \end{aligned} \quad (33)$$

Как мы видим, энергия частицы теперь зависит, вообще говоря, от девяти постоянных размерности массы: m_{xx} , m_{xy} , m_{xz} , m_{yx} , m_{yy} , m_{yz} , m_{zx} , m_{zy} и m_{zz} . Однако поскольку в выражении (33) фактически нет разницы между слагаемыми, содержащими v_xv_y и v_yv_x , v_xv_z и v_zv_x , v_yv_z и v_zv_y , то можно положить $m_{xy} = m_{yx}$, $m_{xz} = m_{zx}$ и $m_{yz} = m_{zy}$. Поэтому в общем случае приходится ввести лишь шесть независимых констант, которые удобно записать в виде таблицы, симметричной относительно диагонали

$$\langle m \rangle = \begin{pmatrix} m_{xx} & m_{xy} & m_{xz} \\ m_{xy} & m_{yy} & m_{yz} \\ m_{xz} & m_{yz} & m_{zz} \end{pmatrix}. \quad (34)$$

При этом обычное выражение для энергии будет соответствовать случаю $m_{xx} = m_{yy} = m_{zz} = m$ и $m_{xy} = m_{xz} = m_{yx} = m_{yz} = m_{zx} = m_{zy} = 0$, так что

$$\langle m \rangle = \begin{pmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{pmatrix}. \quad (35)$$

На первый взгляд, выражение (33) для энергии частиц выглядит несколько искусственным. Во всяком случае, ясно, что в большинстве случаев матрица (34) должна быть диагональной, так что мы возвращаемся к обычному выражению (32). Тем не менее, в некоторых случаях такое обобщение оказывается неизбежным. Например, если мы хотим описать движение электрона, движущегося в кристаллической решетке. Поскольку свойства кристаллической решетки зависят от направления (такая среда называется анизотропной), то за счет взаимодействия электрона с атомами решетки его динамические свойства (попросту говоря, способность сопротивляться внешней силе) начинают зависеть от направления движения. Это и приводит к необходимости ввести уже шесть величин с размерностью массы, характеризующих инерционные свойства частицы¹. При этом чрезвычайно важно, что вид таблицы (матрицы) (34) зависит от внутренней структуры решетки. Иными словами, зная строение элементарных ячеек, из которых состоит кристалл, можно определить количество независимых параметров, определяющих инерционные свойства электрона.

¹Страшно сказать, но при определенных условиях некоторые компоненты тензора масс могут быть даже отрицательными.

Например, для простейшей решетки, состоящей из одинаковых кубических ячеек, "тензор массы", как и в пустом пространстве, становится диагональным с одинаковыми значениями m_{xx} , m_{yy} и m_{zz} , и, следовательно, мы можем обойтись лишь одной постоянной размерности массы. А это изначально совершенно не очевидно, поскольку механические свойства кубического кристалла могли бы зависеть от направления приложения силы. Чтобы показать это, рассмотрим более простой двумерный случай. Иными словами, нам нужно показать, что для двумерного кристалла, состоящего из квадратных ячеек, двумерный "тензор массы"

$$\langle m \rangle = \begin{pmatrix} m_{xx} & m_{xy} \\ m_{yx} & m_{yy} \end{pmatrix}, \quad (36)$$

имеет диагональный вид с равными значениями масс, стоящих на диагонали: $m_{xx} = m_{yy}$, $m_{xy} = m_{yx} = 0$.

Здесь мы должны принять во внимание, что точно также, как в двумерном случае не любые два числа являются компонентами вектора, не любые четыре числа, записанные в таблице 2×2 , являются компонентами тензора. Они обязаны удовлетворять определенным соотношениям при повороте координатной сетки на угол φ :

$$\begin{aligned} m_{x'x'} &= m_{xx} \cos^2 \varphi + m_{xy} \sin \varphi \cos \varphi + m_{yx} \sin \varphi \cos \varphi + m_{yy} \sin^2 \varphi, \\ m_{x'y'} &= -m_{xx} \sin \varphi \cos \varphi + m_{xy} \cos^2 \varphi - m_{yx} \sin^2 \varphi + m_{yy} \sin \varphi \cos \varphi, \\ m_{y'x'} &= -m_{xx} \sin \varphi \cos \varphi - m_{xy} \sin^2 \varphi + m_{yx} \cos^2 \varphi + m_{yy} \sin \varphi \cos \varphi, \\ m_{y'y'} &= m_{xx} \sin^2 \varphi - m_{xy} \sin \varphi \cos \varphi - m_{yx} \sin \varphi \cos \varphi + m_{yy} \cos^2 \varphi. \end{aligned} \quad (37)$$

Кстати, становится понятно, почему для изотропной среды, когда вид тензора массы вообще не должен зависеть от угла φ , он также должен быть пропорционален единичной матрице. Действительно, лишь в случае $m_{xx} = m_{yy} = m$, $m_{xy} = m_{yx} = 0$ получаем тождественно при любом угле φ

$$\begin{aligned} m_{x'x'} &= m, \\ m_{x'y'} &= 0, \\ m_{y'x'} &= 0, \\ m_{y'y'} &= m. \end{aligned}$$

Но оказывается, что для диагональности тензора массы достаточно и того, чтобы ничего не изменилось при повороте на 90° , т.е. при условии $\cos \varphi = 0$, $\sin \varphi = 1$. Действительно, имеем в этом случае

$$\begin{aligned} m_{x'x'} &= m_{yy}, \\ m_{x'y'} &= -m_{yx}, \\ m_{y'x'} &= -m_{xy}, \\ m_{y'y'} &= m_{xx}. \end{aligned}$$

Учитем теперь тот факт, что на самом деле при повороте на 90° наша "решетка" совпадает с той, которая была до поворота. Поэтому тензор массы должен быть в точности таким же, как и до поворота. В результате, получаем вновь условие $m_{xx} = m_{yy} = m$, $m_{xy} = m_{yx} = 0$ (здесь мы воспользовались свойством, что тензор массы является симметричным).

Из сказанного выше можно сделать несколько важных выводов.

1. Общие принципы (симметрия, лоренц-инвариантность) могут помочь ограничить теорию, но в общем случае не определяют ее до конца.
2. При расширении в теорию приходится вводить размерные константы (масса M , скорость c), величины которых могут быть определены только из наблюдений.
3. В предельном случае (в рассмотренном выше примере – при нерелятивистских скоростях $v \ll c$) теория должна сводиться к известной.
4. Одно из возможных обобщений – переход от скаляров (чисел) к тензорам (таблицам).

3.4 Появляются тензоры

3.4.1 Трехмерное пространство

Появление нового объекта – (симметричных) таблиц-матриц, которые в отличие от векторов мы будем обозначать как $\langle m \rangle$, $\langle \sigma \rangle$, *etc.*, в конце предыдущего раздела было не случайным. Более того, как мы увидим, именно используя этот язык удастся достаточно просто сформулировать основные положения общей теории относительности. Поэтому и на этом вопросе следует остановиться более подробно.

Прежде всего, введение симметричных матриц оказывается полезным в тех случаях, когда мы имеем дело с квадратичными зависимостями. Так, в декартовых координатах инвариантный квадрат длины бесконечно малого вектора $(d\mathbf{r})^2$ может быть выражен через любой набор координат. Например в декартовых координатах x, y, z (Рис. 2а) имеем

$$(d\mathbf{r})^2 = (dx)^2 + (dy)^2 + (dz)^2. \quad (38)$$

В цилиндрических координатах r, φ, z (Рис. 2б)

$$(d\mathbf{r})^2 = (dr)^2 + r^2(d\varphi)^2 + (dz)^2. \quad (39)$$

Наконец, в сферических координатах r, θ, φ (Рис. 2в)

$$(d\mathbf{r})^2 = (dr)^2 + r^2(d\theta)^2 + r^2 \sin^2 \theta (d\varphi)^2. \quad (40)$$

Как мы видим, во всех случаях квадрат длины отрезка складывается из суммы квадратов координат, причем во всех случаях координатная сетка является ортогональной, т.е. координатные линии пересекаются под прямым углом. В результате, каждому выбору координат можно сопоставить диагональную таблицу коэффициентов, а именно

$$\langle g \rangle_D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (41)$$

для декартовой системы координат,

$$\langle g \rangle_{\text{cyl}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (42)$$

для цилиндрической, и

$$\langle g \rangle_{\text{sph}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2 \theta \end{pmatrix} \quad (43)$$

для сферической системы координат.

Более того, если имеется любая, не обязательно ортогональная сетка координат p, q, h , то ей будет соответствовать уже не диагональная, но по-прежнему симметричная матрица коэффициентов

$$\langle g \rangle = \begin{pmatrix} g_{pp} & g_{pq} & g_{ph} \\ g_{pq} & g_{qq} & g_{qh} \\ g_{ph} & g_{qh} & g_{hh} \end{pmatrix}, \quad (44)$$

так что

$$\begin{aligned} (d\mathbf{r})^2 &= g_{pp}(dp)^2 + 2g_{pq}(dp)(dq) + g_{qq}(dq)^2 + g_{hh}(dh)^2 \\ &\quad + 2g_{pq}(dp)(dq) + 2g_{ph}(dp)(dh) + 2g_{qh}(dq)(dh). \end{aligned} \quad (45)$$

Например, для "косой линейки" (см. Рис. 3), у которой

$$x' = x - y; \quad y' = y; \quad z' = z, \quad (46)$$

так что

$$(d\mathbf{r})^2 = (dx' + dy')^2 + (dy')^2 + (dz')^2 = (dx')^2 + 2(dy')(dx') + 2(dy')^2 + (dz')^2, \quad (47)$$

получаем

$$\langle g' \rangle = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (48)$$

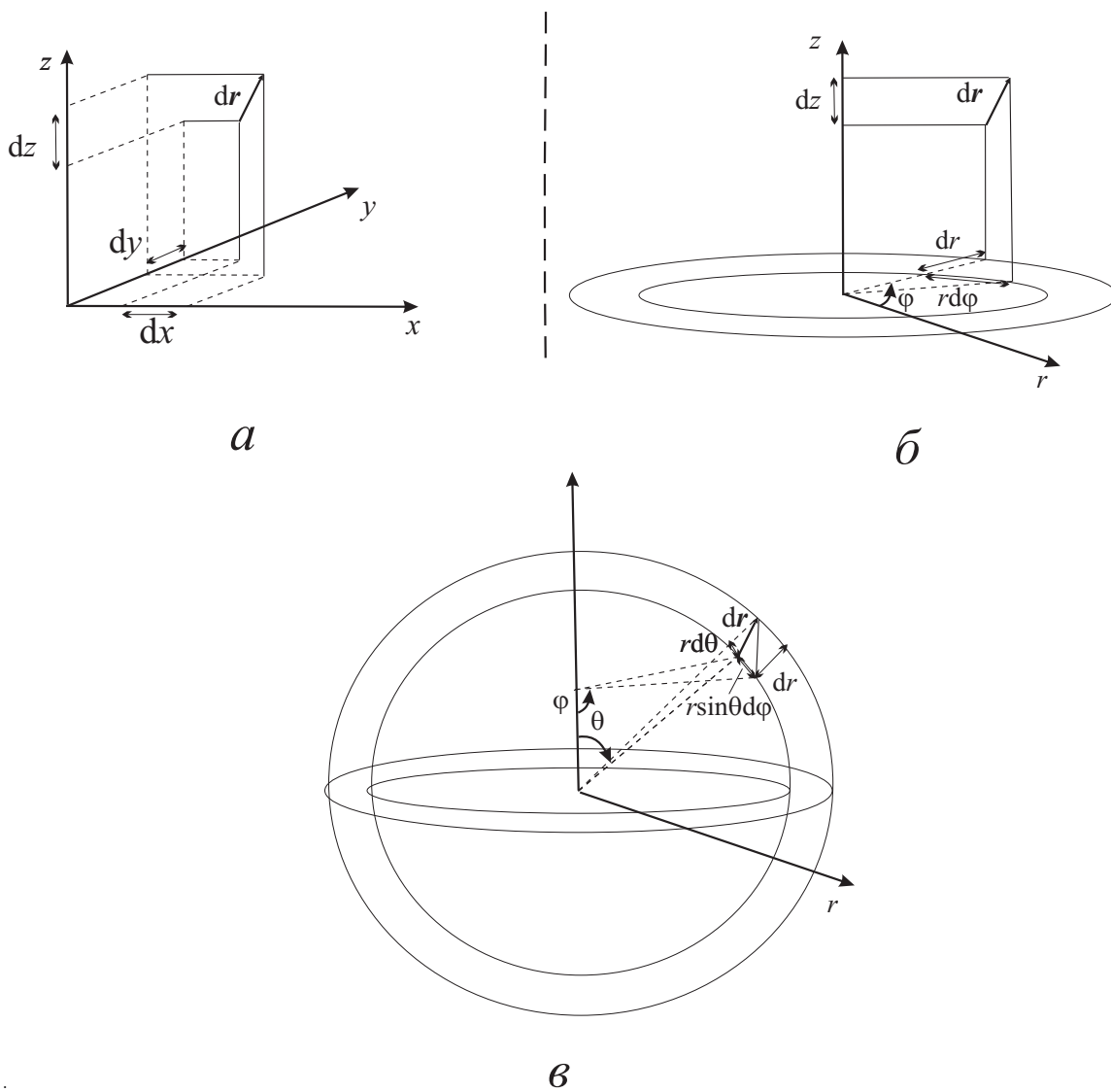


Рис. 2: Элемент длины dr в декартовых (а), цилиндрических (б) и полярных (в) координатах

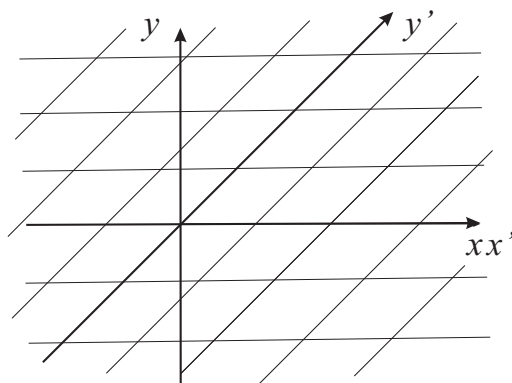


Рис. 3: "Косая линейка" как пример неортогональной системы координат

Но в любом случае выбор координатной сетки однозначно фиксирует и матрицу коэффициентов. В дальнейшем для простоты мы будем часто опускать скобки в подобных выражениях, т.е. писать ds^2 вместо $(ds)^2$.

Появившаяся таблица коэффициентов $\langle g \rangle$, определяющая расстояние между двумя близкими точками, как мы увидим, и играет определяющую роль в теории гравитации. Она называется метрическим тензором и фактически является основным действующим лицом при построении теории гравитационного поля.

Другими примерами, в которых оказывается необходимым ввести тензорные величины, могут служить физические законы, связывающие две векторные величины, например закон Гука $\mathbf{F} = k \cdot \delta \mathbf{r}$ или же закон Ома $\mathbf{j} = \sigma \mathbf{E}$. Выписанные выше соотношения предполагают, что векторы механического смещения $\delta \mathbf{r}$ и электрического тока \mathbf{j} параллельны соответственно приложенной силе \mathbf{F} и внешнему электрическому полю \mathbf{E} . Однако в общем случае это не так, так что, например, закон Ома следует записать в виде

$$\begin{aligned} j_x &= \sigma_{xx}E_x + \sigma_{xy}E_y + \sigma_{xz}E_z, \\ j_y &= \sigma_{yx}E_x + \sigma_{yy}E_y + \sigma_{yz}E_z, \\ j_z &= \sigma_{zx}E_x + \sigma_{zy}E_y + \sigma_{zz}E_z. \end{aligned} \quad (49)$$

Как мы видим, и в этом случае коэффициенты удобно записать в виде квадратной таблицы. Например, ясно, что для диагональной матрицы

$$\langle \sigma \rangle = \begin{pmatrix} \sigma_0 & 0 & 0 \\ 0 & \sigma_0 & 0 \\ 0 & 0 & \sigma_0 \end{pmatrix} \quad (50)$$

закон Ома сводится к обычному выражению $\mathbf{j} = \sigma_0 \mathbf{E}$, так что вектор электрического тока будет параллелен вектору электрического поля. Если же тензор $\langle \sigma \rangle$ имеет вид

$$\langle \sigma \rangle = \begin{pmatrix} 0 & \sigma_1 & 0 \\ \sigma_1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (51)$$

то при приложении к образцу электрического поля вдоль оси x результирующий ток будет направлен вдоль оси y , и наоборот (т.н. эффект Холла, возникающий в сильном магнитном поле). Если же к образцу будет приложено электрическое поле вдоль оси z , то, как видно из формулы (50), электрический ток останется равным нулю. Таким образом, в направлении z образец будет диэлектриком.

Подчеркнем, что на первый взгляд как тензор проводимости $\langle \sigma \rangle$, так и тензор упругости $\langle k \rangle$ не обязательно должны быть симметричными. Тем не менее, дополнительные соображения, связанные с термодинамической устойчивостью, показывают, что

это не так. Поэтому в дальнейшем мы будем обсуждать главным образом лишь симметричные матрицы.

Наконец, линейные преобразования координат (8) и координат векторов (9) при вращении вокруг оси z на угол φ можно представить в виде

$$\begin{aligned}x' &= T_{xx}x + T_{xy}y + T_{xz}z, \\y' &= T_{yx}x + T_{yy}y + T_{yz}z, \\z' &= T_{zx}x + T_{zy}y + T_{zz}z,\end{aligned}\tag{52}$$

и

$$\begin{aligned}A_{x'} &= T_{xx}A_x + T_{xy}A_y + T_{xz}A_z, \\A_{y'} &= T_{yx}A_x + T_{yy}A_y + T_{yz}A_z, \\A_{z'} &= T_{zx}A_x + T_{zy}A_y + T_{zz}A_z,\end{aligned}\tag{53}$$

где девять коэффициентов преобразования вновь могут быть записаны в виде таблицы

$$\langle T \rangle = \begin{pmatrix} \cos \varphi & \sin \varphi & 0 \\ -\sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}.\tag{54}$$

Как мы видим, недиагональные компоненты матрицы поворота являются антисимметричными.

Как уже подчеркивалось, не любой набор из шести чисел можно рассматривать как компоненты симметричного тензора: компоненты матрицы должны преобразовываться по определенным правилам при переходе от одной системы координат к другой. С другой стороны, как и в случае векторов, при преобразованиях координат должны сохраняться определенные инварианты. Для симметричных матриц такими скалярными инвариантами являются

1. "квадрат" матрицы

$$\langle A \rangle^2 = (A_{xx})^2 + 2(A_{xy})^2 + 2(A_{xz})^2 + (A_{yy})^2 + 2(A_{yz})^2 + (A_{zz})^2,\tag{55}$$

2. сумма диагональных элементов – т.н. "след" (по-немецки "шпур" Spur)

$$\text{Sp} \langle A \rangle = A_{xx} + A_{yy} + A_{zz}.\tag{56}$$

Справедливость соотношений (55) и (56) можно проверить, используя явный вид преобразования компонентов тензора (37). На самом деле общие выражения для квадрата и следа матрицы имеют более сложный вид, содержащий компоненты метрического тензора, см. Приложение. В результате, правильным образом определенная операция следа Sp дает ответ 3 для всех метрических тензоров (41)–(43). Здесь мы привели определения лишь для декартовых систем координат, по разному ориентированных в пространстве.

3.4.2 Четырехмерное пространство

Выше мы везде говорили о трехмерных величинах. Поэтому и таблицы коэффициентов представляли собой матрицы размером 3×3 . Однако уже появившиеся ранее примеры, такие как преобразования Лоренца и четырехмерность энергии-импульса частиц, показывают, что нам неизбежно придется иметь дело с матрицами 4×4 , которые будут иметь три пространственные и одну временную координату (которая обычно стоит на первом месте). В этом случае симметричная матрица будет определяться десятью величинами.

Прежде всего, четырехмерным оказывается метрический тензор $\langle g \rangle$. Действительно, выражение для бесконечно малого инвариантного интервала ds в декартовых и сферических координатах может быть записано в виде

$$ds^2 = g_{tt}dt^2 + g_{xx}dx^2 + g_{yy}dy^2 + g_{zz}dz^2, \quad (57)$$

$$ds^2 = g_{tt}dt^2 + g_{rr}dr^2 + g_{\theta\theta}d\theta^2 + g_{\varphi\varphi}d\varphi^2, \quad (58)$$

где

$$\langle g \rangle_{\text{Minkovsky,D}} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (59)$$

$$\langle g \rangle_{\text{Minkovsky,sph}} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2 \sin^2 \theta \end{pmatrix}, \quad (60)$$

Здесь особое внимание следует обратить на знак минус в левом верхнем углу. Именно появление этого знака и приводит к существенному отличию плоского трехмерного пространства и плоского же, но уже четырехмерного пространства-времени. Поэтому хотя внешне форма (58) мало отличается от соответствующего трехмерного выражения $dr^2 = g_{xx}dx^2 + g_{yy}dy^2 + g_{zz}dz^2$, свойства пространства-времени существенно отличаются от обычного пространства. Как уже говорилось, плоское пространство-время с метрикой (59) называется пространством Минковского. В общем же случае четырехмерная симметричная матрица имеет десять независимых компонент.

Далее, и сами преобразования Лоренца (14), оставляющие постоянной инвариант ds^2 , также очень удобно выразить с помощью матрицы 4×4

$$cdt' = \Lambda_{tt}cdt + \Lambda_{tx}dx, \quad (61)$$

$$dx' = \Lambda_{xt}cdt + \Lambda_{xx}dx, \quad (62)$$

$$dy' = \Lambda_{yy}dy, \quad (63)$$

$$dz' = \Lambda_{zz}dz, \quad (64)$$

где коэффициенты $\Lambda_{...}$ можно записать как

$$\langle \Lambda \rangle = \begin{pmatrix} \frac{1}{\sqrt{1-V^2/c^2}} & -\frac{V/c}{\sqrt{1-V^2/c^2}} & 0 & 0 \\ -\frac{V/c}{\sqrt{1-V^2/c^2}} & \frac{1}{\sqrt{1-V^2/c^2}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (65)$$

Здесь мы для удобства добавили множитель c к времени t , после чего размерность всех координат стала одинаковой.

Наконец, как тоже было понято задолго до Эйнштейна, величины, характеризующие энергию и импульс среды (а не отдельной частицы!) в общем случае также должны быть записаны в виде симметричной четырехмерной матрицы, и, следовательно, содержать десять независимых компонент. Постараемся понять причину такого многообразия. На первый взгляд кажется, что энергия и импульс среды, как и одной частицы, вполне определяются четырьмя величинами, а именно ее энергией и тремя компонентами импульса. Здесь однако речь идет не об полной (интегральной) характеристике отдельной частицы, а о локальных (дифференциальных) свойствах среды, т.е. о плотности энергии-импульса в единице объема. А тогда необходимость введения десяти величин становится вполне оправданной.

Дело в том, что, как уже подчеркивалось, законы физики носят интегральный характер, которые нам однако приходится записывать в локальной (дифференциальной) форме. Поэтому закон сохранения энергии – изменение энергии в некотором объеме есть поток энергии через границу этого объема – в дифференциальном виде требует задания как плотности энергии ε (первый параметр, измеряется в эрг/см³), так и трех компонент вектора потока энергии \mathbf{S} в данной точке (еще три параметра, измеряются в эрг/см³·см/с). Лишь зная эти четыре величины мы сможем подсчитать как изменение энергии в любом заданном объеме за единицу времени, так и поток энергии через поверхность этого объема. А для записи в дифференциальной форме закона сохранения импульса – изменение импульса в некотором объеме есть поток импульса через границу этого объема – нам необходимо знать как три компоненты плотности импульса \mathcal{P} (т.е. импульса среды в единице объема, измеряются в г·см/с·1/см³), так и поток каждой из трех компонент импульса через произвольно ориентированную площадку.

Действительно, как показано на Рис. 4, в общем случае изменение x -компоненты импульса в объеме куба связано не только с течением вещества через площадки, перпендикулярные оси x , но и через все остальные грани (если, конечно, скорость среды не направлена вдоль оси x). Поэтому для описания потока векторной величины через произвольно ориентированную поверхность необходимо задать уже $3 \times 3 = 9$ величин. Соответствующий трехмерный тензор, который мы будем обозначать как $\langle t \rangle$, носит название тензора натяжений, а размерность его компонент есть г·см/с·см/с·1/см³.

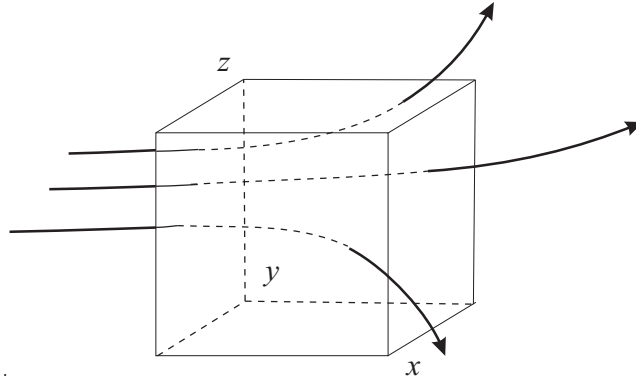


Рис. 4: Изменение x -компоненты импульса в объеме куба связано с течением вещества через все его грани

Таким образом, для записи закона сохранения импульса вообще говоря требуется задать еще двенадцать величин. Однако

1. релятивистская связь $\mathbf{p} = \mathcal{E}\mathbf{v}/c^2$ между энергией $\mathcal{E} = mc^2/\sqrt{1 - v^2/c^2}$ и импульсом $\mathbf{p} = m\mathbf{v}/\sqrt{1 - v^2/c^2}$ приводит к тому, что в релятивистском случае трехмерный вектор потока энергии лишь коэффициентом $1/c$ отличается от трехмерного вектора плотности импульса; если же перейти от переменной t к переменной ct , как это сделано ниже, то эти величины вообще окажутся одинаковыми,
2. как и в случае с тензорами проводимости $\langle \sigma \rangle$ и упругости $\langle k \rangle$, тензор натяжений $\langle t \rangle$ должен быть симметричным, и, следовательно, задаваться лишь шестью величинами.

Подчеркнем, что соотношение $\mathbf{p} = \mathcal{E}\mathbf{v}/c^2$, приводящее к необходимой связи $\mathcal{P} = \varepsilon\mathbf{v}/c^2$ между плотностью импульса \mathcal{P} и плотностью потока энергии $\mathbf{S} = \varepsilon\mathbf{v}$, остается справедливым и для фотонов, движущихся со скоростью света.

В результате, симметричная таблица 4×4 , описывающая плотность энергии-импульса среды (ее называют просто тензором энергии-импульса)

$$\langle T \rangle = \begin{pmatrix} \varepsilon & \mathbf{S} \\ \mathbf{S} & \langle t \rangle \end{pmatrix}, \quad (66)$$

действительно задается десятью независимыми параметрами, а именно скалярной плотностью энергии ε , тремя компонентами потока энергии \mathbf{S} и шестью компонентами симметричного тензора натяжений $\langle t \rangle$, определяющими потоки импульса при движении среды. Столь необычный объект, как мы увидим, также играет ключевую роль при построении релятивистской теории гравитации.

По естественным причинам, тензор энергии-импульса не входит в программу средней школы. Поэтому давайте рассмотрим здесь несколько простых примеров, позволяющих

понять, хотя бы в общих чертах, свойства этой физической величины. Но прежде необходимо объяснить, почему трехмерный тензор $\langle t \rangle$ носит название тензора натяжений. Дело в том, что по самому своему определению компоненты тензора натяжений определяют поток импульса через единичные площадки, по разному ориентированные в пространстве. Например, компонента t_{xx} задает поток x -компоненты импульса через единичную площадку, перпендикулярную оси x , компонента t_{xy} – поток x -компоненты импульса через единичную площадку, перпендикулярную оси y , и т.д. Но поток импульса, как видно из его размерности, есть ни что иное как сила, отнесенная к единице площади. Так и должно быть, поскольку изменение импульса среды в некотором объеме должно быть связана именно с действием силы, приложенной к границе этого объема. В этом смысле ситуация вполне аналогична давлению P , поскольку давление имеет смысл как силы, приложенной к единичной поверхности ($P = \text{сила}/\text{площадь}$), так и плотности энергии ($P = \text{энергия}/\text{объем}$). Единственное отличие – в нерелятивистскую плотность энергии не входит энергия покоя частиц.

В качестве первого примера рассмотрим среду, состоящую из частиц, движущихся со скоростью v вдоль оси x . Тогда, например, плотность их энергии можно записать как $\varepsilon_{\text{part}} = \rho_m c^2 \gamma$, где ρ_m – плотность частиц, а $\gamma = 1/\sqrt{1 - v^2/c^2}$ – их лоренц-фактор. Поток же энергии будет выглядеть как $\mathbf{S} = \rho_m c^2 \gamma (\mathbf{v}/c)$, т.е. произведение плотности энергии на скорость перемещения среды (дополнительный фактор $1/c$ введен для того, чтобы размерность всех величин, входящих в тензор энергии-импульса, были одинаковыми). Как уже отмечалось, такая же величина будет и у плотности импульса. Наконец, поток импульса запишется как произведение плотности импульса \mathbf{S} на скорость v . В результате, тензор энергии-импульса такой среды может быть записан в виде

$$\langle T \rangle_{\text{part}} = \begin{pmatrix} \varepsilon_{\text{part}} & (v/c)\varepsilon_{\text{part}} & 0 & 0 \\ (v/c)\varepsilon_{\text{part}} & (v^2/c^2)\varepsilon_{\text{part}} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (67)$$

Теперь нетрудно сообразить, что в случае произвольной скорости среды имеем

$$\langle T \rangle_{\text{part}} = \begin{pmatrix} \varepsilon_{\text{part}} & (v_x/c)\varepsilon_{\text{part}} & (v_y/c)\varepsilon_{\text{part}} & (v_z/c)\varepsilon_{\text{part}} \\ (v_x/c)\varepsilon_{\text{part}} & (v_x^2/c^2)\varepsilon_{\text{part}} & (v_x v_y/c^2)\varepsilon_{\text{part}} & (v_x v_z/c^2)\varepsilon_{\text{part}} \\ (v_y/c)\varepsilon_{\text{part}} & (v_x v_y/c^2)\varepsilon_{\text{part}} & (v_y^2/c^2)\varepsilon_{\text{part}} & (v_y v_z/c^2)\varepsilon_{\text{part}} \\ (v_z/c)\varepsilon_{\text{part}} & (v_x v_z/c^2)\varepsilon_{\text{part}} & (v_y v_z/c^2)\varepsilon_{\text{part}} & (v_z^2/c^2)\varepsilon_{\text{part}} \end{pmatrix}. \quad (68)$$

Для покоящейся же среды получаем просто

$$\langle T \rangle_{\text{part}} = \begin{pmatrix} \rho_m c^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (69)$$

В качестве другого примера можно рассмотреть плоскую электромагнитную волну, также распространяющуюся вдоль оси x . Для нее тензор энергии-импульса будет иметь вид

$$\langle T \rangle_{\text{em}} = \begin{pmatrix} \varepsilon_{\text{em}} & \varepsilon_{\text{em}} & 0 & 0 \\ \varepsilon_{\text{em}} & \varepsilon_{\text{em}} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (70)$$

где $\varepsilon_{\text{em}} = (E^2 + B^2)/8\pi$ (см. Приложение). Как и следовало ожидать, тензор энергии-импульса плоской волны может быть получен из тензора энергии-импульса частиц заменой плотности энергии частиц $\rho_m c^2 \gamma$ на плотность энергии электромагнитного поля $(E^2 + B^2)/8\pi$ и предельным переходом $v \rightarrow c$. Наконец предполагая, что среда состоит из покоящегося равновесного газа, обладающего внутренней энергией ε и давлением P , получаем

$$\langle T \rangle_{\text{gas}} = \begin{pmatrix} \varepsilon & 0 & 0 & 0 \\ 0 & P & 0 & 0 \\ 0 & 0 & P & 0 \\ 0 & 0 & 0 & P \end{pmatrix}. \quad (71)$$

Этот тензор энергии-импульса используется при построении космологических решений, которые мы кратко рассмотрим в последней части этого раздела. Отметим, что для температур $T \ll 10^{11}$ К, когда тепловые скорости частиц много меньше скорости света, тензор энергии-импульса (71) практически совпадает с тензором (69).

Еще раз подчеркнем, что тензор энергии-импульса позволяет сформулировать как закон сохранения энергии (для этого необходимо воспользоваться величинами из первой строки), так и закон сохранения импульса (три остальные строки). При этом его важнейшим свойством является свойство лоренц-инвариантности. Это значит, что при переходе в другую систему отсчета (т.е. при лоренцевской замене переменных) все компоненты тензора энергии-импульса вновь будут иметь тот же смысл, что и в исходной системе. Иными словами, если в левом верхнем углу в одной системе координат стояла величина, имеющая смысл плотности энергии ε , то и после перехода в другую систему отсчета (при которой, напомним, компоненты тензора должны преобразовываться по вполне определенным правилам!) там окажется величина, совпадающая с плотностью энергии, но уже в другой координатной системе. Более того, если для компонент тензора $\langle T \rangle$ в одной системе выполнен закон сохранения энергии, то он будет выполнен и во всех других системах, связанных с исходной преобразованиями Лоренца. Отметим, наконец, что тензор энергии-импульса обладает простым свойством аддитивности – полный тензор энергии-импульса представляет собой сумму тензоров электромагнитного поля и частиц.

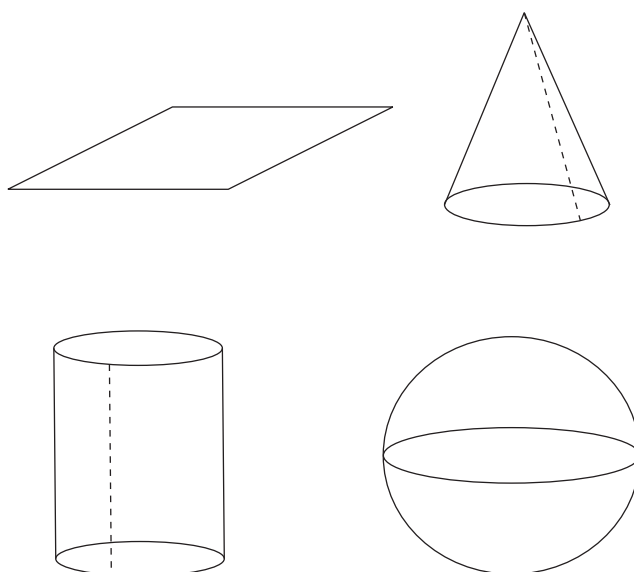


Рис. 5: Плоскость, конус, цилиндр и шар – простейшие примеры двумерных поверхностей. Конус и цилиндр можно развернуть в плоскость вдоль штриховой линии

3.5 Кривизна

3.5.1 Определение кривизны

Вы обратили внимание, что, собираясь заняться теорией гравитации, т.е. теоретической физикой, мы незаметно перешли к обсуждению чисто математических вопросов? И это не случайно. Современные области теории поля требуют использования и самых современных областей математики. Подобная же проблема стояла и перед Эйнштейном. Его гениальная догадка состояла в том, что языком теории гравитации должна быть дифференциальная геометрия – тогда еще достаточно молодая и мало используемая в физике область математики, в разработке которой решающая роль принадлежит К.Ф. Гауссу (1777-1855) и Б. Риману (1826-1866), а также Н.И. Лобачевскому (1792-1856), Э.Б. Кристоффелю (1829-1900) и Г. Риччи-Курбастро (1853-1925). По этой причине нам также придется посвятить еще несколько страниц математической подготовке. Вернемся поэтому к метрическому тензору $\langle g \rangle$ и покажем, как с помощью этой таблицы коэффициентов, определяющей расстояние между двумя близкими точками, можно получить исчерпывающую информацию о геометрических свойствах пространства. Для наглядности в этом разделе мы существенно упростим задачу и будем рассматривать не четырехмерные и даже не трехмерные, а двумерные пространства, т.е. поверхности.

Итак, рассмотрим произвольную поверхность в трехмерном пространстве. Примеры таких поверхностей хорошо известны из курса геометрии средней школы – это плоскость, цилиндр, конус, сфера (см. Рис. 5). Поскольку в этом случае положение любой точки на

поверхности будет определяться заданием двух координат, то, как можно легко догадаться, метрический тензор $\langle g \rangle$ представляет собой симметричную таблицу 2×2 , т.е. задается тремя функциями координат.

Прежде всего, необходимо точно сформулировать, что мы будем понимать под кривизной поверхности. На первый взгляд, свойства цилиндра и конуса ближе к сфере, чем к плоскости, поскольку все они имеют круглые сечения при определенном положении секущей плоскости. Однако в действительности это не так, и свойства конусов и цилиндров во многом гораздо ближе к плоскости, чем к поверхности сферы. Действительно, как конус, так и цилиндр могут быть развернуты в плоскость. Для сферы это сделать нельзя, так что глобус приходится обклеивать многими полосками бумаги. Поэтому локально, в малой окрестности любой точки, за исключением вершины конуса, поверхность конуса и цилиндра будет эквивалентна плоскости, кривизна которой, естественно, равна нулю. Отличие будет проявляться лишь на больших расстояниях, т.к. форма выкройки у конуса и цилиндра будут различны.

Здесь мы неявно использовали одно очень важное предположение: поверхность нельзя, подобно резине, растягивать в разные стороны. Именно это и означает, что на поверхности фиксируется жесткая координатная сетка, однозначно определяющая расстояния между двумя близкими точками. Таким образом, координатная сетка на поверхности цилиндра или конуса в малой окрестности любой точки фактически не будет отличаться от плоской поверхности. В частности, на них можно ввести и декартову систему координат.

Как мы видим, наивное представление о кривизне поверхности не всегда оказывается верным. Кажущаяся кривизна конусов и цилиндров исчезает, если эти поверхности распрямить относительно линии, показанной на Рис. 5. Следовательно, кривизна поверхности в данной точке оказывается нулевой, если в одном (и только в одном) из сечений поверхности плоскостью получается линия, радиус кривизны в этой точке бесконечен². Понятно, что и конус, и цилиндр имеют такие сечения, тогда как при любом сечении сферы получается окружность.

Таким образом, кривизна k должна определяться не одним, а двумя радиусами кривизны R_1 и R_2 , соответствующими двум, наименьшим и наибольшим (и, вообще говоря, разным) радиусам окружностей, возникающим при сечении искомой поверхности

$$k = \frac{1}{R_1 R_2}. \quad (72)$$

Так, для сферы, у которой центры любых двух окружностей находятся по одну сторону от элемента поверхности, кривизна считается положительной. Поэтому сфера считается поверхностью постоянной положительной кривизны. Поверхность же седла, также показанная на Рис. 6, для которой центры окружности находятся по разную сторону от

²Исключение составляет плоскость, для которой любое сечение дает прямую линию.

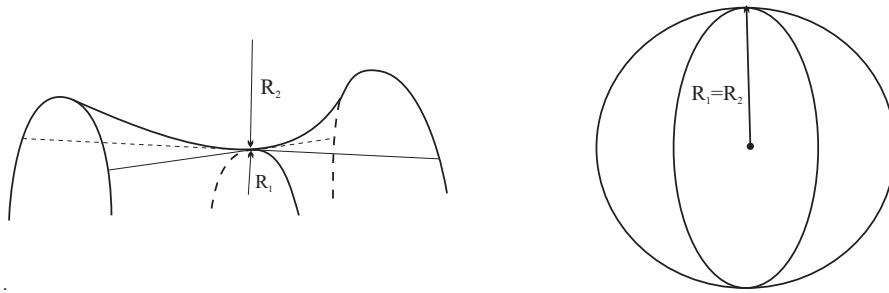


Рис. 6: Поверхности отрицательной (седло) и положительной (сфера) кривизны. Для сферы не существует сечений, отличных от окружности. Для поверхности же седла через каждую точку можно провести две плоскости, пересекающие поверхность по линиям с нулевой кривизной в этой точке

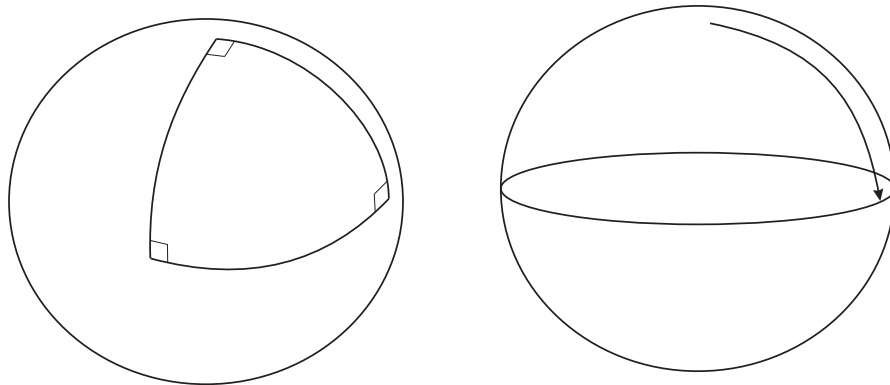


Рис. 7: Треугольник и окружность на поверхности сферы, для которых сумма углов есть $3\pi/2$, а отношение длины окружности к "радиусу" (расстоянию до полюса вдоль поверхности сферы) равно 4

поверхности, кривизна будет отрицательной (в этом случае, как хорошо известно, в каждой точке существует два сечения, дающие линии с нулевой кривизной в этой точке). Если же один из радиусов стремиться к бесконечности, как это имеет место в случае цилиндра или конуса, то кривизна поверхности будет равна нулю.

Отметим теперь следующее важное обстоятельство. При наивном анализе кривизны поверхности мы, фактически, опирались на поведение двумерных поверхностей в трехмерном пространстве. В частности, при измерении радиуса кривизны различных сечений нам приходилось выйти за пределы самой поверхности. Соответственно, распрямление конусов и цилиндров также предполагало их перемещение в трехмерном пространстве. Предположим, однако, что у нас нет возможности покидать саму поверхность. Существует ли способ, позволяющий сформулировать понятие кривизны, оставаясь в пределах рассматриваемой поверхности? Иными словами, могут ли двумерные существа, живущие на двумерной поверхности, определить кривизну своего мира?

Оказывается, такой способ действительно существует. Для этого вспомним две важных теоремы, известные нам из планиметрии.

- Сумма углов в треугольнике Σ равна π и не зависит от размера треугольника.
- Отношение длины окружности к радиусу L_0/R равно 2π и не зависит от размера окружности.

Однако эти утверждения справедливы лишь для поверхностей с нулевой кривизной. Действительно, сумма углов треугольника на поверхности сферы, показанного на Рис. 7, составляет $3\pi/2$, а отношение длины окружности к радиусу (расстоянию до любого полюса вдоль поверхности сферы) для экватора, показанного на Рис. 7, равно четырем. Вместе с тем, для маленьких треугольников и окружностей, у которых их размер много меньше радиуса сферы, классические соотношения будут выполнены, причем с тем большей точностью, чем меньше размеры самих фигур. Следовательно, мы приходим к важному выводу, что мерой кривизны поверхности может служить характер изменения суммы углов треугольника Σ и отношения длины окружности к радиусу L_0/R при увеличении их линейного размера.

3.5.2 Геодезические

Здесь необходимо сделать еще одно (на этот раз последнее, но очень важное) математическое отступление. Как при определении суммы углов треугольника, так и при определении отношения L_0/R , нам необходимо обобщить понятие отрезка на случай поверхностей с ненулевой кривизной. Только в этом случае мы сможем построить треугольник по трем вершинам (и, значит, определить сумму его углов) или же начертить окружность – кривую, все точки которой одинаково отстоят от ее центра.

Последнее обстоятельство наводит нас на мысль, что возможность такого построения становится возможным в том и только в том случае, если мы умеем определять расстояния между двумя произвольными точками. Тогда прямой будем соответствовать такая кривая, длина которой при заданном положении ее концов будет наименьшей. В дифференциальной геометрии такие кривые называются геодезическими и играют столь же фундаментальную роль, как обычные прямые в геометрии на плоскости. Само их название ясно показывает, что впервые они возникли как раз при определении расстояний на поверхности Земли, т.е. на поверхности, свойства которой отличаются от свойств плоскости.

В качестве упражнения, наглядно демонстрирующего понятие геодезической, можно рассмотреть следующую задачу (см. Рис. 8). Альпинист пытается подняться на вершину горы, имеющей идеальную коническую форму. Он закинул веревку за вершину и, постепенно уменьшая ее длину, поднимается вверх. Спрашивается, при каких значениях

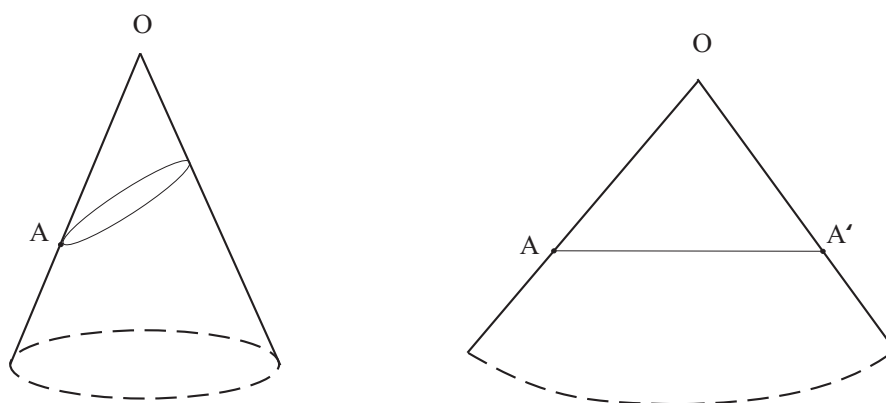


Рис. 8: Альпинист на бесконечно гладкой конической поверхности. Геодезическая линия соответствует минимальной (для каждого положения альпиниста) длине веревки. Поэтому если в развертке конуса угол при вершине оказывается больше 180° , веревка будет соскакивать с вершины

угла α при вершине конуса веревка не будет соскакивать с вершины? Действительно, для малых углов α альпинисту, безусловно, удастся зацепиться за вершину, тогда как при конусе, близком к плоскости, подъем указанным способом окажется невозможным.

Ответ легко может быть получен, если мы заметим, что натянутая веревка описывает кратчайшее расстояние для данного положения альпиниста A на поверхности конуса. Как мы уже поняли, поверхность конуса совершенно аналогична плоскости, и поэтому кратчайшим расстоянием между двумя точками будет прямая. Развернув поверхность конуса по лучу, проходящему через точку A и вершину конуса, мы тем самым расположим альпиниста не только в точке A , но и в точке A' . Геодезической же будет просто прямая, соединяющая эти две точки. Теперь становится понятным, что если угол AOA' в развертке конуса меньше двух прямых, то геодезическая прямая пройдет в пределах самой поверхности, тогда как при угле $AOA' > \pi$ наименьшая прямая лежит "вне поверхности", что и означает, что веревка не сможет зацепиться за вершину. В результате, как легко проверить, критический угол при вершине будет равен 30° .

3.5.3 Треугольники и окружности

Только теперь, определив понятие геодезической, мы можем вернуться к вопросу об определении собственной кривизны поверхности. Рассмотрим сначала вопрос о сумме углов в треугольнике. Предположим, что нам известны координаты трех его вершин A , B и C . Тогда необходимо выполнить следующую процедуру.

1. Для каждой пары точек найти геодезическую, соединяющую эти точки.
2. Найти малые векторы \mathbf{a}_1 и \mathbf{a}_2 , \mathbf{b}_1 и \mathbf{b}_2 , \mathbf{c}_1 и \mathbf{c}_2 , выходящие соответственно из точек

A , B и C и параллельные геодезическим, соединяющим вершины треугольников.

3. Для каждой пары лучей определить скалярное произведение $(\mathbf{a}_1 \cdot \mathbf{a}_2)$, *etc.* и длины векторов $|\mathbf{a}_1| = \sqrt{(\mathbf{a}_1 \cdot \mathbf{a}_1)}$, *etc.*
4. С помощью формулы $\cos \alpha = (\mathbf{a}_1 \cdot \mathbf{a}_2) / (|\mathbf{a}_1| |\mathbf{a}_2|)$, *etc.* найти углы.

Подчеркнем, что в трех первых пунктах вычисления могут быть выполнены лишь в том случае, если известен метрический тензор $\langle g \rangle$. Действительно, рассмотрим пункт 1 более подробно. Для определения геодезической необходимо провести измерение длин различных кривых. А это можно сделать лишь с помощью соотношения

$$\delta l^2 = g_{11}(x_1, x_2) \delta x_1^2 + 2g_{12}(x_1, x_2) \delta x_1 \delta x_2 + g_{22}(x_1, x_2) \delta x_2^2, \quad (73)$$

суммируя величины δl при заданном изменении координат x_1 и x_2 вдоль данной кривой. При этом чрезвычайно важно, что как длина всей кривой $l = \sum \delta l$ (которая полностью соответствует показателю пройденного пути на спидометре автомобиля), так и форма самой геодезической не будут зависеть от выбора координатной сетки. Соответственно, скалярное произведение двух векторов \mathbf{a}_1 и \mathbf{a}_2 , выходящих из начала произвольной координатной сетки (p, q) и имеющих координаты (p_1, q_1) и (p_2, q_2) следует определить как

$$(\mathbf{a}_1 \cdot \mathbf{a}_2) = g_{11}p_1p_2 + g_{12}(p_1q_2 + p_2q_1) + g_{22}q_1q_2. \quad (74)$$

Наконец, квадрат длины вектора будет просто

$$(\mathbf{a}_1 \cdot \mathbf{a}_1) = g_{11}p_1^2 + 2g_{12}p_1q_1 + g_{22}q_1^2, \quad (75)$$

что полностью соответствует соотношению (46).

Как мы видим, процедура определения углов оказалась довольно утомительной. Поэтому в дальнейшем мы будем рассматривать лишь второй пример, а именно отношение длины окружности к радиусу. В этом случае удобно воспользоваться полярными координатами r, φ . Для плоской метрики соответствующий метрический тензор будет выглядеть как

$$\langle g \rangle_{\text{polar}} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}, \quad (76)$$

а выражение для элемента длины – как

$$dl^2 = dr^2 + r^2 d\varphi^2. \quad (77)$$

Как мы видим, и здесь нам потребовалось знание метрического тензора $\langle g \rangle$.

Теперь, зная выражение для элемента длины (77), ничего не стоит определить отношение длины окружности к радиусу. Прежде всего, определим расстояние R между

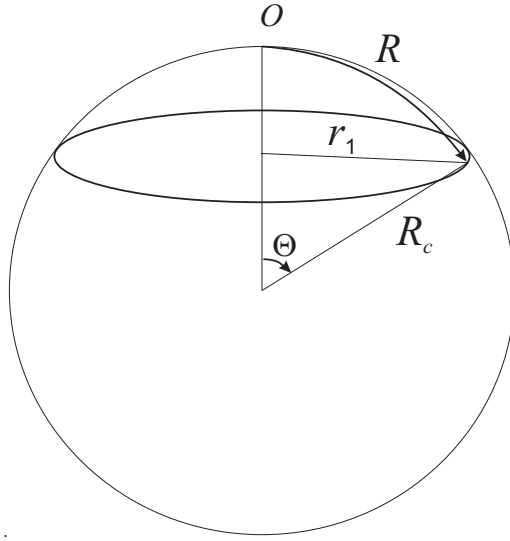


Рис. 9: Метрика на поверхности сферы радиуса R_c

центром окружности O , который мы поместим в начало координат ($r = 0$), и точкой с координатами (r, φ) . Не тратя время на доказательство очевидного факта (хотя на самом деле это необходимо сделать!), что геодезической кривой должна быть линия постоянного угла $\varphi = \text{const}$ (и поэтому вдоль радиуса $\delta\varphi = 0$), получаем с помощью (77), что $\delta l = \delta r$. Складывая теперь приращения длин, имеем, что, как и следовало ожидать, расстояние l между центром окружности и точкой с координатой (r, φ) совпадает с координатой r . Поэтому радиус окружности с центром в начале координат, проходящей через точку с координатой (r, φ) , есть $R = r$.

Рассмотрим теперь приращение длины при движении по окружности, когда по определению $\delta r = 0$. Воспользовавшись вновь соотношением (77), получаем очевидное соотношение

$$\delta l = r \delta \varphi. \quad (78)$$

Складывая теперь элементы длины при изменении угла φ от 0 до 2π , имеем для длины окружности $L_0 = \sum \delta l$

$$L_0 = 2\pi R. \quad (79)$$

Как и следовало ожидать, отношение $L_0/R = 2\pi$ не зависит от радиуса окружности.

Попробуем провести подобную же процедуру на поверхности сферы радиуса R_c . В этом случае, как показано на Рис. 9, "радиусом" R окружности будет длина дуги большого круга

$$R = \Theta R_c, \quad (80)$$

которая не совпадает с "истинным" радиусом

$$r_1 = R_c \sin \Theta. \quad (81)$$

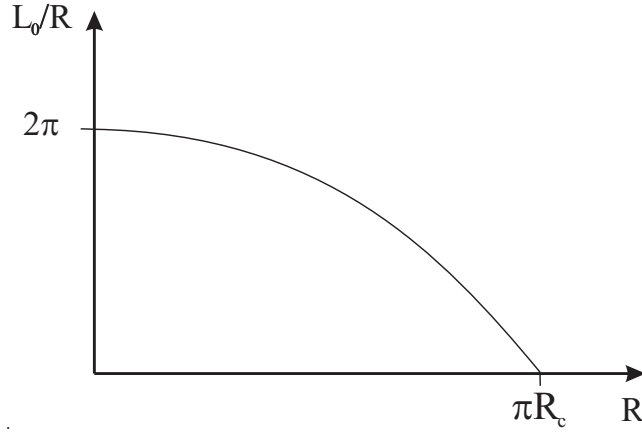


Рис. 10: Зависимость отношения длины окружности L_0 к радиусу R при изменении размера фигуры

Именно такое расстояние между центром окружности и каждой ее точкой отмерят двумерные существа, способные перемещаться лишь по поверхности сферы. В результате, воспользовавшись соотношениями (80) и (81), получаем для малого элемента длины $dl^2 = dr^2 + r_1^2 d\varphi^2$

$$dl^2 = dr^2 + R_c^2 \sin^2(r/R_c) d\varphi^2. \quad (82)$$

Следовательно, метрический тензор имеет вид

$$\langle g \rangle_{\text{sph}} = \begin{pmatrix} 1 & 0 \\ 0 & R_c^2 \sin^2(r/R_c) \end{pmatrix}. \quad (83)$$

При этом необходимо сделать одно важное уточнение: в случае плоскости (метрический тензор (76)) радиальная координата r может принимать любое значение от нуля до бесконечности, тогда как в случае поверхности сферы (метрический тензор (83)) радиальная координата ограничена сверху: $0 < R < \pi R_c$.

Найденный нами метрический тензор содержит всю информацию о кривизне поверхности сферы. В частности, он показывает, при каких размерах окружности мы начнем чувствовать отличие от плоской поверхности. Так, ясно, что при малых радиусах $R \ll R_c$ тензор (83) совпадает с метрикой плоского пространства (76). Чтобы показать это, достаточно воспользоваться известным разложением $\sin x \approx x$. С другой стороны, при увеличении "радиуса" окружности r отношение длины окружности к радиусу будет теперь определяться формулой (проверьте!)

$$\frac{L_0}{R} = 2\pi \frac{R_c}{R} \sin\left(\frac{R}{R_c}\right). \quad (84)$$

Как показано на Рис. 10, это отношение уменьшается при увеличении радиуса R , так что при $R = \pi R_c$ длина окружности обращается в ноль. Как и следовало ожидать, отличие метрики от плоской проявляется при $R \approx R_c$.

Кстати, обратите внимание на то, что структура тензора (83) близка к структуре тензора (76), описывающего плоское двумерное пространство. Они отличаются лишь выражениями в правом нижнем углу. Оказывается, это не случайно. Дело в том, что как уже неоднократно подчеркивалось (и как будет продемонстрировано в следующем разделе), инвариантные величины не должны зависеть от выбора системы отсчета. Но при наличии такой свободы мы всегда можем выбрать такую координатную сетку (а, значит, и такие переменные x_1 и x_2), что две из трех величин, входящих в двумерный метрический тензор, могут быть выбраны из соображения удобства. Рассмотренные выше случаи как раз и соответствуют удобному выбору координат, когда $g_{11} = 1$ и $g_{12} = 0$. Соответственно, в четырехмерном пространстве всегда можно провести такую замену четырех координат, что метрический тензор $\langle g \rangle$ будет иметь не десять, а шесть независимых компонент. Физические же результаты от этого зависеть не будут. Запомним и это замечательное свойство – оно также понадобится нам в дальнейшем.

3.5.4 Гауссова кривизна

Теперь нам осталось понять, как, зная явный вид метрического тензора $\langle g \rangle$, определить кривизну $k = 1/R_1 R_2$ (72). Для двумерных поверхностей ответ был получен еще в начале XIX века К.Ф. Гауссом. Соответствующая формула для т.н. гауссовой кривизны имеет вид ($g = g_{11}g_{22} - g_{12}^2$)

$$\begin{aligned}
 k = \frac{1}{2g} \left(2 \frac{\partial^2 g_{12}}{\partial x_1 \partial x_2} - \frac{\partial^2 g_{11}}{\partial x_2^2} - \frac{\partial^2 g_{22}}{\partial x_1^2} \right) + \frac{1}{4g^2} \left[g_{22} \left(\frac{\partial g_{11}}{\partial x_2} \right)^2 - 2g_{12} \frac{\partial g_{22}}{\partial x_1} \frac{\partial g_{11}}{\partial x_2} + g_{11} \left(\frac{\partial g_{22}}{\partial x_1} \right)^2 \right] \\
 - \frac{1}{4g^2} \left[g_{22} \frac{\partial g_{11}}{\partial x_1} \left(2 \frac{\partial g_{12}}{\partial x_2} - \frac{\partial g_{22}}{\partial x_1} \right) + g_{11} \frac{\partial g_{22}}{\partial x_2} \left(2 \frac{\partial g_{12}}{\partial x_1} - \frac{\partial g_{11}}{\partial x_2} \right) \right] \\
 + \frac{g_{12}}{4g^2} \left[\frac{\partial g_{22}}{\partial x_2} \frac{\partial g_{11}}{\partial x_1} + \left(2 \frac{\partial g_{12}}{\partial x_2} - \frac{\partial g_{22}}{\partial x_1} \right) \left(2 \frac{\partial g_{12}}{\partial x_1} - \frac{\partial g_{11}}{\partial x_2} \right) \right]. \quad (85)
 \end{aligned}$$

Несмотря на свою очевидную громоздкость, выражение (85) для гауссовой кривизны является одной из самых замечательных формул в дифференциальной геометрии. Дело в том, что результат – гауссова кривизна k – совершенно не зависит от выбора системы координат. Можно измерять координату x_1 в километрах а координату x_2 – в миллиметрах, можно взять ортогональную или любую другую координатную сетку – все равно ответ останется одинаковым. В частности, для любой координатной сетки на плоскости мы получим $k = 0$ а в случае сферы $k = R_c^{-2}$. Для найденных выше метрических тензоров этот факт проверяется прямым вычислением. Так, для декартовых координат $x_1 = x$, $x_2 = y$ все метрические коэффициенты постоянны (не зависят от x и y), так что результат $k = 0$ не вызывает сомнения. Для полярных координат $x_1 = r$, $x_2 = \varphi$ ($g_{rr} = 1$,

$g_{\varphi\varphi} = r^2$, $g_{r\varphi} = g_{\varphi r} = 0$) имеем $g = r^2$ и

$$k = -\frac{1}{2g} \frac{\partial^2 g_{\varphi\varphi}}{\partial r^2} + \frac{g_{rr}}{4g^2} \left(\frac{\partial g_{\varphi\varphi}}{\partial r} \right)^2. \quad (86)$$

Все остальные слагаемые в громоздком выражении (85) тождественно равны нулю. Воспользовавшись теперь явным видом величин g и $g_{\varphi\varphi}$, получаем вновь $k = 0$. С другой стороны, для метрического тензора (72), описывающего метрику поверхности сферы, получаем,

$$k = \frac{1}{R_c^2}. \quad (87)$$

Здесь мы воспользовались выражением $g_{\varphi\varphi} = g = R_c^2 \sin^2(r/R_c)$, а также тем фактом, что соотношение (86) остается справедливым и для метрики (72).

Но, как уже говорилось, физический результат (кривизна поверхности) не должен зависеть от выбора координатной сетки. Например, можно в качестве радиальной координаты выбрать величину r_1 . Тогда угловая часть приращения длины dl^2 , как и для плоскости, будет иметь вид $dl^2 = r_1^2 d\varphi^2$. Приращение же длины за счет изменения координаты r_1 запишется теперь как $dl^2 = dr_1^2 / \cos^2 \Theta$. В результате, имеем

$$dl^2 = \frac{dr_1^2}{1 - r_1^2/R_c^2} + r_1^2 d\varphi^2. \quad (88)$$

Поэтому в этом случае

$$\langle g \rangle_{\text{sph}} = \begin{pmatrix} (1 - r_1^2/R_c^2)^{-1} & 0 \\ 0 & r_1^2 \end{pmatrix}. \quad (89)$$

С другой стороны, выражение для гауссовой кривизны теперь должно быть записано как

$$k = -\frac{1}{2g} \frac{\partial^2 g_{\varphi\varphi}}{\partial r_1^2} + \frac{g_{rr}}{4g^2} \left(\frac{\partial g_{\varphi\varphi}}{\partial r_1} \right)^2 + \frac{g_{\varphi\varphi}}{4g^2} \frac{\partial g_{rr}}{\partial r_1} \frac{\partial g_{\varphi\varphi}}{\partial r_1}, \quad (90)$$

что вновь дает $k = R_c^{-2}$. Умеющие дифференцировать могут поупражняться, выбирая другие (не обязательно прямоугольные) сетки координат³.

К сожалению, как мы видим, для точного математического определения кривизны знание дифференциального исчисления является необходимым. Тем не менее, общий смысл рассмотренной выше процедуры не должен вызывать затруднения. Как уже говорилось, она сводится, фактически, к определению характера изменения суммы углов треугольника (или же отношения длины окружности к радиусу) при увеличении размера геометрической фигуры. Тем самым, кривизна поверхности может быть определена

³Удобство выбора координат r_1 , φ , и значит, метрического коэффициента $g_{\varphi\varphi} = r_1^2$ состоит в том, что в этом случае сумма двух первых слагаемых в (90) тождественно равна нулю.

локально, и не выходя за пределы самой поверхности. При этом для поверхностей положительной кривизны (сфера) сумма углов треугольника увеличивается, а отношение длины окружности уменьшается с увеличением размера. И наоборот, для поверхностей отрицательной кривизны (седло) сумма углов треугольника уменьшается, а отношение длины окружности увеличивается при росте размера фигуры. В частности, увеличение отношения L_0/R для поверхности седла связано просто с тем, что при увеличении "радиуса" R при движении по "окружности" $R = \text{const}$ нам приходится спускаться во все более глубокие долины и подниматься на все более высокие взгорья по мере того, как радиус будет увеличиваться. Эти свойства хорошо видны на Рис. 6.

Далее, как показано на Рис. 10, отношение длины окружности к радиусу L_0/R остается практически постоянным при малых размерах окружности по сравнению с радиусом сферы. В результате, кривизна пропорциональна "ускорению" изменения отношения L_0/R при изменении R , т.е. второй производной $d^2(L_0/R)/dR^2$. Именно поэтому в формулах (85) и (86) входит вторая производная метрического тензора по координатам. При этом отличие отношения L_0/R от 2π , как и отличие суммы углов треугольника Σ от π , будет пропорционально площади фигуры δS . Так, например, можно показать, что

$$\Sigma - \pi = k\delta S, \quad (91)$$

где k – гауссова кривизна, а $\delta S \ll R_c^2$. Соответственно, размерность коэффициента пропорциональности k есть $^{-2}$.

Еще одно нетривиальное свойство состоит в том, что кривизна двумерной поверхности полностью определяется лишь одной скалярной величиной – гауссовой кривизной k . Для пространств более высоких размерностей это уже не так, и поэтому необходимо вводить несколько величин (3 для трехмерного и 14 для четырехмерного), характеризующих кривизну многомерного пространства. Наконец, необходимо еще раз подчеркнуть, что формула Гаусса (85) является инвариантной, т.е. она позволяет определить инвариантную геометрическую величину – гауссову кривизну – при любом выборе системы координат. Как уже говорилось, такие универсальные соотношения, не зависящие от выбора системы отсчета и позволяющие вычислять ключевые величины, как раз и необходимы в качестве языка для формулировки инвариантных уравнений. Именно поэтому язык дифференциальной геометрии и был использован Эйнштейном для математической записи своей теории гравитационного поля.

3.6 Прямое усмотрение истины

3.6.1 Наводящие соображения

Теперь, познакомившись кратко с новым математическим аппаратом, можно вернуться к нашей основной задаче. Мы остановились на том, что теория всемирного тяготения

Ньютона является правильной и последовательной, но не универсальной теорией. Она правильно описывает медленное движение в слабых гравитационных полях, но не может объяснить ряд наблюдательных фактов (дополнительное движение перигелия Меркурия, не связанное с действием других планет, или же количественное значение угла отклонения света в гравитационном поле Солнца), когда скорости тел и гравитационные поля становятся достаточно велики. Кроме того, эта теория не удовлетворяет некоторым общим свойствам, например, условию инвариантности записи дифференциальных уравнений, описывающих движение тел, которым должна была бы удовлетворять последовательная теория.

Кстати, интуитивно были ясны и характерные величины скорости и гравитационного потенциала, при которых начинает чувствоваться ограниченность теории. Вспомните появление размерных параметров при анализе формулы, определяющей кинетическую энергию. Малость скорости означает ее малость по отношению к скорости света $c \approx 300000$ км/с, поскольку специальная теория относительности должна быть частью общей теории относительности в слабых гравитационных полях. Малость же гравитационного потенциала φ_g (его размерность, как мы помним, есть квадрат скорости) означает его малость по отношению к квадрату скорости света c^2 . Поэтому характерной малой величиной, эквивалентной отношению скорости тел к скорости света в специальной теории относительности, в теории гравитации должна быть величина $\varepsilon_g = \varphi_g/c^2$. Для гравитационного поля тела с массой M и радиусом R эта величина имеет вид

$$\varepsilon_g = \frac{GM}{c^2 R}. \quad (92)$$

При $\varepsilon_g \ll 1$ общая теория относительности должна переходить в обычную теорию Всемирного тяготения Ньютона.

Любопытно, что выражение для гравитационного радиуса тела массы M , т.е. такого размера тела, при котором его поверхность не может покинуть даже луч света, было получено П.С. Лапласом (1749-1827) еще в 1795 году на основе ньютоновской теории тяготения. Действительно, если приравнять вторую космическую скорость $v_2 = \sqrt{2GM/R}$ к скорости света c , то мы получаем

$$r_g = \frac{2GM}{c^2}. \quad (93)$$

Удивительно, что выражение (93) остается справедливым и в общей теории относительности. Запомним эту формулу, она еще не раз возникнет в дальнейшем!

Таким образом, уже на этом этапе можно сделать вывод, что эффекты общей теории относительности должны проявляться лишь для достаточно сильных гравитационных полей. При этом необходимы не столько большие массы гравитирующих тел, сколько их чрезвычайно малые размеры. Иными словами, сильное отличие предсказаний ОТО от предсказаний ньютоновской теории гравитации будет иметь место лишь

при размерах тела, близких к гравитационному радиусу. При этом повседневный опыт здесь оказывается совершенно бесполезным, поскольку как на Земле, так и в солнечной системе параметр ε_g чрезвычайно мал. Действительно, для звезды с массой Солнца $M_\odot \approx 2 \sim 10^{33}$ г гравитационный радиус составляет лишь 3 км, так что для поверхности Солнца $\varepsilon_\odot \sim r_g/R_\odot \approx 10^{-5}$. Соответственно, для Земли (масса $M \approx 6 \sim 10^{27}$ г) гравитационный радиус составляет лишь 0.6 см, так что $\varepsilon_g \sim 10^{-8}$. Неудивительно поэтому, что в большинстве случаев нам оказываются доступны лишь первые поправки в соответствующих разложениях по величине ε_g .

Отметим, наконец, еще одно немаловажное обстоятельство. Как видно из выражения (93), общая теория относительности помимо гравитационной постоянной G , известной из нерелятивистской теории, и скорости света c не требует никаких дополнительных размерных констант. В результате, можно предположить, что гравитационный потенциал при малых значениях параметра ε_g должен иметь вид

$$\Phi_g = c^2 \left[1 + \frac{\varphi_g}{c^2} + A_1 \left(\frac{\varphi_g}{c^2} \right)^2 + \dots \right], \quad (94)$$

где A_1 – безразмерный коэффициент порядка единицы. При этом появление нулевого члена, очень похожего, кстати, на постоянный дополнительный член mc^2 в выражении для кинетической энергии, не может привести к наблюдательным эффектам, т.к. в ньютоновской теории гравитационный потенциал определяется с точностью до произвольной константы. Конечно же, как и в случае с выражением для кинетической энергии, определить коэффициент A_1 можно лишь зная истинный вид уравнений гравитационного поля.

На этот вопрос можно посмотреть и с другой стороны. Всемирная теория тяготения требует задания одной скалярной величины – гравитационного потенциала φ_g . Мы уже видели, что расширение теории может быть связано с переходом от скалярных величин в векторным, и от векторных – к тензорным. Простейшим обобщением ньютоновской гравитации мог бы быть, например, переход к векторной теории, когда гравитационный потенциал является на самом деле временной компонентой четырехмерного потенциала подобно тому, как энергия частицы является временной компонентой четырехмерного вектора энергии-импульса. Кстати, именно такую структуру имеет теория электромагнитного поля Максвелла. Действительно, при медленном движении частиц (по сравнению со скоростью света) электрическое поле полностью определяется скалярным электрическим потенциалом Φ_e , причем сила взаимодействия между зарядами согласно закону Кулона в точности эквивалентна закону Всемирного тяготения.

Оказалось, однако, что такой подход не может привести к успеху. Действительно, в электромагнетизме одноименные заряды отталкиваются, тогда как в гравитации, наоборот, притягиваются. Тем не менее, как мы увидим, некоторые следствия общей теории

относительности оказываются удивительно схожи с хорошо известными свойствами электромагнитного поля. Таким образом, одним из первых вопросов, возникших при построении общей теории относительности, был вопрос о структуре теории; приведенные выше аргументы показывают, что последовательная теория гравитации должна быть тензорной, т.е. описываться симметричной матрицей 4×4 .

Но если это действительно так, то здесь можно сделать еще один шаг, проясняющий контуры будущей теории. Дело в том, что, как видно из самого определения интервала ds^2 (58), матричный элемент g_{tt} при отсутствии смещения в пространстве (т.е. когда $dr^2 = 0$) связывает интервал ds^2 и ход времени dt удаленного наблюдателя:

$$ds^2 = -g_{00}c^2dt^2. \quad (95)$$

Здесь для удобства мы ввели величину $g_{00} = -g_{tt} > 0$. Но в этом случае величина $\sqrt{-ds^2}$ имеет смысл собственного времени τ , измеряемого наблюдателем, покоящимся в данной точке. Следовательно, мы можем записать

$$d\tau = \sqrt{g_{00}}dt. \quad (96)$$

Поэтому если $g_{00} < 1$, то время в этой точке будет течь медленнее по сравнению с бесконечно удаленной точкой, где $g_{00} = 1$.

С другой стороны, изменение собственного времени можно напрямую связать с величиной гравитационного потенциала. Действительно, замедление времени должно немедленно привести к уменьшению частот колебаний всех процессов, происходящих в гравитационном поле, например частот колебаний пружинных маятников в обычных часах, частот колебаний электронов в атомах, частот квантов света, распространяющихся вблизи массивных тел. Ясно, однако, что сам наблюдатель, находящийся в гравитационном поле, по своим часам не должен обнаружить никакого изменения. Для него частоты колебаний пружин и частоты атомных переходов будут такими же, как и в отсутствие гравитационного поля. В этом, кстати, и проявляется принцип эквивалентности, гарантирующий тождественность результатов экспериментов во всех системах отсчета. Отличие будет иметь место лишь с точки зрения другого наблюдателя, находящегося вне гравитационного поля.

Но мы знаем, что частота ν любого колебания связана с энергией соотношением $\mathcal{E} = h\nu$, где h – постоянная Планка. А это означает, что и энергия фотона \mathcal{E}_∞ , вышедшего из области с ненулевым гравитационным полем, т.е. измеренная удаленным наблюдателем, окажется меньше его энергии \mathcal{E} при излучении, причем теперь

$$\mathcal{E}_\infty = \sqrt{g_{00}}\mathcal{E}. \quad (97)$$

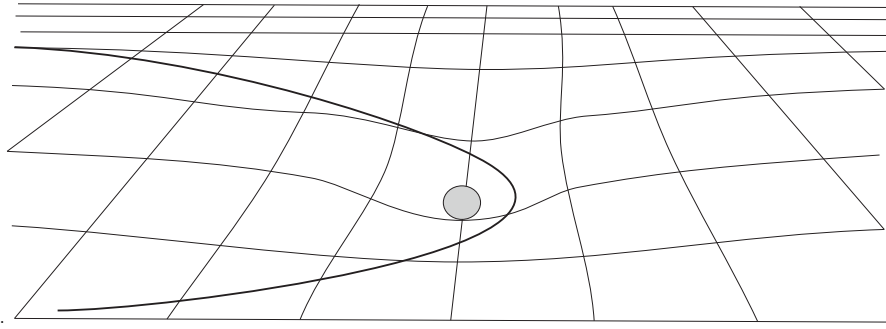


Рис. 11: В присутствии гравитационных полей траектория свободно движущихся тел искривляется, что можно связать с искривлением самого пространства

Естественно связать такое изменение энергии с появлением гравитационного поля, в присутствии которого закон сохранения энергии имеет вид

$$\mathcal{E}_\infty = \mathcal{E} + m\varphi_g, \quad (98)$$

где здесь $m = \mathcal{E}/c^2$ – ”масса” фотона. Хорошо известный механический аналог – кинетическая энергия тела уменьшается при удалении от гравитирующих тел. Поэтому более точно следовало бы сказать, что изменяется ”кинетическая” энергия фотона $h\nu$, а его полная энергия \mathcal{E}_∞ остается постоянной. В результате, сравнивая соотношения (97) и (98), мы приходим к заключению, что матричный элемент g_{00} в слабом гравитационном поле должен выглядеть как

$$g_{00} \approx 1 + \frac{2\varphi_g}{c^2}. \quad (99)$$

Это соотношение, как мы видим, действительно близко по форме к выражению (94), полученному ранее из более общих соображений.

Следующей основополагающей догадкой была гипотеза о связи гравитационного поля и кривизны самого пространства. Основание этого положения в общем не требует комментариев, поскольку гравитационное поле приводит к искривлению траекторий. Хорошей наглядной иллюстрацией здесь может служить двумерная аналогия, показанная на Рис. 11. Как мы видим, наличие тяготеющего центра, искривляющего двумерную мембрану, приводит к отклонению движения тела от прямолинейного.

С другой стороны, и в этом вопросе было далеко не все понятно. Действительно, искривленность траекторий различных тел у поверхности Земли, где, очевидно, само гравитационное поле можно считать постоянным, заведомо зависит от начальной скорости этих тел (см. Рис. 12). Поэтому не представляется возможным напрямую связать кривизну траекторий с напряженностью гравитационного поля (подобную теорию пытался развить Риман). В чем же причина такого отличия в кривизне траекторий для одного и того же гравитационного поля? И можно ли так определить кривизну траектории, чтобы она была одинаковой для всех возможных случаев и определялась лишь

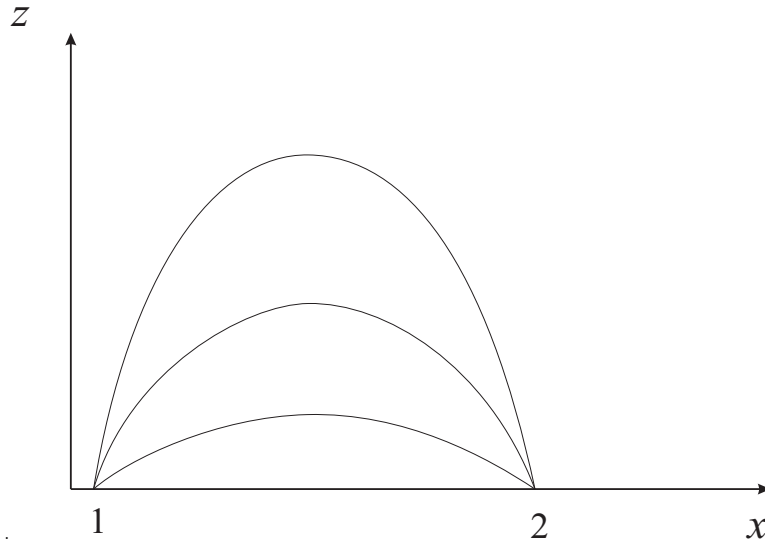


Рис. 12: Траектории движения тел, брошенных из точки 1 в точку 2 под разными углами к горизонту. Кривизна траекторий в трехмерном пространстве оказывается различной

напряженностью гравитационного поля \mathbf{g} ? Оказывается, можно. И здесь нам на помощь действительно приходит четвертое измерение.

Обсудим этот вопрос подробнее – он очень поучителен⁴. Рассмотрим движение тела в однородном гравитационном поле \mathbf{g} , брошенного из точки x_1 в точку x_2 . При этом мы будем исследовать движение тела не между двумя пространственными точками \mathbf{r}_1 и \mathbf{r}_2 , а между двумя пространственно-временными точками (\mathbf{r}_1, t_1) и $(\mathbf{r}_2, t_1 + \delta t)$. При этом для удобства мы будем рассматривать ось времени как пространственную координату. Это можно легко сделать, перейдя от переменной t к переменной ct .

Обратите внимание – в самой постановке задачи содержится новый элемент. Мы не хотим определить траекторию тела, брошенного из точки \mathbf{r}_1 с определенной скоростью \mathbf{v} , а обсуждаем траекторию тела, которое, начав движение в точке \mathbf{r}_1 , через некоторое время δt будет находиться в точке с координатой \mathbf{r}_2 . В этом случае начальная скорость тела \mathbf{v}_0 (вертикальная начальная скорость V_0) уже не является свободным параметром, а должна зависеть от времени движения δt . Понятно, что траектория движения как и ее кривизна, будут также зависеть от параметра $\delta t = 2V_0/g$. Так, например, наибольшая высота подъема есть

$$z_{\max} = \frac{g(\delta t)^2}{8}. \quad (100)$$

В результате, как показано на Рис. 13, кривизна высокой траектории, на прохождении которой необходимо больше времени, и кривизна низкой траектории окажутся одинаковыми. Иными словами, на двумерных плоскостях, расположенных в четырехмерной

⁴Этот пример взят из замечательной книги У.Дж. Кауфмана "Космические рубежи теории относительности", М.: Мир, 1981.

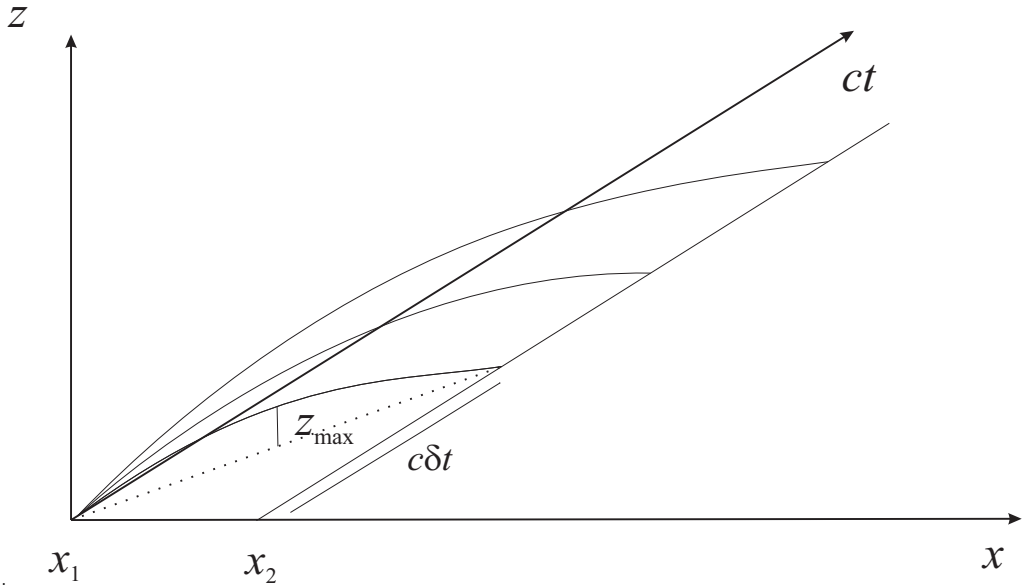


Рис. 13: Траектории движения тел, брошенных из точки 1 в точку 2 в четырехмерном пространстве. Кривизна всех траекторий оказывается одинаковой

пространстве, любая траектория будет представлять собой часть окружности, радиус которой будет одинаков. Глядя на рисунок и учитывая нерелятивистский характер движения (так что $c\delta t \gg z_{\max}$ и $c\delta t \gg x_2 - x_1$), нетрудно сообразить, что

$$R_c = \frac{c^2}{g}. \quad (101)$$

Для гравитационного поля на поверхности Земли он равен двум световым годам. Это, в частности, относится и к световому лучу, для которого в трехмерном пространстве траектория будет исчезающе мало отличаться от прямой линии. Именно эту кривизну, одинаковую для всех движущихся тел, и следует связать с гравитационным полем Земли. Таким образом, еще одной гениальной догадкой Эйнштейна была необходимость рассматривать искривленное четырехмерное пространство. А для описания кривизны такого пространства, как мы уже видели, как раз необходимо ввести симметричный метрический тензор 4×4 .

Следующее наводящее соображение было связано с геодезическими кривыми. Действительно, равенство кривизны четырехмерных траекторий для всех тел, включая и световой луч, наводит на мысль о том, что и сами траектории, т.е. кривые, вдоль которых происходит движение тел, должны обладать какой-то универсальностью. А, как мы видели, такими универсальными кривыми являются геодезические. Еще раз подчеркнем, что речь идет о геодезических в четырехмерном пространстве, т.е. о кратчайших расстояниях, связывающих не две пространственные, а две пространственно-временные точки. Поэтому мерой длины траектории должна быть не трехмерная длина dl , т.е. показатель

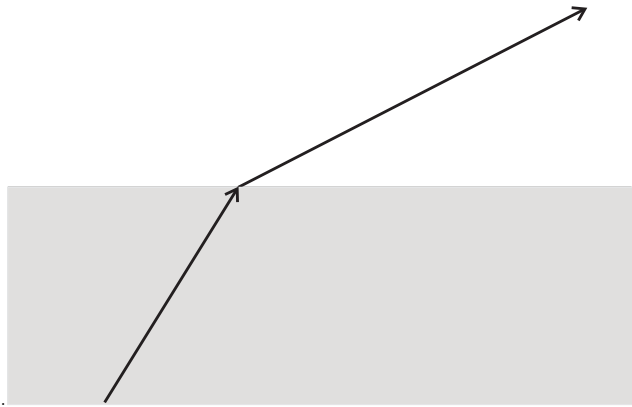


Рис. 14: "Принцип наименьшего действия" на простейшем примере принципа Ферма. Прогрывая на длине траектории, луч света проходит весь путь между начальной и конечной точками за кратчайшее время

пройденного пути на спидометре, а интервал ds (16). Действительно, в рассмотренном нами случае трехмерные длины высокой и низкой траектории между двумя пространственными точками \mathbf{r}_1 и \mathbf{r}_2 могут быть совершенно различными. Поэтому условие минимальности трехмерного расстояния приведет нас к неверному результату – к прямой линии, соединяющей эти две точки. В гравитационном поле такая линия не может быть траекторией реальных тел, включая и световой луч.

Если же мы в качестве показаний спидометра выберем интервал s , то "расстояние" между двумя пространственно-временными точками 1 и 2 окажется наименьшей именно на траектории свободно движущейся частицы. Здесь под "расстоянием" мы должны понимать сумму малых приращений $\sqrt{-(\delta s)^2}$, вычисленных по формуле (16) вдоль разных четырехмерных кривых, соединяющих эти точки. При этом следует иметь в виду, что поскольку скорость тела $|\mathbf{dr}/dt|$ не может превышать скорость света c , для всех траекторий реальных тел должно быть выполнено условие $(\delta s)^2 < 0$. Такие траектории называются времени-подобными. Поэтому под понятием "расстояние" мы выше имели в виду модуль изменения интервала. Мировые же линии, для которых $(\delta s)^2 > 0$ (они называются пространственно-подобными) не могут быть траекторией движения никаких физических тел, поскольку в этом случае движение должно было бы происходить со скоростью, превышающей скорость света. И лишь для светового луча должно быть выполнено условие $(\delta s)^2 = 0$. Не стоит напоминать, что определение геодезической линии также невозможно без знания метрического тензора $\langle g \rangle$.

Здесь можно сделать еще одно замечание исторического характера. Идея о том, что истинные траектории соответствуют кривым, для которых некоторый параметр должен оказаться минимальным, впервые был сформулирован П. Ферма (1601-1665) еще в XVII веке. Согласно принципу Ферма, траектория распространения света между двумя про-

странственными точками соответствует минимальному времени распространения сигнала. В вакууме такими траекториями, будут прямые линии, соединяющие начальную и конечную точки траектории. Если же луч света пересекает границу двух сред, в которых скорости распространения отличаются друг от друга, то, как показано на Рис. 14, происходит всем хорошо известное явление преломления света. При этом величина излома оказывается в точности такой, что, проигрывая на длине траектории, луч света проходит весь путь между начальной и конечной точками за кратчайшее время.

В дальнейшем идея о том, что на реальной траектории тел некоторая величина должна принимать наименьшее из возможных значений, была обобщена П.Л.М. де Мопертюи (1698-1759), Л. Эйлером (1707-1783) и Ж.Л. Лагранжем (1736-1813) на все остальные физические явления и стала одной из ключевых идей теоретической физики. Тем самым, был сформулирован общий рецепт построения теории, известный как принцип наименьшего действия. Именно этот принцип был использован при формулировке уравнений гравитационного поля. Не удивительно поэтому, что, как мы увидим, в рамках построенной таким образом общей теории относительности траектории свободно движущихся частиц являются геодезическими.

Наконец, последним наводящим соображением, позволившим построить последовательную теорию гравитационного поля, стал принцип эквивалентности гравитационной и инертной масс. Мы уже обсуждали это свойство в самом начале. Действительно, еще с опытов Галилея было известно, что все тела движутся в гравитационном поле с одинаковым ускорением g (4) (конечно же, при условии, что мы пренебрегаем силами трения и другими силами, приводящими к выделению энергии). Впоследствии Л. Этвеш (1848-1919) с помощью крутильных весов произвел сравнение восьми разных материалов с платиной, принятой за эталон. Если бы инертная и гравитационная массы различных тел отличались друг от друга, то отличались бы и центробежная (неинерциальная) и центростремительная (гравитационная) силы при их движении относительно Солнца, так что различные тела начинали бы двигаться с относительным ускорением. И тогда, прикрепив разные тела к коромыслу Кавендиша, можно было бы увидеть, что одно из тел повернется по направлению к Солнцу и перекрутит кварцевую нить. На самом деле, однако, нить не перекручивается, и, с точностью $5 \cdot 10^{-9}$, которую дал этот опыт, притяжение двух предметов строго пропорционально центробежному эффекту, который обусловлен инерцией. В 1961-1964 годах опыты Р. Дикке подтвердили равенство обоих видов масс уже с точностью до $3 \cdot 10^{-11}$, а в 1971 году В.Б. Брагинский и В.И. Панов довели точность до 10^{-12} . В опытах с радиоактивными материалами никакого отличия также не было найдено.

Но если точное равенство гравитационной и инертной массы действительно имеет место, то это означает, что если мы перейдем в неинерциальную систему отсчета, свободно

движущуюся в гравитационном поле (всем хорошо известный пример – борт орбитальной станции, хотя сам Эйнштейн рассматривал свободно падающий лифт), то в такой лаборатории гравитационное поле будет, фактически, отсутствовать. Иными словами, гравитационное поле, по крайней мере локально, может быть всегда устранено подходящим выбором системы координат. И никакими экспериментами, проводимыми в такой системе отсчета, мы не сможем определить, является ли наша система отсчета инерциальной, или же она движется с ускорением во внешнем гравитационном поле. И наоборот, наблюдая ускоренное движение пробных тел, экспериментатор не сможет установить, находится ли его лаборатория во внешнем гравитационном поле или же гравитационное поле отсутствует, но сама лаборатория движется с ускорением.

Но если гравитация может быть устранена, по крайней мере в малой окрестности данной точки, переходом в ускоренную систему координат (и наоборот, гравитационное поле может быть симитировано переходом в неинерциальную систему отсчета), то это означает, что при описании гравитационного поля мы должны нарушить одно из главных положений классической механики, а именно предположить, что между инерциальными и неинерциальными системами отсчета нет никаких существенных различий. Поэтому принцип эквивалентности можно сформулировать и как принцип эквивалентности инерциальных и неинерциальных систем отсчета.

Кроме того, принцип эквивалентности приводит к еще одному неожиданному выводу. Действительно, если гравитационное поле может быть устранено соответствующим выбором системы отсчета, то не имеет смысла говорить и о таких величинах, как плотность энергии и плотность импульса гравитационного поля. Иными словами, гравитация должна быть включена в рассмотрение не как некоторая среда в дополнение к материи (тензор энергии-импульса $\langle T \rangle_{\text{part}}$ (67)) и электромагнитному полю (тензор энергии-импульса $\langle T \rangle_{\text{em}}$ (70)), а как-то иначе. Но как?..

3.6.2 Уравнения Эйнштейна

Теперь становится понятна основная идея Эйнштейна. Он предположил, что гравитационные силы, приводящие к ускоренному движению тел, связаны с искривленностью четырехмерного пространства-времени; это свойство, как мы видели, полностью определяется метрическим тензором $\langle g \rangle$. При этом движение тел происходит вдоль геодезических кривых, соединяющих две пространственно-временные точки (\mathbf{r}_1, t_1) и (\mathbf{r}_2, t_2) . Источником же кривизны пространства-времени является материя, т.е. материальные тела и поля не гравитационной природы. Иными словами, ключевое физическое утверждение Эйнштейна (не зависящее от выбора системы отсчета!) состояло в том, что

- гравитация есть проявление кривизны пространства,

- кривизна определяется материей.

Дело теперь было за малым – записать дифференциальные уравнения, выражающие эту ключевую физическую идею. Здесь, однако, возникала серьезная математическая проблема. Дело в том, что физическая необходимость признать эквивалентность инерциальных и неинерциальных систем отсчета потребовала отказаться от лоренц-инвариантности теории, поскольку преобразования Лоренца, строго говоря, относятся лишь к инерциальным системам координат. Поэтому принцип лоренц-инвариантности теории должен быть заменен более общим принципом, позволяющим формулировать уравнения на языке инвариантных величин. Этот принцип называется принципом ковариантности (см. Приложение). Фактически, мы уже встречались с этим свойством на примере выражения (85) для гауссовой кривизны поверхности k , которое, хотя и очень громоздко, выглядит одинаково во всех системах отсчета и дает один и тот же результат.

Вспомним теперь наш пример с кинетической энергией тела и попробуем смоделировать, хотя бы в общих чертах, ход рассуждений Эйнштейна, которые привели его к формулировке уравнений гравитационного поля. Итак, уравнение, описывающее гравитационное взаимодействие, должно было выглядеть как тензорное равенство, связывающее кривизну пространства-времени и материю. При этом правая часть такого уравнения не вызывала затруднений – там должен был стоять полный тензор энергии-импульса среды $\langle T \rangle$. Но что должно быть в левой части уравнения? Здесь и возникли основные трудности.

Дело в том, что, как уже говорилось, в четырехмерном пространстве кривизна не может быть описана одной скалярной функцией. В общем случае полная информация о кривизне содержится в т.н. тензоре Римана $\langle\langle \mathcal{R} \rangle\rangle$, который, как и гауссова кривизна, характеризует изменение суммы углов треугольника (отношение длины окружности к радиусу) в различных двумерных сечениях четырехмерного пространства-времени. И, как и гауссова кривизна, компоненты тензора Римана выражаются через вторые производные метрического тензора по координатам.

Для определения тензора Римана нам понадобится третий способ определения кривизны поверхности, который может быть осуществлен не выходя за ее пределы. Оказывается, как это показано на Рис. 15, в искривленном пространстве при параллельном переносе вектора вдоль замкнутого контура (точное определение этой процедуры нам сейчас несущественно⁵) начальное и конечное направление вектора не совпадают. При этом отличие $\delta\mathbf{A}$ будет прямо пропорционально площади контура δS . Поскольку рассогласование $\delta\mathbf{A}$ совсем не обязательно параллельно самому вектору \mathbf{A} , то, как мы уже знаем, компоненты этих двух (четырёхмерных!) векторов будут связаны некоторой мат-

⁵Если контур состоит из набора геодезических, то параллельный перенос означает, что угол между вектором и геодезической остается постоянным.

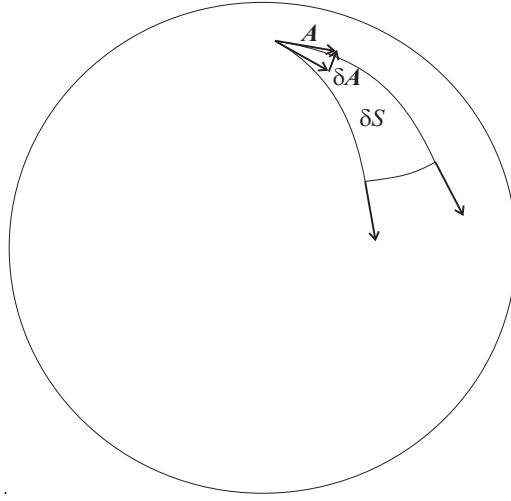


Рис. 15: Параллельный перенос вектора \mathbf{A} по замкнутому контуру (который в этом примере состоит из трех геодезических). В искривленном пространстве начальное и конечное направление вектора не совпадают, причем отличие $\delta\mathbf{A}$ будет прямо пропорционально площади контура δS

рицей $\langle r \rangle$ размером 4×4 , которую можно определить как

$$\delta\mathbf{A} = \langle r \rangle \mathbf{A} \delta S. \quad (102)$$

При этом соотношение (102) будет различным при различных ориентациях контура. Поэтому для определения кривизны пространства нам необходимо знать не один, а несколько таблиц $\langle r \rangle$, соответствующих различным ориентациям контуров δS .

Здесь, однако, возникает еще одно усложнение. Мы привыкли, что элемент площади является вектором, поскольку его ориентация задается вектором нормали. Но это справедливо лишь в трехмерном пространстве, когда можно ввести операцию векторного произведения. Ни в двумерном, ни в четырехмерном пространствах это невозможно. Поэтому в общем случае элемент площади определяется не вектором нормали, а двумя векторами, задающими этот элемент поверхности, т.е. сам является таблицей 4×4 . В результате, для полного определения кривизны нам потребуется знание тензора $\langle r \rangle$ при ориентации контура в плоскости xy , xz , xt , yz и т.д. Поэтому тензор Римана состоит из $4 \times 4 \times 4 \times 4 = 256$ компонент. При этом, поскольку плоскости xy и yx и т.д. совпадают, а плоскости xx и т.д. вырождаются в линию и не дают вклада в рассогласование $\delta\mathbf{A}$, то многие компоненты тензора Римана либо равны нулю, либо выражаются друг через друга. В итоге, четырехмерный тензор Римана содержит 14 физически независимых компонент. В двумерном же пространстве соответствующий тензор $2 \times 2 \times 2 \times 2$ имеет лишь одну независимую компоненту \mathcal{R}_{1212} , т.е. все остальные либо совпадают с \mathcal{R}_{1212} ,

либо равны нулю. Сама же величина \mathcal{R}_{1212} совпадает с гауссовой кривизной

$$\mathcal{R}_{1212} = k. \quad (103)$$

Как бы то ни было, тензор Римана является еще более сложным объектом, чем обсуждавшиеся выше матрицы 4×4 , и поэтому его нельзя было рассматривать в качестве кандидата на роль левой части уравнения гравитационного поля.

С другой стороны, в дифференциальной геометрии можно определить тензор 4×4 , также содержащий информацию о кривизне. Это так называемый тензор Риччи $\langle \mathcal{R} \rangle$, который получается из тензора Римана путем определенного суммирования его компонент⁶. Поэтому компоненты тензора Риччи, как и компоненты тензора Римана, выражаются примерно таким же сложным образом через компоненты метрического тензора $\langle g \rangle$, как и гауссова кривизна k (85). Лучше сказать – еще более сложным образом, поскольку теперь в нашем распоряжении уже не две, а четыре независимых координаты и не три, а десять компонент метрического тензора. Тем не менее, тензор Риччи по-прежнему выражается через вторые производные по координатам от компонент метрического тензора $\langle g \rangle$, и поэтому его размерность см^{-2} совпадает с размерностью гауссовой кривизны k . Этот тензор к тому же удовлетворял и необходимым свойством ковариантности, т.е. он позволял вычислять инвариантные характеристики искривленного пространства-времени. При этом, правда, равенство нулю всех компонент тензора Риччи еще не означает, что кривизна пространства-времени равна нулю. Для этого необходимо более сильное условие равенство нулю тензора Римана.

Надо сказать, что сам Эйнштейн, начиная работать над построением общей теории относительности, потратил не один месяц на изучение основ дифференциальной геометрии, с которой прежде не был знаком. При этом он консультировался у своего друга математика Марселя Гроссмана (1878-1936), и поэтому часто называл для себя тензор Римана тензором Гроссмана. Именно М. Гроссман рассказал Эйнштейну, что уже существующий аппарат дифференциальной геометрии действительно позволяет на математически строгом (и ковариантном!) языке сформулировать физические идеи, которые Эйнштейн хотел заложить в фундамент своей теории. Чтобы отметить роль Гроссмана в построении общей теории относительности, в настоящее время регулярно проводятся большие международные конференции по общей теории относительности и гравитации, посвященные его памяти.

Итак, перед Эйнштейном был набор из трех тензоров $\langle T \rangle$, $\langle \mathcal{R} \rangle$ и $\langle g \rangle$ и основополагающая идея, что кривизна пространства должна определяться материей. Кажется, естественно было бы записать ожидаемое уравнение гравитационного поля в виде

$$\langle \mathcal{R} \rangle = \langle T \rangle. \quad (104)$$

⁶Эта процедура является обобщением оператора Sp.

Конечно, поскольку размерность тензора Риччи (см^{-2}) и размерность тензора энергии-импульса ($\text{эрг}/\text{см}^3$) не совпадают, необходимо было ввести некоторую размерную постоянную, позволяющую связать эти величины. Однако, имея две размерные константы, c и G , сделать это не представляет труда. Как легко проверить, коэффициент пропорциональности должен иметь вид G/c^4 , так что уравнение могло быть записано как

$$\langle \mathcal{R} \rangle = \lambda(G/c^4) \langle T \rangle, \quad (105)$$

где λ – численный коэффициент порядка единицы. Как и в примере о кинетической энергии, он мог бы быть определен из естественного предположения, что в нерелятивистском случае мы должны вернуться к ньютоновской теории Всемирного тяготения.

Однако поскольку в нашем распоряжении имеется еще и метрический тензор $\langle g \rangle$, ситуация становится гораздо более неопределенной. Дело в том, что, как было специально подчеркнуто выше, из любого тензора 4×4 можно сконструировать инвариантный скаляр $\mathcal{R} = \text{Sp} \langle \mathcal{R} \rangle$, представляющий собой сумму его диагональных компонент. Поэтому к левой части уравнения можно было добавить выражение вида $\lambda_1 \mathcal{R} \langle g \rangle$, где λ_1 – еще одна безразмерная постоянная. При этом комбинация скаляр \times тензор не должна вызвать затруднения. Умножение тензора на число производится также, как и умножение на число вектора – все компоненты тензора умножаются на это число.

Такая процедура уже слишком буквально повторяла известную карикатуру, на которой изображен Эйнштейн, стоящий у доски с мелом в руках. Он записал уравнение $E = ma^2$, но оно его не удовлетворило. Зачеркнув его, он написал другое соотношение $E = mb^2$. Опять неудача. И лишь соотношение $E = mc^2$ выглядело убедительно⁷. Все, знающие историю создания общей теории относительности, не станут отрицать, насколько мало шутки в этой шутке.

Конечно, в этой шутке все же значительно больше шутки. Окончательное уравнение было получено Эйнштейном из самых общих соображений, причем именно с использованием принципа наименьшего действия, о котором мы уже упомянули. При этом важно, что это удалось сделать лишь после того, как Эйнштейн убедился, что принцип эквивалентности действительно требует полной независимости всех систем отсчета, в том числе и неинерциальных. Однако Эйнштейн многие годы считал, что все же существует некоторые выделенные системы отсчета. Именно этот пункт не позволял ему закончить построение теории на протяжении нескольких лет! Дело в том, что вид левой части уравнения Эйнштейна получается из принципа наименьшего действия лишь в том случае, если мы не накладываем никаких ограничений на выбор координат.

⁷ Другой вариант: Эйнштейн пишет $E = ma^2 + mb^2 = m(a^2 + b^2) = mc^2$.

В результате, уравнения гравитационного поля, полученные Эйнштейном в ноябре 1915 года, записываются как

$$\langle \mathcal{R} \rangle - \frac{1}{2} \mathcal{R} \langle g \rangle = \frac{8\pi G}{c^4} \langle T \rangle . \quad (106)$$

Это и есть итог десяти лет работы, прошедших с момента создания специальной теории относительности. При этом коэффициент $-1/2$ в левой части уравнения (106) непосредственно следовал из принципа наименьшего действия. Что же касается коэффициента 8π , то он вытекал из условия соответствия, согласно которому в случае слабых гравитационных полей $\varepsilon_g \ll 1$ уравнения Эйнштейна должны переходить в уравнение (21) для гравитационного потенциала, соответствующее закону Всемирного тяготения Ньютона. Правда, более точно было бы сказать, что величина коэффициента $8\pi G/c^4$ должна определяться из наблюдений, как и сами размерные мировые константы c и G .

Действительно, рассмотрим уравнение, возникающие при сравнении величин, стоящих в левых верхних углах всех четырехмерных матриц. Справа в тензоре энергии-импульса там находится плотность энергии вещества, включающая в себя и энергию покоя (см. (69)). Поэтому для нерелятивистских частиц имеем просто $T_{tt} = \rho_m c^2$. Что же касается левой части, то там стоит в общем случае чрезвычайно сложная комбинация производных от компонент метрического тензора $\langle g \rangle$, которая однако для слабых гравитационных полей сводится к производным лишь от компоненты g_{tt} . Воспользовавшись теперь явным видом g_{tt} (99), получаем в сферически симметричном случае

$$\frac{1}{r^2} \frac{d}{dr} \left[r^2 \frac{d}{dr} \left(1 + \frac{2\varphi_g}{c^2} \right) \right] = \frac{8\pi G}{c^4} \cdot (\rho_m c^2). \quad (107)$$

Мы видим, что после устранения из этого уравнения скорости света c (что и должно было быть в нерелятивистской теории) и сокращения на 2 оно действительно переходит в нерелятивистское уравнение (21).

Рассмотренный выше пример показывает, что уравнения Эйнштейна, связывающие симметричные матрицы 4×4 , представляют собой в действительности десять скалярных уравнений. Казалось бы, десятью неизвестными функциями должны быть как раз десять компонент метрического тензора $\langle g \rangle$, поскольку, как уже подчеркивалось, в левой части уравнений стоят вторые (частные) производные по координатам от величин, составляющих метрический тензор $\langle g \rangle$. Однако если мы действительно встали на точку зрения, что физический результат не должен зависеть от выбора системы отсчета, то, как уже было подчеркнуто, соответствующей заменой четырех координат мы всегда можем выбрать четыре компоненты тензора $\langle g \rangle$ произвольным образом. Например, положить четыре из шести недиагональных компонент равными нулю. Поэтому в действительности определяющими оказываются лишь шесть компонент метрического тензора, задающие шесть параметров, определяющих физические свойства пространства-времени. Именно

в этом и состоит принцип общей ковариантности. Следовательно, помимо определения кривизны, уравнения Эйнштейна должны нести еще какую-то дополнительную информацию.

Чтобы разобраться в этом вопросе, необходимо воспользоваться некоторыми общими теоремами дифференциальной геометрии. Дело в том, что левая часть уравнения (106) обладает замечательным свойством. Ее компоненты тождественно удовлетворяют четырем законам сохранения для любого метрического тензора $\langle \mathcal{R} \rangle$. Это означает, например, что изменение во времени величины $\mathcal{R}_{tt} - (1/2)\mathcal{R}g_{tt}$, стоящей в левом верхнем углу на месте плотности энергии среды, тождественно равно потоку вектора, компоненты которого занимают остальные позиции в первой строке тензора $\langle \mathcal{R} \rangle - (1/2)\mathcal{R} \langle g \rangle$. То есть на том самом месте, где в тензоре энергии-импульса находятся компоненты потока энергии. Это свойство носит название тождеств Бианки.

Но если это так, то законам сохранения должен удовлетворять и тензор $\langle T \rangle$, стоящий в правой части уравнения, причем удовлетворять тождественно. Следовательно, мы приходим к важнейшему заключению, что законы сохранения энергии и импульса содержатся в самих уравнениях Эйнштейна. А это означает, что сама материя, приводящая к искривлению пространства-времени, не может двигаться произвольным образом. Поэтому уравнения Эйнштейна автоматически содержат не только законы гравитационного поля, но и законы движения. Следовательно, можно сказать, что общая теория относительности является обобщением как закона Всемирного тяготения, так и законов Ньютона.

Кстати, теперь можно ответить и на вопрос о том, где содержится информация о движении тел вдоль геодезических (конечно, при отсутствии электромагнитных полей; в общем случае движение будет более сложным). Она зашифрована не в левой, а в правой части уравнений (106). Действительно, автоматическое выполнение закона сохранения импульса как раз и означает, что движение тел оказывается свободным. Но свободным не в плоском, а в искривленном пространстве-времени. А это и означает, что движение тел будет происходить по геодезическим.

Здесь, впрочем, совершенно необходимо сделать еще несколько замечаний исторического характера. Прежде всего, нужно прямо сказать, что в 1915 году Эйнштейн не знал о тождествах Бианки. Поэтому тот факт, что сформулированные им уравнения содержат и уравнения движения, стали ясны Эйнштейну несколько позже. Далее, необходимо отметить, что Эйнштейн был далеко не один, кто в десятые годы XX века разрабатывал теорию гравитационного поля. В этой области активно работали М. Абрагам (1875-1922), Г. Нордстрем (1883-1923), Г. Ми (1868-1957). Однако никто из них не догадался выйти за рамки лоренц-инвариантных теорий.

Особенно интересным было взаимодействие Эйнштейна с Д. Гильбертом (1862-1943). Осенью 1915 года Эйнштейн жил в Берлине, а Гильберт – в Геттингене, но они были в дружеских отношениях и переписывались практически каждый день. При этом почта, несмотря на военное время, работала настолько хорошо, что если Эйнштейн за утренним кофе писал Гильберту открытку, то она в тот же день доходила до адресата. Летом 1915 года Эйнштейн гостил у Гильберта в Геттингене, и они достигли полного согласия относительно общих принципов построения теории гравитации. В результате, заинтересовавшись математическими аспектами теории, Гильберт тоже активно включился в работу,.. и был близок к тому, чтобы получить уравнение (106) на неделю раньше самого Эйнштейна.

Нельзя сказать, что Эйнштейн был от этого в восторге, но дружеские отношения были вскоре полностью восстановлены. Дело в том, что Гильберт никогда не претендовал на приоритет в создании общей теории относительности. Он всегда признавал исключительную роль Эйнштейна как автора физической теории гравитационного поля, а свои работы рассматривал как попытку сформулировать уравнения гравитационного поля на строгом математическом языке. Вместе с тем, именно Гильберту принадлежит заслуга последовательного вывода уравнений Эйнштейна из вариационного принципа наименьшего действия, который и излагается теперь во всех учебниках⁸. В результате, взаимодействие двух выдающихся людей позволили не только создать одну из самых изящных физических теорий, но придать ей математически безупречную форму.

В заключение этого раздела необходимо подчеркнуть, что общая теория относительности, сформулированная в виде уравнения (106), не является единственно возможной теорией гравитации даже в рамках самой геометрической идеи Эйнштейна. Дело в том, что, как было специально подчеркнуто в самом начале, общие принципы построения теории позволяют существенно сузить возможный класс моделей, но не всегда определяют их до конца. Оказалось, что к правой части уравнения (106) можно безболезненно добавить слагаемое $\Lambda < g >$, где Λ должна рассматриваться как новая фундаментальная константа. Безболезненно в том смысле, что такое дополнительное слагаемое не нарушает ни один из ключевых принципов, которым должна удовлетворять последовательная теория.

При этом из самого определения ясно, что постоянная Λ имеет такую же размерность см^{-2} , как и стоящий в левой части уравнения тензор Риччи. Поэтому линейный размер $L_\Lambda = \Lambda^{-1/2}$ можно рассматривать как характерный масштаб, на котором роль дополнительного слагаемого становится существенной. Сама же величина Λ , как и все

⁸Как уже отмечалось, принцип наименьшего действия использовал и Эйнштейн, но его вывод не был математически строгим. Дело в том, что знание общих законов не всегда позволяет получить окончательный ответ. Точно также закон сохранения энергии оказывается бесполезным, если мы не знаем правильные выражения для кинетической и потенциальной энергии тела

остальные мировые постоянные, может быть определена лишь из наблюдений. Именно наблюдения и показали, что величина Λ не превышает значений 10^{-54} – 10^{-55} см⁻². Поэтому для физических процессов, масштабы которых ограничены нашей Галактикой (10^{23} см) и даже ближайшими галактиками (10^{25} см), уравнение (106), безусловно, остается справедливым. Лишь в масштабах Вселенной этот дополнительный член начинает играть заметную роль.

С дополнительным слагаемым $\Lambda < g >$ также связана интересная история. Дело в том, что такая добавка существенно меняет структуру космологических решений. Напомним, что в начале XX века считалось, что наша Вселенная является статической, т.е. не изменяется во времени. Именно поэтому Эйнштейн был вынужден добавить слагаемое $\Lambda < g >$ в свое уравнение (106), что позволило получать решения, не зависящие от времени. С другой стороны, в 1922 году русский метеоролог Александр Фридман (1888-1925) показал, что уравнения Эйнштейна (106) с тензором энергии-импульса (71) действительно не имеют статических решений, так что размер Вселенной должен изменяться во времени. А в 1929 году после наблюдений Эдвина Хаббла (1889-1953), измерившего скорости ближайших галактик, стало ясно, что мы живем в расширяющемся мире и наша Вселенная заведомо не является статической. Поэтому Эйнштейн, который, кстати, хотя и не сразу, но все же принял результаты Фридмана, в дальнейшем никогда не рассматривал свои уравнения с поправкой $\Lambda < g >$. Тем не менее, в последние годы появились прямые наблюдательные свидетельства в пользу того, что реальное расширение Вселенной нельзя объяснить на основе уравнения (106). Поэтому сейчас вопрос о природе мировой постоянной Λ вновь широко обсуждается. Однако это уже совсем другая история.

Для нас же здесь будет полезно сделать еще одно замечание об единственности теории гравитации. Начнем издалека. Если теория должна предсказывать связь двух величин, A и B , пусть даже и одинаковой размерности, то совсем не обязательно, что математически она должна выражаться связью $A = B$. Существует бесчисленное количество других связей, например $A = KB^2$ (K – размерная константа) и т.д., которые не совпадают с равенством $A = B$. Если же величины A и B безразмерны, то возможны и теории вида $A = \sin B$, $A = \operatorname{tg} B$ и т.д., совпадающие с теорией $A = B$ лишь при малых A и B , но начинающие существенно отличаться от нее при $A \sim B \sim 1$.

Таким образом, мы видим, что общая теория относительности представляет собой лишь простейшую версию теории, в которой кривизна пространства определяется материей. Удача же заключалась в том, что для уравнения (106), где такая связь предполагалась линейной, были автоматически выполнены все общие требования к последовательной теории, т.е. условие ковариантности, предельный переход к нерелятивистской теории тяготения, которые мы обсуждали выше. Что же касается альтернативных теорий гравитации (а их позже был предложен не один десяток, и не только в рамках геомет-

рической идеи Эйнштейна), то с теоретической точки зрения все они имеют такое же право на существование, как и общая теория относительности. Лишь эксперимент может ответить на вопрос, какая из теорий гравитации является истинно верной. Мы подробно рассмотрим этот вопрос в следующей части нашего повествования.

3.7 Бескрайние возможности

Нет сомнения, что общая теория относительности является вершиной теоретической физики, в чем-то непревзойденной до сих пор. Для самого Эйнштейна, который был прежде всего физиком, ее успех был связан с объяснением векового движения перигелия Меркурия и с точным предсказанием отклонения луча света в гравитационном поле Солнца, которое было подтверждено в 1919 году во время солнечного затмения. Ниже мы подробно обсудим эти эффекты. Однако на теорию Эйнштейна можно посмотреть гораздо шире. Сформулированные им уравнения позволили существенно расширить рамки возможностей теоретической физики.

Одним из таких примеров была космология. Общая теория относительности позволила перевести ее из разряда наблюдательной в область точных наук. Действительно, если предположить, что Вселенная является однородной и изотропной, то удобно выбрать метрику в виде

$$ds^2 = -c^2 d\tau^2 + \frac{dr^2}{1 - \kappa r^2/R_c^2} + r^2(d\theta^2 + \sin^2 \theta d\varphi^2), \quad (108)$$

где τ – собственное время, а $R_c(\tau)$ – т.н. масштабный фактор, зависящий лишь от времени τ . При этом величина κ определяет кривизну пространства ($\kappa = 0$ для плоского мира, $\kappa = 1$ для пространства положительной и $\kappa = -1$ для пространства отрицательной кривизны). В частности, для закрытой Вселенной, имеющей конечный объем, ее пространственная часть подобна двумерной сфере. Именно поэтому в пространственной части метрики появился фактор $(1 - r^2/R_c^2)^{-1}$, полностью эквивалентный соответствующему выражению в метрике двумерной сферы (88). В этом случае $\tau\tau$ и rr компоненты уравнений Эйнштейна мгновенно приводят к уравнениям Фридмана для однородной изотропной Вселенной (точка означает дифференцирование по времени τ)

$$3 \frac{\dot{R}_c^2}{R_c^2} + 3 \frac{\kappa c^2}{R_c^2} = \frac{8\pi G}{c^2} \varepsilon, \quad (109)$$

$$\frac{\ddot{R}_c}{R_c} = -\frac{4\pi G}{3c^2} (\varepsilon + 3P). \quad (110)$$

При этом, конечно же, был выбран вполне определенный вид тензора энергии-импульса, а именно тензор (71). Как уже говорилось, эта система уравнений не имеет статических

решений $R_c = \text{const}$. Если же воспользоваться уравнениями Эйнштейна с Λ -членом

$$\langle \mathcal{R} \rangle - \frac{1}{2} \mathcal{R} \langle g \rangle = \frac{8\pi G}{c^4} \langle T \rangle + \Lambda \langle g \rangle, \quad (111)$$

то уравнения, описывающие эволюцию Вселенной, переписутся в виде

$$3 \frac{\dot{R}_c^2}{R_c^2} + 3 \frac{\kappa c^2}{R_c^2} = \frac{8\pi G}{c^2} \varepsilon + \Lambda c^2, \quad (112)$$

$$\frac{\ddot{R}_c}{R_c} = \frac{\Lambda c^2}{3} - \frac{4\pi G}{3c^2} (\varepsilon + 3P). \quad (113)$$

Здесь уже возможны решения, в которых масштабный фактор Вселенной R_c остается неизменным. Выбирая другие возможные (и невозможные) выражения для тензора энергии-импульса, можно получать другие решения. Их анализ, однако, выходит за рамки нашего изложения.

Другим примером является предсказание существования гравитационных волн. Для их анализа следует предположить, что гравитационные волны есть малое возмущение плоской метрики Минковского $\langle M \rangle = \text{diag}(-1, 1, 1, 1)$ (59), т.е. выбрать метрику в виде

$$\langle g \rangle = \langle M \rangle + \langle h \rangle, \quad (114)$$

где $\langle h \rangle \ll 1$, т.е. все компоненты тензора $\langle h \rangle$ много меньше единицы. В этом случае уравнения Эйнштейна приводят к обычным волновым уравнениям

$$\frac{\partial^2}{\partial \mathbf{r}^2} \langle h \rangle - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \langle h \rangle = 0. \quad (115)$$

Как мы видим, из уравнения (115) автоматически следует, что скорость распространения гравитационных волн совпадает со скоростью света. Это есть прямое следствие общей ковариантности уравнений общей теории относительности (что для малых возмущений метрики Минковского означает ее лоренц-инвариантность), благодаря чему в теории отсутствуют иные постоянные размерности скорости за исключением скорости света.

Что же касается механизма излучения гравитационных волн, то его природа в целом мало отличается от механизма излучения волн электромагнитных. В обоих случаях ключевым моментом является конечность скорости распространения сигнала, благодаря чему наблюдатель, находящийся в точке A в момент времени t (см. Рис. 16), будет регистрировать сигнал, характеризующий источник в момент времени $t - r/c$. Если массы в источнике движутся с ускорением, то на больших расстояниях от источника гравитационное поле будет иметь вид сферической волны, уносящей энергию на бесконечность.

Так, для двух точечных масс, вращающихся по эллиптическим орбитам друг относительно друга, выражение для мощности гравитационного излучения имеет вид (см.

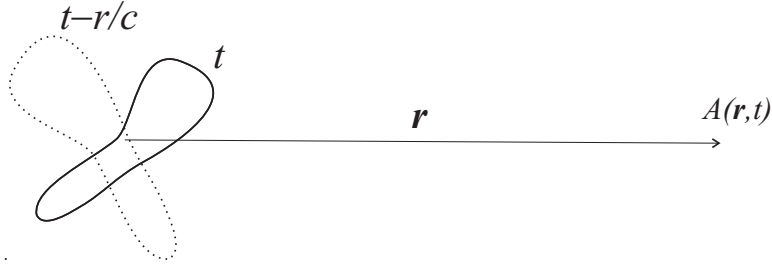


Рис. 16: Запаздывание сигнала при излучении гравитационных волн. Наблюдатель, находящийся в точке A в момент времени t , будет регистрировать сигнал, характеризующий источник в момент времени $t - r/c$

Приложение)

$$W_g = \frac{32}{5} \frac{G^4 M_1^2 M_2^2 (M_1 + M_2)}{c^5 a^5} f(e). \quad (116)$$

Здесь M_1 и M_2 – массы обеих звезд, a – среднее расстояние между звездами, а

$$f(e) = \left(1 + \frac{73}{24} e^2 + \frac{37}{96} e^4\right) (1 - e^2)^{-7/2} \quad (117)$$

– достаточно сложная функция эксцентриситета орбиты e . В частности, для круговой орбиты $f(0) = 1$. В настоящее время построены уже несколько детекторов, способных зарегистрировать гравитационное излучение от реальных космических источников.

Наконец, следствием общей теории относительности было и предсказание черных дыр – совершенно нового класса объектов. Это решение было получено К. Шварцшильдом (1873-1916). Для центрально симметричного гравитационного поля решение уравнения (106) для случая $\langle T \rangle = 0$, т.е. вне гравитирующих масс, имеет вид

$$ds^2 = - \left(1 - \frac{2GM}{c^2 r}\right) c^2 dt^2 + \left(1 - \frac{2GM}{c^2 r}\right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\varphi^2. \quad (118)$$

Особенность этой метрики при $r = r_g = 2GM/c^2$ как раз и приводят к удивительным свойствам черных дыр. Ниже это решение будет подробно проанализировано. Однако здесь следует еще раз напомнить, что физические свойства не должны зависеть от выбора координатной сетки. Поэтому существуют другие координаты, описывающие пространство-время вблизи черной дыры, причем в некоторых из них особенность при $r = r_g$ отсутствует. Поэтому особенность в метрике (118) не соответствует реальной физической особенности в том смысле что никакие инвариантно определенные физические величины (например, плотность газа в системе отсчета, относительно которой газ покоится) не обращаются в бесконечность. В частности, при $r = r_g$ остается конечной и кривизна \mathcal{R} .

В некоторых случаях удобно вместо координаты r использовать координату ρ , такую,

что

$$r = \rho \left(1 + \frac{r_g}{4\rho} \right). \quad (119)$$

В этом случае метрика имеет вид

$$ds^2 = - \left(\frac{1 - r_g/4\rho}{1 + r_g/4\rho} \right)^2 c^2 dt^2 + \left(1 + \frac{r_g}{4\rho} \right)^4 (d\rho^2 + \rho^2 d\theta^2 + \rho^2 \sin^2 \theta d\varphi^2). \quad (120)$$

Как мы видим, пространственная часть метрики с точностью до коэффициента совпадает с метрикой плоского пространства, записанной в сферических координатах. Поэтому такие координаты называются сферическими изотропными.

Наконец, необходимо отметить, что метрика Шварцшильда описывает не только черные дыры, но и метрику всех гравитирующих тел. Только для обычных объектов (звезды, планеты) поправки к плоской метрике оказываются незначительными. Но они и определяют отличие в предсказаниях общей теории относительности и Всемирной теории тяготения Ньютона, к анализу которых мы и переходим.

4 Наблюдательные проявления общей теории относительности

4.1 Эффекты в слабых гравитационных полях

Итак, обсудим эффекты, которые возникают в достаточно слабых гравитационных полях. Это т.н. постньютоновские эффекты, которые связаны с отличием от предсказаний классической теории Всемирного тяготения. Как было показано, поправки к гравитационному потенциалу должны быть порядка

$$\varepsilon_g = \frac{GM}{c^2 R}, \quad (121)$$

т.е. они обратно пропорциональны квадрату скорости света. Уже отсюда следует, что для значимого отличия необходимо либо движение со скоростью, сравнимой со скоростью света, либо близость размера тела к гравитационному радиусу r_g . Однако, как уже говорилось, на поверхности Солнца $\varepsilon_g \approx 2 \cdot 10^{-6}$, а на поверхности Земли $\varepsilon_g \approx 7 \cdot 10^{-10}$. Поэтому в реальных условиях постньютоновские поправки оказываются малыми, и для их выявления обычно требуется кропотливая обработка наблюдательных данных.

Поскольку нашей задачей будет проверка предсказаний общей теории относительности, то хотелось бы сразу учесть все возможные отклонения от предсказаний ОТО. Для этого обычно используются три независимых коэффициента α , β и γ , стоящие перед первыми членами разложения метрики по степеням величины ε_g . Например, в изотропных сферических координатах

$$ds^2 = - \left(1 - \alpha \frac{r_g}{\rho} + \beta \frac{r_g^2}{2\rho^2} + \dots \right) c^2 dt^2 + \left(1 + \gamma \frac{r_g}{\rho} + \dots \right) (d\rho^2 + \rho^2 d\theta^2 + \rho^2 \sin^2 \theta d\varphi^2), \quad (122)$$

а для координат метрики Шварцшильда

$$ds^2 = - \left(1 - \alpha \frac{r_g}{r} + \frac{\beta - \alpha\gamma}{2} \frac{r_g^2}{r} + \dots \right) c^2 dt^2 + \left(1 + \gamma \frac{r_g}{r} + \dots \right) dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\varphi^2. \quad (123)$$

Предсказаниям же общей теории относительности будут соответствовать значения

$$\alpha = \beta = \gamma = 1. \quad (124)$$

Отметим сразу, что величину α можно всегда положить равной единице. Действительно, как уже неоднократно подчеркивалось, ньютоновский потенциал целиком содержится в метрическом коэффициенте $g_{00} = 1 - 2GM/c^2 r$. Поэтому для $\alpha \neq 1$ гравитационный потенциал был бы равен $\varphi_g = -\alpha GM/r$. Но поскольку у нас нет никакого другого способа измерить гравитационную массу тела кроме как через измерение его гравитационного

потенциала φ_g (ускорения свободного падения $g = \alpha GM/r^2$), то фактически в любом таком эксперименте будет измеряться лишь произведение αM . А это и означает, что можно всегда определить массу тела таким образом, чтобы α была равна единице.

Вернемся, однако, к обсуждению основных отличий в предсказаниях общей теории относительности и Всемирной теории тяготения. Для $\varepsilon_g \ll 1$ метрика Шварцшильда (118) переписывается в виде

$$ds^2 = - \left(1 - \frac{2GM}{c^2 r}\right) c^2 dt^2 + \left(1 + \frac{2GM}{c^2 r}\right) dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\varphi^2. \quad (125)$$

При $\varepsilon_g \ll 1$ получаем

$$\varphi_g = c^2 \sqrt{g_{00}} \approx c^2 - \frac{GM}{r} - \frac{1}{2} \frac{G^2 M^2}{c^2 r^2}. \quad (126)$$

Здесь первое слагаемое есть постоянная величина, которая не приводит к появлению дополнительных сил, второе слагаемое в точности соответствует ньютоновскому потенциалу, а третье слагаемое как раз и описывает малое отклонение от предсказаний теории Всемирного тяготения. С другой стороны, как видно из формулы (125), при скоростях, близких к скорости света c , когда длины dr становятся одного порядка с cdt , возникает необходимость принимать во внимание и метрический коэффициент g_{rr} . При этом в метрическом коэффициенте g_{rr} появляется точно такая же поправка $2GM/c^2 r$, как и в величине g_{00} . Таким образом, в слабых гравитационных полях поправки к плоской метрике Минковского в выражениях для g_{00} и g_{rr} будут одинаковыми.

Но самое важное отличие теории относительности от ньютоновской теории состоит в том, что помимо обычной силы гравитации $\mathbf{F} = m\mathbf{g}$ на любые тела начинает действовать сила, зависящая от их скорости \mathbf{v} . При этом в слабом гравитационном поле имеет место замечательная аналогия между гравитационными и электромагнитными уравнениями. Действительно, полную силу $\mathbf{F} = m\mathbf{a}$, действующую на тело массы m , можно переписать в виде

$$\mathbf{F} = m \left(\mathbf{g} + \frac{\mathbf{v}}{c} \times \mathbf{H} \right), \quad (127)$$

которое внешне совпадает с электромагнитной силой (25). Только теперь вместо заряда частицы e нужно подставить ее массу m и, кроме того, вместо магнитного поля \mathbf{B} учесть действие совершенно нового поля \mathbf{H} , которое по аналогии с электромагнетизмом называют гравитомагнитным.

Напомним, что выражение для силы (127) следует из уравнений Эйнштейна, поскольку они, как было показано, содержат не только уравнения для гравитационного поля, но и уравнения движения. Удивительно, что другие компоненты уравнения Эйнштейна в слабом гравитационном поле оказываются полностью эквивалентными уравнениям Максвелла. Эта аналогия настолько красива, что невозможно не привести эти уравнения в строгом математическом виде.

В дифференциальной форме уравнения Максвелла записываются как

$$\operatorname{div} \mathbf{E} = 4\pi\rho_e, \quad (128)$$

$$\operatorname{rot} \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} = 0, \quad (129)$$

$$\operatorname{div} \mathbf{B} = 0, \quad (130)$$

$$\operatorname{rot} \mathbf{B} - \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} = \frac{4\pi}{c} \mathbf{j}. \quad (131)$$

Здесь ρ_e есть плотность электрического заряда, а $\mathbf{j} = \rho_e \mathbf{v}$ – плотность электрического тока. При этом физический смысл этих уравнений хорошо известен. Первое уравнение говорит о том, что источником электрического поля являются заряды. В интегральной форме это и есть теорема Гаусса, о которой говорилось выше. Второе уравнение есть закон Фарадея – ЭДС, наводимая в контуре, определяется скоростью изменения магнитного потока через этот контур. Это свойство используется во всех электростанциях. Далее, третье уравнение говорит о том, что магнитные силовые линии всегда либо замкнуты, либо уходят на бесконечность. А последнее уравнение показывает, что источником стационарного магнитного поля являются электрические токи. При этом величина магнитного поля определяется из закона Био-Савара.

Что же касается уравнений Эйнштейна, то в слабых гравитационных полях они записываются в виде

$$\operatorname{div} \mathbf{g} = -4\pi G\rho_m, \quad (132)$$

$$\operatorname{rot} \mathbf{g} + \frac{1}{c} \frac{\partial \mathbf{H}}{\partial t} = 0, \quad (133)$$

$$\operatorname{div} \mathbf{H} = 0, \quad (134)$$

$$\operatorname{rot} \mathbf{H} - \frac{1}{c} \frac{\partial \mathbf{g}}{\partial t} = -\frac{16\pi}{c} G\rho_m \mathbf{v}. \quad (135)$$

Структура этих уравнений очевидным образом следует из уравнений Эйнштейна. Справа стоят источники поля, содержащиеся в тензоре энергии-импульса. Слева же все многообразие вторых производных от метрических коэффициентов удается свести к первым производным от компонент трехмерных векторов \mathbf{g} и \mathbf{H} , имеющих прозрачный физический смысл. Как мы видим, отличие от уравнений Максвелла состоит лишь в знаках в первом и последнем уравнениях (одноименные заряды в гравитации притягиваются). Иными словами, гравитационное поле \mathbf{g} аналогично электрическому полю, а гравитомангнитное поле \mathbf{H} – магнитному полю \mathbf{B} . При этом в стационарном случае источником гравитоэлектрического поля \mathbf{g} являются массы, а источником гравитомангнитного поля \mathbf{H} – потоки масс. В частности, уравнение (132), следующее из tt компоненты уравнений Эйнштейна, в сферически симметричном случае совпадает с уравнением (23). Наконец, вдали от гравитирующих тел уравнения Эйнштейна

$$\operatorname{rot} \mathbf{g} = -\frac{1}{c} \frac{\partial \mathbf{H}}{\partial t}, \quad (136)$$

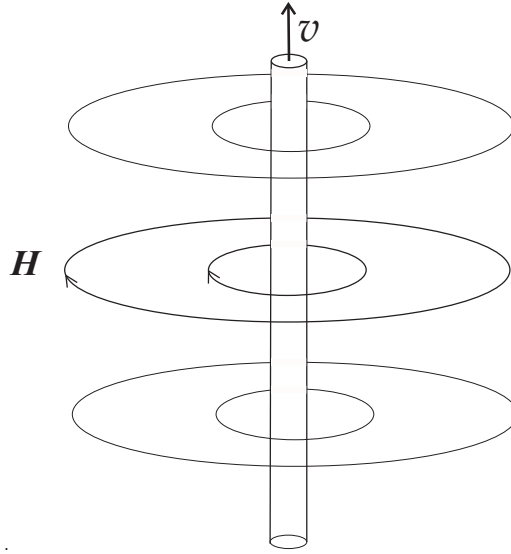


Рис. 17: Гравитомагнитное поле \mathbf{H} вблизи однородного цилиндра, движущегося вдоль оси z со скоростью v

$$\text{rot } \mathbf{H} = \frac{1}{c} \frac{\partial \mathbf{g}}{\partial t}, \quad (137)$$

будут полностью эквивалентны соответствующим уравнениям Максвелла, описывающих электромагнитные волны, распространяющиеся со скоростью c .

Появление дополнительной гравитомагнитной силы и является наиболее важным следствием движения тел в общей теории относительности. Например, покоящийся бесконечно длинный массивный цилиндр (который мы для простоты направим вдоль вертикальной оси z) будет создавать вокруг себя радиальное гравитационное поле

$$\mathbf{g} = -\frac{2G}{r_{\perp}} \mu_g \mathbf{e}_{r_{\perp}}, \quad (138)$$

где μ_g (г/см) – масса на единице длины, r_{\perp} – расстояние от оси z , а $\mathbf{e}_{r_{\perp}}$ – единичный вектор, направленный от оси z . Если же теперь цилиндр будет двигаться вдоль своей оси со скоростью v (см. Рис. 17), то в дополнение к гравитоэлектрическому полю \mathbf{g} появится и гравитомагнитное поле \mathbf{H} , величину которого, согласно (135), можно определить в точном соответствии с законом Био-Савара

$$\mathbf{H} = \frac{8Gv}{r_{\perp} c} \mu_g \mathbf{e}_{\varphi}. \quad (139)$$

При этом основное отличие от случая электромагнетизма состоит в том, что в теории гравитации существуют частицы лишь одного знака заряда, и поэтому гравитомагнитная сила всегда в v^2/c^2 меньше обычной силы гравитации. В электродинамике же возможна ситуация, когда провод с током не содержит полного заряда, т.е. когда существует

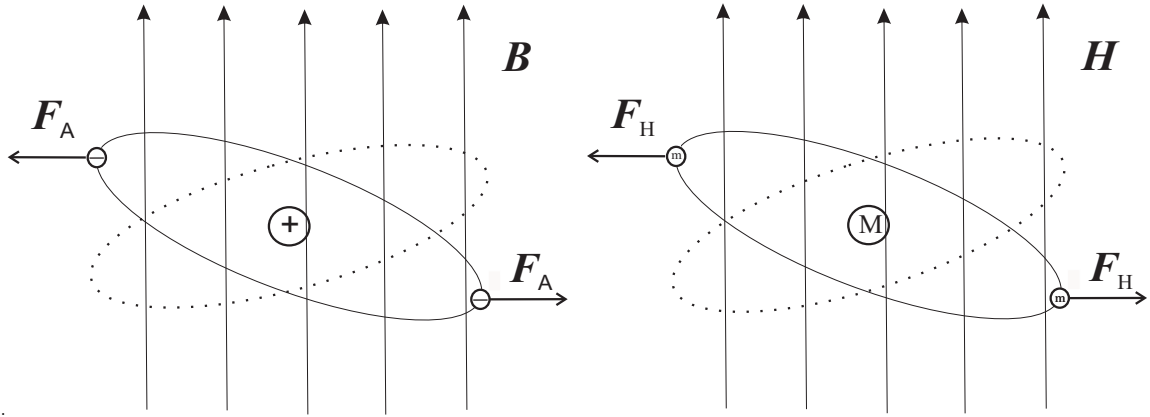


Рис. 18: Прецессия Лармора в магнитном поле \mathbf{B} и прецессия гироскопа в гравитомагнитном поле \mathbf{H} . Момент силы Ампера $\mathbf{r} \times \mathbf{F}_A$, где $\mathbf{F}_A = e\mathbf{v} \times \mathbf{B}/c$ приводит к прецессии орбиты электрона вокруг направления внешнего магнитного поля. Аналогичный эффект имеет место и в гравитационном поле

лишь магнитное поле, а электрическое поле вообще не возникает. Интересно, что два параллельных цилиндра, движущиеся в одном направлении вдоль своих осей, будут не притягиваться, как два параллельных тока, а, наоборот, отталкиваться друг от друга. Однако сила отталкивания, связанная с гравитомагнитным полем \mathbf{H} , всегда будет меньше силы притяжения, связанной с гравитационным полем \mathbf{g} .

Другой пример – вращающийся однородный шар с массой M и моментом импульса $\mathbf{J}_r = 2/5MR^2\Omega\mathbf{e}_z$ будет создавать поля

$$\mathbf{g} = -\frac{GM}{r^2}\mathbf{n}, \quad (140)$$

$$\mathbf{H} = \frac{2G}{c} \frac{\mathbf{J}_r - 3\mathbf{n}(\mathbf{J}_r\mathbf{n})}{r^3}, \quad (141)$$

т.е. обычное поле гравитации и дипольное гравитомагнитное поле.

Наконец, еще одним следствием существования гравитомагнитного поля является гравитомагнитная прецессия гироскопов. Можно сказать, что этот эффект есть дальний родственник знаменитого эффекта Ж.Б.Л. Фуко (1819-1868). Только в опыте Фуко поворот плоскости маятника связан с вращающейся Землей, которая не является инерциальной системой отсчета. В результате появляются неинерциальные силы, приводящие к повороту плоскости маятника. Гравитомагнитная же прецессия возникает за счет появления дополнительных сил, которые обусловлены движением тел; в результате, система отсчета, покоящаяся относительно удаленных звезд, становится неинерционной.

Подобный эффект хорошо известен и в электродинамике и называется прецессией Лармора. Как показано на Рис. 18, сила Ампера $\mathbf{F}_A = e\mathbf{v} \times \mathbf{B}/c$, действующая на электрон, движущийся во внешнем магнитном поле \mathbf{B} по замкнутой орбите, например, в

атоме⁹, приводит к появлению момента сил, приводящего к вращению плоскости орбиты электрона вокруг направления магнитного поля с частотой

$$\Omega_L = \frac{eB}{2m_e c}. \quad (142)$$

Поскольку же, как мы видели, сила, действующая в гравитационном поле на движущееся тело $\mathbf{F}_H = m\mathbf{v} \times \mathbf{H}/c$ подобна силе Лоренца, то такая же прецессия должна иметь место и для вращающегося гироскопа, ось которого не совпадает с направлением гравитомагнитного поля. А так как частота Лармора не зависит от радиуса орбиты (и, значит, каждый элемент массы прецессирует с одинаковой частотой), то угловая скорость прецессии будет одинаковой для любой формы вращающегося тела. Поэтому в выражении (142) вместо заряда e следует просто подставить массу тела m . В результате, частота вращения гироскопа запишется в виде

$$\Omega_g = \frac{H}{2c}. \quad (143)$$

Поскольку и само гравитомагнитное поле \mathbf{H} (141) пропорционально c^{-1} , эффект прецессии также является постньютоновским: $\Omega_g \propto c^{-2}$.

Вместе с тем, нужно четко понимать, что аналогия между гравитацией и электромагнетизмом не является полной. Действительно, мы видели, что ключевым свойством гравитации является возможность перейти в такую неинерциальную систему отсчета, в которой гравитация будет отсутствовать. Ниже мы увидим, что физически выделенными будут те системы отсчета, в которых отсутствует гравитомагнитное поле \mathbf{H} , т.е. те системы, в которых отсутствует вращение гироскопов. С другой стороны, если электромагнитное поле существует в одной системе отсчета, то оно будет существовать и во всех других. Для инерциальных систем отсчета это становится ясно, если мы вспомним о существовании инварианта $I_1 = \mathbf{B}^2 - \mathbf{E}^2$, который, будучи не равным нулю в одной из систем координат, никогда не сможет быть равным нулю в какой-нибудь другой. Такое отличие связано с тем, что законы преобразования при переходе из одной системы в другую для электромагнитных и гравитационных полей отличаются друг от друга.

4.2 Классические постньютоновские эксперименты

4.2.1 Гравитационное красное смещение

Как мы видели, эффект гравитационного красного смещения существует и в ньютоновской теории. Фактически, он означает просто то, что "кинетическая" энергия фотона

⁹Эта сила предполагается малой по сравнению силой кулоновского взаимодействия электрона с атомным ядром.

Таблица 1: Гравитационное красное смещение в спектрах белых карликов. Для удобства результат представлен для эквивалентной скорости v ; $\Delta\nu/\nu = v/c$

Белый карлик	M (M_{\odot})	R (R_{\odot})	$v^{(\text{obs})}$ (км/с)	$v^{(\text{th})}$ (км/с)
Сириус В	1.05 ± 0.03	0.0074 ± 0.0006	89.0 ± 1.6	90 ± 8
40 Эридана В	0.48 ± 0.02	0.0124 ± 0.0005	23.9 ± 1.3	24 ± 1

$h\nu$, как и любого тела, уменьшается при выходе из гравитационного поля. Поэтому для малых ε_g выражение (97) дает для изменения частоты фотона ν , испущенного на поверхности звезды радиуса R и принимаемого на бесконечности

$$\frac{\Delta\nu}{\nu} = \frac{GM}{c^2 R}. \quad (144)$$

Если же фотон излучается на расстоянии R_1 от центра, а регистрируется на высоте R_2 , то тогда

$$\frac{\Delta\nu}{\nu} = \frac{GM}{c^2} \left(\frac{1}{R_1} - \frac{1}{R_2} \right). \quad (145)$$

Впервые гравитационное красное смещение было обнаружено в земной лаборатории. И это несмотря на то, что для характерной разности высот $H = R_2 - R_1 \approx 10$ м изменение частоты фотона $\Delta\nu/\nu$ составляет величину порядка 10^{-15} (поэтому для удобства результат часто представляют для эквивалентной скорости v ; $\Delta\nu/\nu = v/c$). Однако этот фантастически малый эффект был действительно измерен Р. Паундом и Дж. Ребкой в 1960 году для гамма-лучей. Это стало возможным благодаря тому, что гамма-лучи, испускаемые возбужденными ядрами, имеют исключительно малый разброс по энергиям, что позволяет точно определить частоты испущенного и поглощенного фотона. Башня в Гарвардском университете имела высоту 22.5 м, а частота используемых гамма-квантов $\nu = 2.2 \cdot 10^{19} \text{ с}^{-1}$. В этом случае оказалось, что эквивалентная скорость $v = 7.5 \cdot 10^{-5} \text{ см/с}$. При этом измеренное и предсказанное значения $\Delta\nu/\nu$ находились в хорошем согласии

$$\frac{\Delta\nu^{(\text{obs})}}{\Delta\nu^{(\text{th})}} = 1.05 \pm 0.10. \quad (146)$$

В последствии согласие наблюдений с предсказаниями общей теории относительности удалось довести до 0.02 %.

Гравитационное красное смещение было обнаружено и у белых карликов. Здесь также можно провести уверенное сравнение теории и наблюдений, поскольку радиус белого карлика R с хорошей точностью оценивается по светимости звезды L и ее температуре T ($L = 4\pi R^2 \sigma T^4$), а масса M – по орбитальному движению в двойной системе. В табл. 1

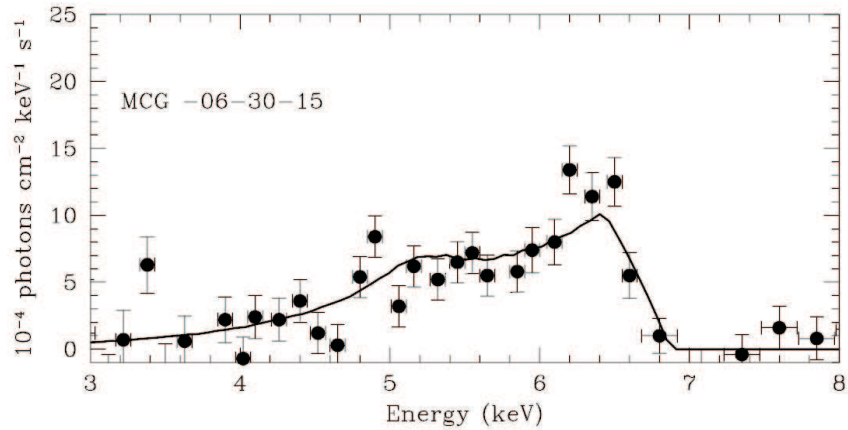


Рис. 19: Профиль линии K_α сильно ионизированных атомов железа, наблюдаемый в спектрах активных галактических ядер

приведены наблюдаемые значения красных смещений и красные смещения, предсказанные по формуле (144).

Понятно, что еще большие смещения должны иметь место у нейтронных звезд. Однако здесь пока проверка предсказаний общей теории относительности проведена быть не может, поскольку в настоящее время не существует надежных методов точного определения радиусов нейтронных звезд. К тому же в тепловом излучении нейтронных звезд до последнего времени не удавалось отождествить линии, частоты которых определялись бы известными атомными переходами между различными уровнями энергии. Циклотронные же линии, связанные с вращением частиц в магнитном поле, не могут быть использованы для определения красного смещения, поскольку точная величина магнитного поля нам неизвестна.

Наконец, не исключено, что гравитационное красное смещение было зарегистрировано в спектрах активных галактических ядер, которые, как будет подробно рассказано ниже, по-видимому являются сверхмассивными черными дырами. Для этого использовалась линия K_α сильно ионизированных атомов железа (т.е. ядра железа с одним электроном), имеющая энергию 6.4 кэВ. Дело в том, что вещество, падающее на черную дыру, должно сильно нагреваться и эффективно излучать. При этом энергии фотонов, наблюдаемых у Земли, которые были испущены вблизи горизонта, должны быть существенно меньше энергии фотонов, излучаемых из внешних областей. Следовательно, профиль линии должен иметь асимметричную форму, показанную на Рис. 19. Сейчас такой профиль линии железа обнаружен уже у нескольких десятков объектов, однако неопределенность в интерпретации не позволяет пока надежно определить параметры черной дыры и проверить предсказания общей теории относительности.

В заключение, отметим, что гравитационное красное смещение может быть сформулировано и на языке хорошо известного "парадокса близнецов". Если один человек проведет какое-то время в более сильном гравитационном поле, то он окажется моложе своего брата, находившегося все время вдали от гравитирующих тел. Справедливость этого замечательного явления также не раз проверялось экспериментально. В одном из таких экспериментов, проведенных Мерилендским Университетом, одни из двух одинаковых цезиевых часов были установлены в самолете, в то время как другие часы оставались на Земле. После пятнадцатичасового полета оказалось, что часы, находившиеся в самолете, ушли вперед на 47 наносекунд. Этот результат находится в превосходном согласии с предсказаниями теории, причем такая величина объясняется комбинацией двух эффектов, связанных с релятивистским изменением хода часов. Действительно, поскольку часы, установленные в самолете, находились на большой высоте и, следовательно, в менее интенсивном гравитационном поле Земли, они ушли вперед на 53 нс (гравитационное красное смещение), но в то же время они отстали на 6 нс, поскольку самолет двигался относительно неподвижных часов (замедление времени).

4.2.2 Отклонение света в поле Солнца

Сам факт отклонения луча света в гравитационном поле также не является свойством лишь только общей теории относительности. Как и любое материальное тело, обладающее энергией (и, значит, массой), луч света должен отклоняться в гравитационном поле. Иными словами, отклонение света существует уже в рамках теории Всемирного тяготения. Этот эффект, по-видимому, был известен и Ньютону, однако впервые четко был сформулирован И.Г. Зольднером (1776-1833) в самом начале XIX века. Простейший расчет показывает, что в этом случае угол отклонения составляет

$$\Delta\theta = \frac{2GM}{c^2 R_{\min}}, \quad (147)$$

где R_{\min} есть минимальное сближение луча света с гравитирующим телом. Для Солнца при $R_{\min} = R_{\odot}$ угол отклонения равен 0.8 угловых секунды.

Однако, как мы видели, при скоростях, близких к скорости света c , необходимо учитывать отличие от единицы не только в метрическом коэффициенте g_{00} , но и в метрическом коэффициенте g_{rr} . Как было показано, в слабых гравитационных полях поправки к плоской метрике Минковского в величинах g_{00} и g_{rr} в рамках ОТО совпадают. При этом обе поправки имеют одинаковый знак. Поэтому для луча света, у которого скорость в точности равна c , имеем

$$\Delta\theta = (1 + \gamma) \frac{2GM}{c^2 R_{\min}}, \quad (148)$$

т.е. для ОТО точно в два раза больше, чем это дает нерелятивистская теория.

Поэтому А.С. Эддингтон (1882-1944), проводивший в 1919 году измерения во время солнечного затмения, имел четкую программу действий. Ему предстояло выяснить, какое из трех утверждений

1. свет не отклоняется в поле Солнца,
2. свет отклоняется на угол $2GM/c^2R$ согласно теории Всемирного тяготения, что составляет 0.874 угловых секунд для края Солнца,
3. свет отклоняется на угол $4GM/c^2R$ согласно общей теории относительности, т.е. на 1.749 угловых секунд,

является истинным. Как показали измерения, проведенные в двух местах, величина отклонения составила 1.98 ± 0.30 и 1.61 ± 0.30 угловых секунд. Иными словами, с точностью порядка 15% результат совпадал с предсказанием ОТО. Собственно говоря, именно после экспедиции Эддингтона, ответившей "да" на третий вопрос, теория относительности Эйнштейна и начала свое триумфальное шествие по миру. В последствии согласие с теорией было существенно улучшено и сейчас мы имеем

$$\frac{1}{2}(1 + \gamma) = 1.001 \pm 0.003. \quad (149)$$

4.2.3 Движение перигелия Меркурия

Как уже говорилось, движение перигелия Меркурия (порядка 43 угловых секунды за столетие) было единственным строго установленным отличием от предсказаний нерелятивистской теории, которое было известно к началу XX века. Физически этот эффект связан с тем, что замкнутая эллиптическая орбита имеет место лишь для ньютоновского потенциала $\varphi_g = -GM/r$ (а также для потенциала пружины $\varphi_g \propto r^2$, который не может быть реализован в астрофизике). Любое небольшое отличие от этого потенциала приводит к тому, что эллипс, по которому движется планета или звезда, сам начинает медленно вращаться, поскольку тело не возвращается точно в ту же точку, из которой оно начинало свое движение.

Точное выражение для углового смещения орбиты тела, масса которой много меньше чем масса центральной звезды M , за один оборот имеет вид

$$\Delta\varphi = 2\pi(2 + 2\gamma - \beta) \frac{GM}{c^2 a(1 - e^2)}, \quad (150)$$

где a – большая полуось орбиты, а e – ее эксцентриситет. Результаты астрономических наблюдений векового смещения перигелия для некоторых планет солнечной системы приведены в табл. 2. Радиолокационные же измерения смещения перигелия Меркурия дают еще большую точность

$$\frac{1}{3}(2 + 2\gamma - \beta) = 1.003 \pm 0.005. \quad (151)$$

Таблица 2: Вековое смещение перигелия для планет солнечной системы

планета	$\Delta\theta^{(\text{th})}$	$\Delta\theta^{(\text{obs})}$	$(2 + 2\gamma - \beta)/3$
Меркурий	43.03	43.11 ± 0.45	1.003 ± 0.005
Венера	8.63	8.4 ± 4.8	0.97 ± 0.55
Земля	3.84	5.0 ± 1.2	1.3 ± 0.3
Марс	1.0	1.3 ± 0.4	1.3 ± 0.4
Икар	10.3	9.8 ± 0.8	0.95 ± 0.08

Как мы видим, здесь также существует полное согласие предсказаний общей теории относительности и наблюдений.

Сейчас вековое смещение перигелия известно и у ряда других планет солнечной системы. При этом, как видно из табл. 2, во всех случаях результаты наблюдений не противоречат предсказаниям общей теории относительности. С другой стороны, остальные планеты не могут соперничать с Меркурием по точности совпадения теории и наблюдений.

4.3 Новые эксперименты

4.3.1 Задержка Шапиро

Бурное развитие наблюдательной астрофизики в начале шестидесятых годов привело не только к возможности более точно проверить классические постньютоновские эффекты, но и сформулировать и проверить новые. Прежде всего, был сформулирован четвертый постньютоновский эффект, который носит имя руководителя проекта радиолокационного зондирования планет И. Шапиро. Если внутренняя планета, Меркурий или Венера, находится в соединении (наименьшем угловом расстоянии), и при этом на противоположной части орбиты, то радиосигнал будет распространяться в непосредственной близости от Солнца, гравитационное поле которого сокращает промежутки времени и удлиняет расстояния. Поэтому сигналу, фактически, нужно пройти большее расстояние, чем то, которое было бы в плоском пространстве, не искаженном гравитационным полем Солнца.

В результате, на элементе длины δl , находящегося на расстоянии r от Солнца, временная задержка должна составить

$$c\delta t_S = (1 + \gamma) \frac{2GM}{c^2 r} \delta l. \quad (152)$$

Просуммировав теперь задержку от всей длины луча, получаем

$$\Delta t_S = (1 + \gamma) \frac{2GM}{c^3} \ln \frac{r_E + r_P + R}{r_E + r_P - R}, \quad (153)$$

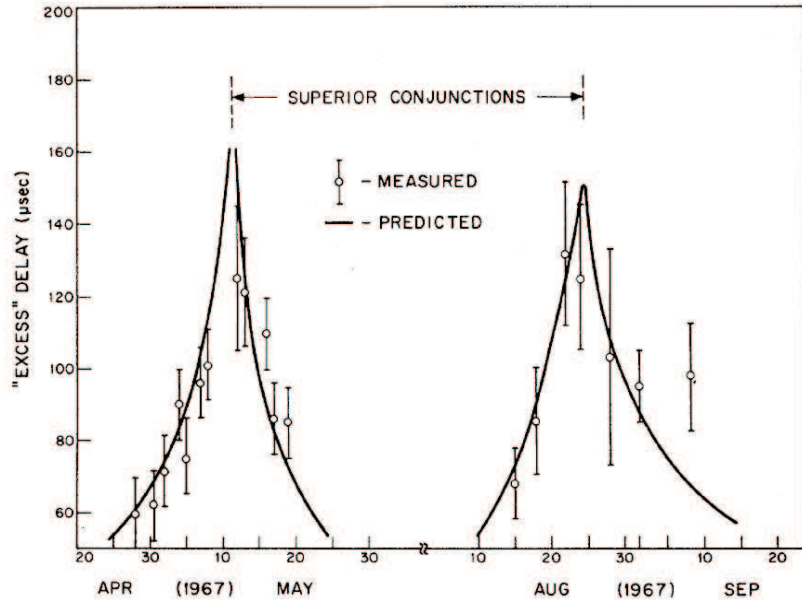


Рис. 20: Задержка Шапиро при радиолокации планет. Сигнал возвратится на Землю чуть позже, чем если бы гравитационное поле Солнца не искажало пространство-время в своей окрестности

где r_E и r_p – расстояние от Солнца до Земли и планеты, а R – расстояние между Землей и планетой. Формально этот эффект пропорционален c^{-3} , однако соотношение (152) показывает, что это такой же эффект $\sim c^{-2}$, как и классические постньютоновские эффекты, а появление дополнительной степени c связан с тем, что нас интересует время задержки, а не безразмерный угол.

Проведенные наблюдения показали, что для Меркурия от элонгации до верхнего соединения время задержки Δt изменяется от 15 до 240 мкс. При этом согласие с предсказаниями ОТО составляет

$$\frac{1}{2}(1 + \gamma) = 1.01 \pm 0.02. \quad (154)$$

Характерный ход временной задержки показан на Рис. 20.

4.3.2 Радиопульсары в тесных двойных системах

Уникальной лабораторией для проверки эффектов общей теории относительности оказались радиопульсары, входящие в состав тесных двойных систем. Действительно, поскольку радиопульсары представляют собой точные часы, движущиеся в гравитационном поле звезды-компаньона, то они позволяют получить уникальную информацию о кривизне пространства-времени. Дело в том, что среди более ста ”двойных” радиопульсаров известны девять случаев, в которых звезда-компаньон также является нейтронной звездой. А в

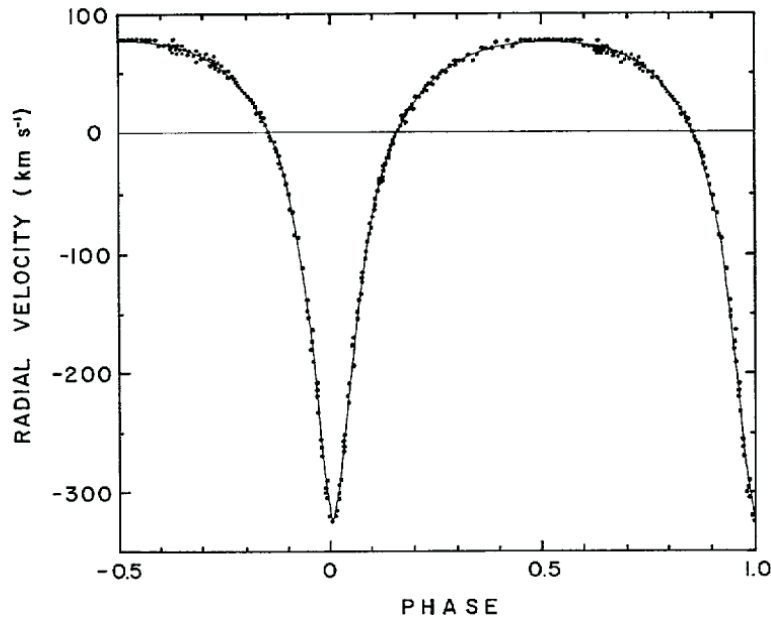


Рис. 21: Кривая лучевой скорости $K_1(t)$ радиопульсара 1913+16, измеренная по эффекту Доплера. Форма кривой позволяет определить не только орбитальный период P_b , но и эксцентриситет e и функцию масс $f(M)$ (см. Приложение)

системе J0737–3039 обе нейтронные звезды наблюдаются как радиопульсары. При этом в семи случаях орбита оказывается достаточно тесной (орбитальный период составляет всего несколько часов), так что в ней удастся зарегистрировать все постньютоновские эффекты, такие как изменение долготы периастра, гравитационное красное смещение и задержку Шапиро. В этом случае удастся определить все элементы орбиты, а также массы обеих звезд. На сегодняшний день это наиболее точное определение массы звезды, достигнутое в астрономии.

В частности, это становится возможным благодаря тому, что у всех пульсаров, входящих в тесные двойные системы, эксцентриситет орбиты оказывается чрезвычайно высок. Поэтому, например, гравитационное красное смещение – изменение угловой скорости вращения нейтронной звезды, измеряемая удаленным наблюдателем – существенно зависит от времени, поскольку радиопульсар движется в переменном гравитационном поле своего компаньона. В итоге, хотя сам эффект и составляет лишь несколько тысячных долей секунды, он может быть уверенно выделен. Здесь свою роль сыграли и чрезвычайно узкие средние профили импульса радиоизлучения (у пульсара 1913+16 – доли миллисекунды), что позволило с большой точностью провести временную привязку.

Напомним, что классические методы наблюдательной астрономии, определяющие с помощью эффекта Доплера радиальную, т.е. параллельную лучу зрения, компоненту

скорости одной из звезд $K_1(t)$, показанную на Рис. 21, позволяют найти лишь часть параметров, характеризующих двойную систему. Это прежде всего орбитальный период P_b и эксцентриситет e , одинаковый у обеих звезд, а также комбинация $a_1 \sin i$, где a_1 есть большая полуось эллипса, по которому движется наблюдаемая звезда, а i – угол наклона плоскости орбиты к картинной плоскости (плоскости, перпендикулярной лучу зрения). Кроме того, определяется еще т.н. функция масс (подробнее см. Приложение)

$$f(M) = \frac{M_2^3 \sin^3 i}{(M_1 + M_2)^2}. \quad (155)$$

Иными словами, не хватает еще трех уравнений для определения всех параметров двойной системы. Три постньютоновские поправки как раз и позволяют получить необходимые дополнительные соотношения.

Прежде всего, это скорость движения периастра (150), которая может быть переписана в виде

$$\dot{\omega} = 3 \left(\frac{P_b}{2\pi} \right)^{-5/3} \left(\frac{G}{c^3} \right)^{2/3} (M_1 + M_2)^{2/3} (1 - e^2)^{-1}. \quad (156)$$

Эта величина была определена практически у всех тесных двойных систем, содержащих нейтронные звезды, поскольку для ее определения достаточно лишь нескольких месяцев наблюдений. Как мы видим, величина $\dot{\omega}$ зависит лишь от полной массы системы, что позволило достаточно точно оценить массы нейтронных звезд. При этом оказалось, что для всех объектов их массы с точностью порядка 10% совпадают с чандрасекаровским пределом $M_{Ch} \approx 1.4M_\odot$. Этот результат, несомненно, имеет фундаментальное значение для теории образования нейтронных звезд. Напомним, что с теоретической точки зрения нейтронная звезда может иметь гораздо более широкий диапазон масс – от $0.1M_\odot$ до $(2 - 3)M_\odot$.

Второе дополнительное соотношение на параметры орбиты дает гравитационное красное смещение, которое необходимо анализировать вместе с поперечным эффектом Доплера (см. Приложение). Амплитуда дополнительной задержки времени прихода импульсов γ_p имеет вид

$$\gamma_p = e \left(\frac{P_b}{2\pi} \right)^{1/3} \frac{G^{2/3}}{c^2} \frac{M_2(M_1 + 2M_2)}{(M_1 + M_2)^{4/3}}. \quad (157)$$

Для самых тесных двойных систем, содержащих радиопульсары, она составляет тысячные доли секунды.

Наконец, еще одним измеряемым эффектом является задержка Шапиро (153), которая может быть переписана в виде

$$\Delta t_S = -2r_S \ln \left[1 - e \cos E - s[\sin \omega(\cos E - e) + (1 - e^2)^{1/2} \cos \omega \sin E] \right]. \quad (158)$$

Здесь E – эксцентрическая аномалия – есть угол, определяющий положение тела на орбите (см. Приложение). В общем случае задержка Шапиро зависит от двух параметров, r_S и

Таблица 3: Основные параметры двойных систем, содержащих две нейтронные звезды

Пульсар	P (сек)	P_b (дни)	e	$\dot{\omega}$ (град/год)	\dot{P}_b (10^{-12})	M (M_\odot)
$J0737-3039A$	0.0227	0.102	0.088	16.88	-1.25(1)	1.3381(7)
$J0737-3039B$	2.7735					1.2489(7)
$J1141-6545$	0.3930	0.198	0.172			1.30(2) 0.98(2)
$J1518+4904$	0.0409	8.634	0.249	0.011		1.56 1.05
$B1534+12$	0.0379	0.421	0.274	1.756	-0.152(3)	1.335(2) 1.344(2)
$J1756-2251$	0.0285	0.319	0.18	2.585		1.40(3) 1.18(3)
$J1906+0746$	0.1441	0.166	0.085	7.57		1.31(5) 1.31(5)
$B1913+16$	0.0590	0.323	0.617	4.227	-2.425(1)	1.4414(2) 1.3867(2)
$B2127+11C$	0.0305	0.335	0.681	4.462		1.349(40) 1.363(40)
$B2303+46$	1.0664	12.34	0.658	0.010		1.30 1.34

s

$$r_s = \frac{GM_2}{c^3}, \quad (159)$$

$$s = \sin i, \quad (160)$$

однако обычно удается уверенно определить лишь амплитуду r_s .

Как видно из табл. 3, благодаря уникальным физическим параметрам, постньютоновские эффекты в четырех системах оказываются чрезвычайно большими. Так, скорость движения периастра $\dot{\omega}$ может достигать нескольких градусов в год, что на четыре порядка превышает аналогичную скорость движения перигелия Меркурия¹⁰. В результате, за 25 лет, прошедших с момента обнаружения двойного пульсара 1913 + 16, его орбита

¹⁰Фактически, сам эффект только на два порядка сильнее, поскольку имеет смысл говорить лишь об изменении положения периастра за один период.

повернулась более чем на 90° . Поражает и точность определения масс нейтронных звезд, достигающая четырех (!) значащих цифр.

Наконец, в трех случаях было зарегистрировано изменение орбитального периода \dot{P}_b , природу которого естественно связать с потерями энергии, связанными с излучением гравитационных волн. Иными словами, впервые удалось, по крайней мере косвенно, подтвердить предсказания общей теории относительности в более высоком порядке ($\propto 1/c^5$) чем постньютоновские поправки ($\propto 1/c^2$). Подчеркнем, что сама возможность проверки связана с тем, что постньютоновские эффекты позволили определить все параметры двойной системы. Лишь в этом случае величины M_1 , M_2 и $a = a_1 + a_2$ могут быть подставлены в формулу Эйнштейна (116) для мощности излучения гравитационных волн, которая и определяет величину \dot{P}_b

$$\dot{P}_b = -\frac{192\pi}{5} \left(\frac{P_b}{2\pi}\right)^{-5/3} \frac{G^{5/3} M_1 M_2}{(M_1 + M_2)^{1/3}} \left(1 + \frac{73}{24} e^2 + \frac{37}{96} e^4\right) (1 - e^2)^{-7/2}. \quad (161)$$

Оказалось, что предсказания общей теории относительности для пульсара 1913 + 16 оказались справедливыми с точностью, составляющей лишь доли процента

$$\frac{\dot{P}_b^{(\text{obs})}}{\dot{P}_b^{(\text{th})}} = 1.0013 \pm 0.0021. \quad (162)$$

Как известно, за работы в этой области Р. Халсу и Дж. Тейлору в 1993 году также была присуждена Нобелевская премия по физике. Действительно, открытие двойной системы, время жизни которой (200 млн. лет) существенно меньше времени жизни Вселенной, стало одним из важнейших событий последних десятилетий. Это означало, что слияние нейтронных звезд должно быть достаточно частым явлением. В настоящее время модель слияния двух нейтронных звезд рассматривается как одна из наиболее вероятных причин космологических гамма-всплесков; этот же процесс представляется наиболее перспективным в связи с проблемой обнаружения гравитационных волн. Удивительно хорошее согласие имеет место и для двойного пульсара J0737–3039, для которого удалось выделить оба параметра, r_s и s , определяющих величину задержки Шапиро и, следовательно, проверить уже два независимых параметра

$$\frac{s^{(\text{obs})}}{s^{(\text{th})}} = 0.9999 \pm 0.0005, \quad (163)$$

$$\frac{\dot{P}_b^{(\text{obs})}}{\dot{P}_b^{(\text{th})}} = 1.003 \pm 0.012. \quad (164)$$

Это стало возможным благодаря более удачному расположению плоскости орбиты двойной системы – она видна практически с ребра – что приводит к существенному увеличению эффектов общей теории относительности при прохождении сигнала вблизи самой поверхности звезды-компаньона.

Что же касается двойной системы, содержащей пульсар 1534 + 12, то здесь ситуация оказалась менее тривиальной. Дело в том, что для этой системы также удалось выделить оба параметра, определяющих величину задержки Шапиро. Но если для величины s теория и наблюдения находятся в хорошем согласии

$$\frac{s^{(\text{obs})}}{s^{(\text{th})}} = 1.010 \pm 0.008, \quad (165)$$

то для скорости изменения орбитального периода предсказание и теория явно не соответствуют друг другу

$$\frac{\dot{P}_b^{(\text{obs})}}{\dot{P}_b^{(\text{th})}} = 0.87 \pm 0.09. \quad (166)$$

Несмотря на это противоречие, авторы все же не склонны связывать его с возможным нарушением самой теории относительности. Дело в том, что на определение параметров двойной системы должно оказывать заметное влияние ее ускорение относительно Земли, которое, в свою очередь, зависит от расстояния до двойной системы. Поскольку же точность определения расстояния до радиопульсаров в настоящее время невелика (она зависит от плотности свободных электронов на луче зрения), то не исключено, что именно это обстоятельство и приводит к наблюдаемому рассогласованию.

Отметим, наконец, что возможность использования формулы Эйнштейна для гравитационного излучения реальных объектов также не является очевидной. В частности, изменение орбитального периода может быть обусловлено приливными эффектами, возникающими за счет конечного размера звезд (формула Эйнштейна справедлива лишь для точечных масс). Но, как было показано, двойная система, состоящая из двух нейтронных звезд, является все же достаточно "чистой" физической лабораторией, так что нейтронные звезды, благодаря своей компактности, можно рассматривать как точечные.

4.3.3 Гравитационные линзы

Еще одним новым физическим явлением (впрочем, тесно связанным с эффектом отклонения света в поле Солнца), открытым в конце XX века, стало обнаружение линзирования далеких квазаров гравитационным полем более близких галактик. Фактически, здесь речь идет об еще одном следствии уравнения (148). Первым на возможность линзирования гравитационным полем указал в 1924 году петербургский физик О.Д. Хвольсон (1852-1934). Публикация же Эйнштейна на эту тему относится лишь к 1936 году. Однако и Хвольсон, и Эйнштейн рассматривали лишь возможность линзирования на звездах. Поэтому они считали, что гравитационное линзирование (т.е. наблюдение второго изображения или же кольца) наблюдаться не может. На возможность наблюдать эффект гравитационного линзирования на далеких галактиках обратил внимание в 1937 году

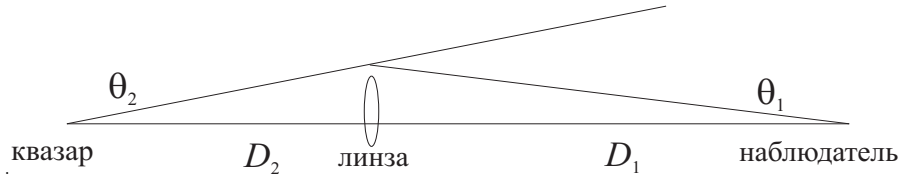


Рис. 22: Определение радиуса Хвольсона-Эйнштейна

Ф. Цвикки (1898-1974). Однако потребовалось еще сорок лет, чтобы наблюдение гравитационных линз стало реальностью. Первая гравитационная линза была обнаружена в 1979 году, а сейчас каталог содержит уже несколько десятков таких объектов (см. табл. 4).

Для определения характерного углового радиуса линзирования следует вновь воспользоваться формулой (148), которая легко позволяет определить угловой радиус кольца Θ_{HE} , возникающего в том случае, если все три тела, источник, линза и Земля, находятся на одной прямой. Поскольку угол отклонения света равен сумме углов θ_1 и θ_2 (см. Рис. 22), то угол $\Theta_{\text{HE}} = \theta_1$, определяющий радиус кольцевого изображения, запишется как

$$\Theta_{\text{HE}} = \left[\frac{4GM}{c^2} \frac{D_2}{D_1(D_1 + D_2)} \right]^{1/2}. \quad (167)$$

Здесь D_1 есть расстояние между наблюдателем и линзой а D_2 – между линзой и источником. Это и есть знаменитое выражение для радиуса кольца Хвольсона-Эйнштейна. Для характерных расстояний до квазаров и масс линзирующих галактик, угловое расстояние между изображениями должно составлять несколько угловых секунд.

В действительности, конечно же, идеально соосный случай никогда не реализуется. Кроме того, сама линзирующая галактика не является точечным объектом. Поэтому обычно изображение далекого квазара представляет собой не кольцо Хвольсона-Эйнштейна, а несколько протяженных образований, число которых зависит от распределения массы в линзе. Существует даже теорема, согласно которой для прозрачной неточечной линзы число изображений должно быть нечетным. Однако во всех случаях формула (167) дает правильную оценку расстояния между изображениями.

К сожалению, неопределенность в расстояниях D_1 и D_2 , как и в массе линзирующей галактики, не позволяет использовать гравитационные линзы для проверки общей теории относительности. Вместе с тем, гравитационные линзы позволяют получить другую, часто уникальную информацию. Например, если излучение квазара сильно переменное (а такая переменность на масштабах времени от дней до нескольких лет и является одним из особенностей этого класса объектов), то по временной задержке можно независимо определять расстояния до квазара и линзирующей галактики. Временная задержка при этом должна составлять от нескольких месяцев до года; у нескольких гравитационных

Таблица 4: Гравитационные линзы

название	число изобр.	$\Delta\theta$ угл. секунд	z_l	z_{QSO}	метод	год
0957 + 561	2	6.1	0.36	1.41	опт./радио	1979
1115 + 080	4	2.3	0.29	1.72	опт.	1980
2016 + 112	3	3.8	1.01	3.27	опт.	1984
2237 + 0305	4	1.8	0.04	1.41	опт.	1985
0142 + 100	2	2.2	0.49	2.72	опт.	1987
3C324	3	2.0	0.84	1.21	опт./радио	1987
1413 + 117	4	1.1	1.40	2.55	опт./радио	1988
1414 + 0534	4	3.1	1.	2.63	опт./радио	1989

линз такая задержка действительно была обнаружена.

4.3.4 Микролинзирование

Качественно, теория микролинзирования полностью совпадает с теорией обычных гравитационных линз. Отличие состоит лишь в том, что термин "микролинзирование" используется в том случае, если как источник света, так и линзирующее тело представляют собой обычные звезды. Для средних расстояний между звездами в нашей Галактике угловое расстояние между изображениями должно быть порядка 10^{-6} угловой секунды. Поэтому, как уже говорилось, и Эйнштейн, и Хвольсон считали, что зарегистрировать линзирование на звездах практически невозможно. Лишь в 1970 году А.В.Бялко показал, что при микролинзировании может существенно изменяться светимость более далекой звезды. В результате становится возможным зарегистрировать событие по непродолжительному усилению ее яркости. Однако вероятность такого события была очень мала, и поэтому казалось, что никакой надежды на прогресс в этой области нет.

Тем не менее, здесь также был достигнут несомненный успех. И заслуга в этом, безусловно, принадлежит Б. Пачиньскому (1940-2007). Он понял, что микролинзирование может быть обнаружено, если одновременно наблюдать большое количество звезд. Оказалось, что при известной плотности звезд в нашей Галактике для регистрирования нескольких событий в год необходимо одновременно наблюдать за 5 миллионами звезд. Причем наблюдать если не ежедневно, то по крайней мере раз в неделю. И, что не менее важно, он также показал, что уровень приемной аппаратуры, достигнутый к концу XX века, позволяет решить эту задачу. Авторитет Б. Пачиньского, известного астрофизика-теоретика, был высок. Задача, поставленная им, чрезвычайно интересна. Игра стоила

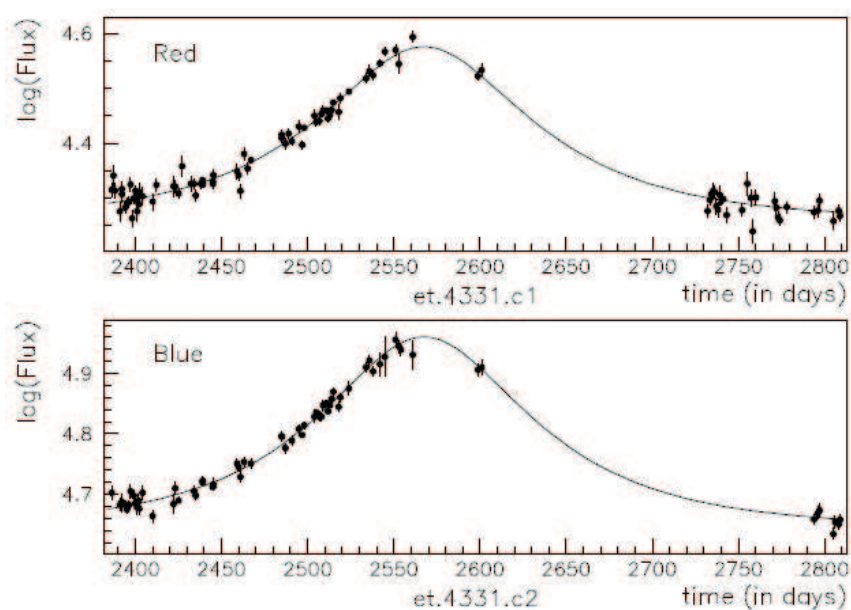


Рис. 23: Характерная кривая блеска при микролинзировании

свеч. Несколько групп организовали наблюдения, и в 1993 году первые объекты микролинзирования были обнаружены. А к 2005 году было зарегистрировано уже несколько десятков таких событий.

На Рис. 23 показана характерная кривая блеска при микролинзировании на звездах нашей Галактики. Линия соответствует теоретическому предсказанию. Как мы видим, кривая блеска должна быть полностью симметрична и иметь характерную колоколообразную форму, которая, в свою очередь, не должна зависеть от длины волны. Этим она существенно отличается от переменных звезд, кривые блеска которых всегда асимметричны. Продолжительность же микролинзирования составляет обычно 20 – 50 суток.

Как и для линзирования на галактиках, микролинзирование не позволяет проводить проверку предсказаний общей теории относительности. С другой стороны, наблюдая объекты микролинзирования можно получить дополнительную информацию как о распределении звезд по массам, так и о свойствах темной материи. Чрезвычайно интересно было бы обнаружить также линзирование на двойных звездах. В этом случае кривая блеска не должна иметь гладкое поведение, показанное на Рис. 23.

4.3.5 Прецессия гироскопов

Наконец, как уже говорилось, еще одним замечательным предсказанием общей теории относительности является прецессия гироскопов. Выше мы рассмотрели прецессию, связанную с дополнительными гравитомагнитными силами, которые возникают за счет движения гравитирующих масс. Понятно, что такая прецессия отсутствует в ньютоновской

теории гравитации, ее нет и в метрике Шварцшильда. Однако если гравитирующее тело, например звезда или даже Земля, вращается вокруг своей оси, то в общей теории относительности возникают силы, приводящие к регулярному повороту оси свободно вращающегося тела. Как можно получить из соотношений (141) и (143), угловая частота вращения на экваторе имеет вид

$$\Omega_g = \frac{1 + \gamma}{2} \frac{GJ_r}{r^3 c^2}, \quad (168)$$

где J_r – момент импульса центрального тела, а r – радиус орбиты. Для Земли скорость поворота оси вращения составляет лишь 0.041 угловых секунды за один год.

Этот эффект, однако, значительно меньше т.н. геодезической прецессии, которая возникает за счет искривления пространства вблизи гравитирующих масс. Иными словами, увеличение длин отрезков приводит не только к вращению эллиптических орбит, но и к повороту оси вращения гироскопа, движущегося по такой орбите. В этом случае угловая скорость вращения гироскопа должна составлять

$$\Omega_{\text{geo}} = \frac{1 + 2\gamma}{2} \frac{GMv}{r^2 c^2}, \quad (169)$$

где v – скорость движения тела по орбите. Для спутника, вращающегося вокруг Земли, угловая частота геодезической прецессии должна быть порядка 6.6 угловых секунды за один год.

Как мы видим, скорость прецессии гироскопов вблизи Земли должна быть чрезвычайно мала. Тем не менее, современный уровень измерений в принципе позволяет выделить подобные эффекты. В 2004 году был запущен спутник Gravity Probe B, специально спроектированный для обнаружения релятивистской прецессии. После 16 месяцев работы эксперимент был успешно завершен, однако точность обработки данных позволила подтвердить лишь наличие геодезической прецессии.

С другой стороны, геодезическая прецессия в тесных двойных системах, содержащих радиопульсары, должна быть значительно больше. Действительно, вращающаяся нейтронная звезда также представляет собой гироскоп, и к ней применимы все формулы, рассмотренные выше. В результате, согласно (169), угловая скорость прецессии нейтронной звезды должна быть порядка одного градуса в год. Следовательно, за несколько лет наблюдений ось вращения радиопульсара может существенно повернуться, в результате чего не исключено, что радиолуч вообще перестанет попадать на Землю. Во всяком случае, геодезическая прецессия должна приводить к достаточно быстрому изменению средних профилей радиоизлучения. В настоящее время имеются четкие указания на такие изменения средних профилей пульсаров, входящих в двойные системы, однако отсутствие точной теории радиоизлучения не позволяет пока определить скорость прецессии из наблюдений.

5 Черные дыры

5.1 Метрика Шварцшильда

Наконец, мы можем перейти к подробному обсуждению свойств черных дыр. При этом нашей основной задачей будет анализ точных количественных соотношений, следующих из уравнений Эйнштейна. Что же касается качественного анализа явлений, в том числе и методики определения масс черных дыр (как солнечных масс, так и сверхмассивных) то их можно найти, например, в недавно изданной книге А.М. Черепашука и А.Д. Чернина "Вселенная, жизнь, черные дыры".

Безусловно, предсказание о существовании черных дыр является одним из наиболее интригующих предсказаний ОТО. Начать хотя бы с того, что черная дыра характеризуется лишь тремя параметрами – массой M , угловым моментом импульса J_r и электрическим зарядом Q (последним, однако, для астрофизических приложений можно пренебречь). Можно сделать черную дыру из холодного или горячего газа, из электронно-протонной или электрон-позитронной плазмы, из любого другого вещества – все равно вся информация об исходном материале будет утеряна.

Естественно начать изложение с наиболее простого случая – невращающейся черной дыры ($J_r = 0$). Она описывается метрикой Шварцшильда (118) – сферически симметричным решением уравнения Эйнштейна для случая $\langle T \rangle = 0$

$$ds^2 = -\alpha^2 c^2 dt^2 + \alpha^{-2} dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\varphi^2, \quad (170)$$

где

$$\alpha^2 = 1 - \frac{2GM}{c^2 r}. \quad (171)$$

Отметим сразу, что метрика (170) записана для определенной координатной сетки, в которой радиус черной дыры соответствует координате $r = r_g$. Но это есть лишь одна из возможных координатных сеток и, кстати, не всегда самая удобная, поскольку она имеет особенность при $r = r_g$. В частности, она не описывает движение свободно падающего наблюдателя, пересекающего горизонт черной дыры. Физические же свойства, как неоднократно подчеркивалось выше, не должны зависеть от выбора координатной сетки. И действительно, можно показать, что инвариантная характеристика – скалярная кривизна пространства-времени \mathcal{R} – при $r = r_g$ остается конечной несмотря на расходимость в метрическом коэффициенте g_{rr} .

Прежде всего, постараемся понять, что же означает сама величина гравитационного радиуса $r_g = 2GM/c^2$, на котором метрика Шварцшильда имеет особенность. Для этого удобно вновь воспользоваться двумерной аналогией, показанной на Рис. 24. Здесь по горизонтальной оси отложена координата r , а само пространство искривлено. В результате, реальное расстояние δl между точками с координатами r и $r + \delta r$, т.е. то расстояние,

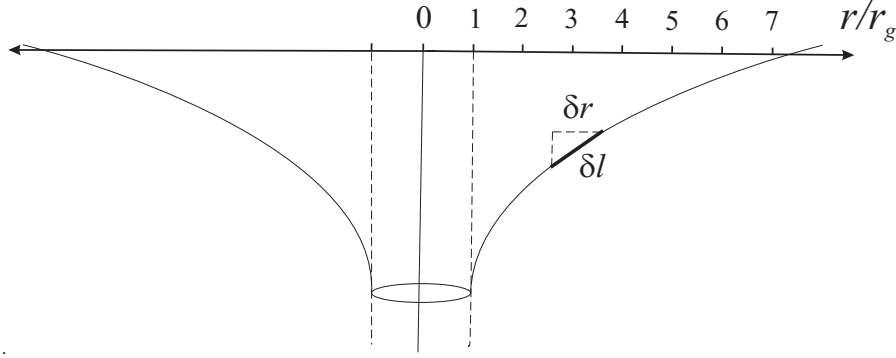


Рис. 24: Структура пространства вблизи черной дыры

которое своим сантиметром измерит наблюдатель, находящийся в этом месте, конечно же, не будет равно δr . Согласно (170), оно будет равно $\sqrt{ds^2}$ при $dt = 0$, $d\theta = 0$ и $d\varphi = 0$, так что

$$\delta l = \frac{\delta r}{\alpha}. \quad (172)$$

Как мы видим, реальное расстояние δl будет больше, чем координатное расстояние δr между этими двумя точками. Поэтому "радиус черной дыры" r_g никак не имеет смысла расстояния между ее центром и поверхностью. С другой стороны, при постоянных t , r и $\theta = \pi/2$ (т.е. на экваторе) имеем, как и в плоском пространстве $ds^2 = r^2 d\varphi^2$, т.е. $\delta l = r \delta\varphi$. Поэтому длина окружности с $r = r_g$ будет равна $L_0 = 2\pi r_g$. Таким образом, правильнее говорить, что черная дыра не есть объект радиуса r_g , а объект, длина экватора которого есть $2\pi r_g$. А величину r_g лучше называть "окружным радиусом".

Соответственно, в гравитационном поле будут сокращаться и промежутки времени

$$\delta\tau = \alpha \delta t, \quad (173)$$

где $\delta\tau$ и δt – соответственно промежутки времени, измеряемые наблюдателем вблизи черной дыры и на бесконечности. Это, как мы видели, в точности соответствует известному "парадоксу близнецов": время в сильном гравитационном поле течет медленнее, чем вдали от гравитирующих тел. В итоге, мы теперь с легкостью можем определить координатную скорость света, распространяющегося вертикально вдоль радиуса r . Действительно, как мы подробно обсуждали выше, инвариантность скорости света означает то, что интервал ds между пространственно-временными точками, проходимыми световым лучом, должен быть равен нулю. При этом скорость $\delta l / \delta\tau$, измеренная любыми наблюдателями, в том числе находящимися вблизи поверхности $r = r_g$, всегда будет равна c . Но это совсем не означает, что остается инвариантной координатная скорость света $V_c = \delta r / \delta t$. Действительно, воспользовавшись выражением (170) для $d\theta = 0$ и $d\varphi = 0$, находим, что условие $ds^2 = 0$ будет выполнено, когда $dr/dt = \alpha^2 c$, так что

$$V_c = \alpha^2 c. \quad (174)$$

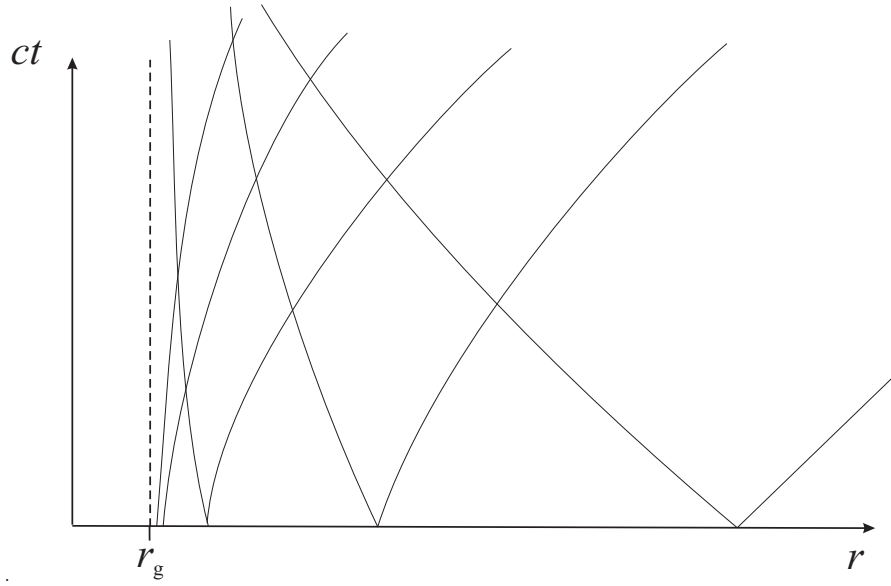


Рис. 25: Траектории лучей света, распространяющихся в гравитационном поле черной дыры

В частности, при $r = r_g$ получается $V_c = 0$. Иными словами, можно сказать, что на поверхности $r = r_g$ гравитация останавливает световой луч, так что никакая информация не может быть передана из области $r < r_g$. Поэтому эта поверхность называется горизонтом событий. Повторим, однако, еще раз, что речь здесь идет только о координатной скорости. Скорость света относительно любых материальных тел, движущихся вблизи черной дыры, по-прежнему останется равной c .

Воспользовавшись теперь соотношением (174), можно определить траекторию луча света, испущенного у поверхности черной дыры. Решая уравнение

$$\frac{dr}{dt} = \pm \left(1 - \frac{2GM}{c^2 r}\right) c, \quad (175)$$

получаем

$$r - r_0 + r_g \ln \left(\frac{r - r_g}{r_0 - r_g} \right) = \pm ct. \quad (176)$$

Здесь r_0 – координата луча света в момент времени $t = 0$, а два знака соответствуют движению луча от и к поверхности черной дыры.

Проанализируем полученное решение. В случае, когда свет излучается на большом расстоянии от черной дыры, можно пренебречь слагаемым, содержащим натуральный логарифм. В этом случае имеем

$$r - r_0 = \pm ct, \quad (177)$$

т.е. траектории лучей света на плоскости (радиус r) – (время t) \times (скорость света c) представляют собой прямые, наклоненные к горизонтальной оси под углом 45° . Они образуют

важнейшую характеристику – световой конус, который определяет границу возможных траекторий любых материальных тел. Действительно, поскольку скорости тел всегда меньше скорости света, их траектории должны лежать в пределах такого светового конуса.

С другой стороны, по мере приближения к горизонту черной дыры координатная скорость распространения света уменьшается, а на самом горизонте и вовсе обращается в ноль. При этом важно, что, согласно соотношению (176), время, которое нужно свету, чтобы пройти от точки с координатой $r_0 \approx r_g$ до точки с координатой r , приближается к бесконечности при $r_0 \rightarrow r_g$

$$\Delta t \approx \frac{r_g}{c} \ln \left(\frac{r - r_g}{r_0 - r_g} \right). \quad (178)$$

Таким образом, информация о том, что происходит вблизи самого горизонта, станет известно удаленному наблюдателю со значительной задержкой. Но это, конечно же, не означает, что само пересечение горизонта космическим кораблем действительно происходит в бесконечном будущем. Просто информация об этом дойдет до нас со значительным опозданием.

Конечно же, замедление времени, описанное выше, соответствует просто задержке Шапиро. Только для черной дыры эта задержка уже не является малой поправкой к времени распространения сигнала, а сама определяет время распространения. С другой стороны, гравитационное поле черной дыры приводит и к гравитационному красному смещению. Иными словами, помимо временной задержки, частота фотонов, излученных на частоте ν_0 в точке с координатой r_0 , будет регистрироваться удаленным наблюдателем на частоте

$$\nu_\infty = \left(1 - \frac{2GM}{c^2 r_0} \right)^{1/2} \nu_0. \quad (179)$$

При этом частота принимаемого сигнала, излученного вблизи горизонта, будет близка к нулю. Именно этот эффект и обсуждался при анализе профиля линии железа 6.4 кэВ, наблюдаемого в некоторых активных галактических ядрах.

На формулу (178) можно посмотреть и с другой стороны. Она показывает, что для удаленного наблюдателя время, которое необходимо свету (и, тем более, любому материальному телу), чтобы достигнуть горизонта, также бесконечно велико. Таким образом, удаленный наблюдатель никогда не увидит, как, например, космический корабль пересекает горизонт черной дыры. Это связано с двумя обстоятельствами. Во-первых, информация об этом достигнет удаленного наблюдателя лишь в бесконечном будущем. Во-вторых, частоты сигналов (и, значит, их энергия) будут настолько низкими, что их невозможно будет зарегистрировать.

Еще одним существенным свойством гравитационного поля черной дыры является отсутствие устойчивых орбит, окружные радиусы которых достаточно малы. В метрике

Шварцшильда радиус такой орбиты дается формулой

$$r_0 = 3r_g. \quad (180)$$

Несмотря на простоту, для вывода этого соотношения нужно провести определенные вычисления. Прежде всего, воспользуемся законом сохранения механической энергии, которое в общей теории относительности можно записать в виде

$$\mathcal{E} = mc^2\alpha\gamma = \text{const}. \quad (181)$$

Здесь γ – лоренц-фактор частицы, измеренный покоящимся наблюдателем, расположенным в той точке, где находится тело. В слабых гравитационных полях и для малых скоростей $v \ll c$ имеем

$$\alpha \approx 1 - \frac{GM}{c^2 r} \quad (182)$$

и

$$\gamma \approx 1 + \frac{1}{2} \frac{v^2}{c^2}, \quad (183)$$

так что

$$\mathcal{E} \approx mc^2 + \frac{mv^2}{2} - \frac{GMm}{r}, \quad (184)$$

что с точностью до энергии покоя mc^2 совпадает с обычным выражением для механической энергии. Поскольку же для движения в экваториальной плоскости определение лоренц-фактора $\gamma^2(1 - v^2/c^2) = 1$ может быть переписано как

$$\gamma^2 = \frac{v_r^2}{c^2}\gamma^2 + \frac{v_\varphi^2}{c^2}\gamma^2 + 1, \quad (185)$$

то закон сохранения энергии можно записать в виде

$$\frac{v_r^2}{c^2}\gamma^2 = \left(\frac{e'}{\alpha}\right)^2 - 1 - \frac{v_\varphi^2}{c^2}\gamma^2. \quad (186)$$

Здесь мы ввели удельную (отнесенную к единице массы) энергию частицы $e' = \mathcal{E}/mc^2$. Кроме того, появившаяся здесь величина v_r соответствует радиальной скорости, измеренной по своим часам наблюдателем, находящимся в точке r , т.е. $v_r = \delta l_r / \delta \tau$. Она связана с координатной скоростью dr/dt соотношением (174)

$$\frac{dr}{dt} = \alpha^2 v_r. \quad (187)$$

С другой стороны, благодаря сохранению удельного момента импульса, должно быть выполнено соотношение

$$L = r v_\varphi \gamma = \text{const}. \quad (188)$$

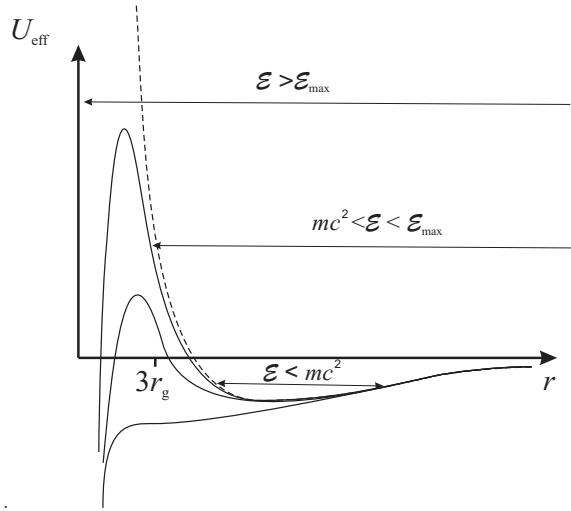


Рис. 26: Эффективный потенциал $U_{\text{eff}}(r)$ для различных значений энергии \mathcal{E} и углового момента L . При $\mathcal{E} < mc^2$, т.е. при $e' < 1$ движение тел финитно, а при $mc^2 < \mathcal{E} < \mathcal{E}_{\text{max}}$, орбита незамкнута, т.е. тело с положительной энергией $mv^2/2 - GMm/r > 0$ вновь уходит на бесконечность. Наконец, при $\mathcal{E} > \mathcal{E}_{\text{max}}$ становится возможным гравитационный захват тел, приходящих из бесконечности. Штриховой линией показан эффективный потенциал (193) в нерелятивистской теории

В итоге, получаем

$$(e')^2 \frac{v_r^2}{c^2} = (e')^2 - V(r), \quad (189)$$

где

$$V(r) = \left(1 - \frac{2GM}{c^2 r}\right) \left(1 + \frac{L^2}{c^2 r^2}\right). \quad (190)$$

В результате, уравнение (189) можно переписать в "нерелятивистской" форме

$$\frac{m}{2} \left(\frac{dr}{dt}\right)^2 + U_{\text{eff}}(r) = \mathcal{E} - mc^2, \quad (191)$$

где

$$U_{\text{eff}}(r) = -\frac{mc^2 \alpha^4}{2(e')^2} [(e')^2 - V(r)] + \mathcal{E} - mc^2 \quad (192)$$

есть эффективный потенциал. В пределе $v \ll c$ величина $U_{\text{eff}}(r)$ совпадает с нерелятивистским выражением

$$U_{\text{eff}}(r) = -\frac{GMm}{r} + \frac{mL^2}{c^2 r^2}, \quad (193)$$

следующим из закона Всемирного тяготения.

Дальнейший анализ проводится совершенно аналогично нерелятивистскому случаю. Поскольку величина $(dr/dt)^2$ должна быть неотрицательной, то при заданной энергии \mathcal{E} тело может находиться лишь в области $\mathcal{E} - mc^2 > U_{\text{eff}}$ (см. Рис. 26). Прежде всего,

при $\mathcal{E} < mc^2$ тело не может уйти на бесконечность, а будет находиться в ограниченной области $r_{\min} < r < r_{\max}$. При этом, в отличие от нерелятивистского движения, в общем случае орбита будет незамкнутой. В частности, круговые орбиты $v_r = 0$ должны соответствовать минимуму эффективного потенциала $U_{\text{eff}}(r)$. Для нерелятивистского случая (штриховая линия на Рис. 26) минимум энергии может существовать при любом радиусе r . Для черной же дыры уменьшение величины $U_{\text{eff}}(r)$ при $r \rightarrow 0$ приводит к тому, что эффективный потенциал не может иметь минимума при достаточно малых радиусах r ни при каких значениях \mathcal{E} и L .

Анализ зависимости U_{eff} от r показывает, что последняя устойчивая орбита существует как раз при $r = 3r_g$. В этот момент не только первая, но и вторая производная эффективного потенциала обращается в нуль. При этом удельная энергия должна быть равна

$$e' = \sqrt{8/9}. \quad (194)$$

Соответственно, энергия связи, приходящаяся на единицу массы, составляет

$$\delta e' = 1 - \sqrt{8/9} \approx 5.7\%. \quad (195)$$

Это есть доля энергии, которая должна быть потеряна частицей, перемещающейся на последнюю устойчивую круговую орбиту. Как мы видим, эффективность выделения гравитационной энергии при этом оказывается существенно выше эффективности ядерного горения, которая не превышает долей процента.

При $mc^2 < \mathcal{E} < \mathcal{E}_{\max}(L)$, где величина $\mathcal{E}_{\max}(L)$ зависит от углового момента L , орбита незамкнута, т.е. тело с положительной (нерелятивистской) энергией $mv^2/2 - GMm/r > 0$ вновь уходит на бесконечность. Наконец, при $\mathcal{E} > \mathcal{E}_{\max}(L)$ в общей теории относительности становится возможным еще один эффект, отсутствующий в ньютоновской теории тяготения. Это гравитационный захват частиц. Действительно, в нерелятивистской теории падение тела на гравитирующий центр может происходить лишь при прямом соударении, когда формально расстояние в афелии (наименьшем расстоянии от гравитирующего центра) становится меньше размера тела¹¹. В противном случае тело должно вновь уйти на бесконечность. Это свойство вновь связано с формой эффективного потенциала (штриховая линия), неограниченно возрастающего при $r \rightarrow 0$.

Однако, как мы видим, в релятивистской теории эффективный потенциал имеет максимум. Поэтому тела с достаточно большой энергией \mathcal{E} (или, более точно, тела с достаточно малым удельным угловым моментом L при заданной величине энергии) будут неограниченно приближаться к черной дыре. В результате, сечение захвата оказывается

¹¹Здесь мы говорим о траектории точечного тела. Тело же конечных размеров за счет приливного взаимодействия, приводящего к нагреву и потере энергии, может быть захвачено и в рамках нерелятивистской теории.

больше, чем чисто геометрические размеры черной дыры. С другой стороны, существуют траектории в районе максимума эффективного потенциала, для которых тело, приходящее из бесконечности, совершает много оборотов вокруг черной дыры и затем вновь уходит на бесконечность.

5.2 Вращающиеся черные дыры

Если черная дыра обладает собственным угловым моментом J_r , то ее свойства становятся еще богаче, поскольку теперь возникают дополнительные эффекты, связанные с действием гравитомагнитных сил. Для сильных гравитационных полей это означает, что в метрическом тензоре $\langle g \rangle$ появляются недиагональные элементы. Обычно метрический тензор (который, напомним, есть решение уравнений Эйнштейна в пустом пространстве $\langle T \rangle = 0$ вне вращающейся черной дыры) записывают в виде

$$ds^2 = -\alpha^2 c^2 dt^2 + g_{rr} dr^2 + g_{\theta\theta} d\theta^2 + g_{\varphi\varphi} (d\varphi - \omega dt)^2. \quad (196)$$

В этих координатах имеет место лишь один перекрестный член $d\varphi dt$. Впервые решение (196) было получено Р. Керром в 1963 году. Поэтому метрику вращающейся черной дыры обычно называют метрикой Керра.

Для вращающейся черной дыры гравитационное красное смещение α , определяющее замедление хода времени, имеет вид

$$\alpha = \frac{\rho_1}{\Sigma} \sqrt{\Delta}, \quad (197)$$

где мы ввели стандартные обозначения

$$\Delta = r^2 + a^2 - 2GM r/c^2, \quad (198)$$

$$\rho_1^2 = r^2 + a^2 \cos^2 \theta, \quad (199)$$

$$\Sigma^2 = (r^2 + a^2)^2 - a^2 \Delta \sin^2 \theta. \quad (200)$$

Здесь и далее мы используем выражение

$$a = \frac{J_r}{Mc}. \quad (201)$$

Оно имеет размерность длины и характеризует удельный угловой момент (т.е. угловой момент, отнесенный к единичной массе). Далее,

$$g_{rr} = \frac{\rho_1^2}{\Delta}, \quad g_{\theta\theta} = \rho_1^2, \quad g_{\varphi\varphi} = \frac{\Sigma^2}{\rho_1^2} \sin^2 \theta. \quad (202)$$

Наконец, угловая скорость

$$\omega = \frac{2aGM r}{\Sigma^2 c}. \quad (203)$$

Несмотря на громоздкий вид метрики Керра, она легко позволяет определить основные характеристики вращающейся черной дыры. Прежде всего, размер черной дыры теперь начинает зависеть от углового момента J_r . Вспоминая, что радиальная координата горизонта черной дыры определяется из условия бесконечного замедления времени $\alpha = 0$, т.е. $\Delta = 0$, получаем

$$r_g = \frac{GM}{c^2} + \sqrt{\frac{G^2 M^2}{c^4} - a^2}. \quad (204)$$

Как мы видим, с увеличением углового момента "радиус" черной дыры уменьшается. При этом наибольший угловой момент черной дыры соответствует значению $a = GM/c^2$. Его можно рассматривать как угловой момент, при котором скорость вращения "поверхности черной дыры" приближается к скорости света. В рассматриваемых здесь координатах "радиус" черной дыры не зависит от широты θ . Однако если мы вычислим гауссову кривизну поверхности горизонта (85), чья двумерная метрика получается из (196) при условии $dt = 0$, $dr = 0$

$$k = \frac{(r_g^2 + a^2)(r_g^2 - 3a^2 \cos^2 \theta)}{r_g^2 + a^2 \cos^2 \theta}, \quad (205)$$

то мы с удивлением увидим, что кривизна поверхности вращающейся черной дыры отличается от $1/r_g^2$. Более того, при $a/M > \sqrt{3}/2$ в районе полюсов гауссова кривизна становится отрицательной. Этот пример еще раз наглядно показывает, насколько нужно быть осторожным, используя те или иные координаты при описания искривленного пространства-времени.

Соответственно, изменяется радиус последней устойчивой орбиты r_0 , которые теперь могут существовать лишь в экваториальной плоскости. К тому же величина r_0 начинает зависеть от направления, в котором вращается тело. Точное выражение имеет вид

$$r_0 = \frac{GM}{c^2} \left[3 + Z_2 \mp \sqrt{(3 - Z_1)(3 + Z_1 + 2Z_2)} \right], \quad (206)$$

где

$$Z_1 = 1 + \left(1 - \frac{a^2 c^4}{G^2 M^2} \right)^{1/3} \left[\left(1 + \frac{ac^2}{GM} \right)^{1/3} + \left(1 - \frac{ac^2}{GM} \right)^{1/3} \right], \quad (207)$$

$$Z_2 = \left(3 \frac{a^2 c^4}{G^2 M^2} + Z_1^2 \right)^{1/2}. \quad (208)$$

Так, для предельного вращения $a = GM/c^2$ получаем $r_0 = GM/c^2$ для прямого и $9GM/c^2$ для обратного вращения. Для невращающейся же черной дыры мы вновь возвращаемся к выражению $r_0 = 6GM/c^2 = 3r_g$. Обратите внимание, что при $a = GM/c^2$ (и для прямого вращения) радиус последней устойчивой орбиты формально совпадает с радиусом

горизонта черной дыры. Однако это лишь кажущееся совпадение. Если вычислить расстояние между горизонтом и последней устойчивой орбитой, то это расстояние окажется бесконечным. На Рис. 24 это означало бы, что пространство вблизи предельно быстро вращающейся черной дыры подобно цилиндру, все точки которого имеют координату r_g . Однако окружность, соответствующая горизонту, находится бесконечно далеко внизу по отношению к окружности, соответствующей последней устойчивой орбите.

В результате, для прямого вращения удельная энергия на последней устойчивой орбите становится равной

$$e' = \sqrt{1/3}. \quad (209)$$

Соответственно, энергия связи, приходящаяся на единицу массы, составляет уже

$$\delta e' = 1 - \sqrt{1/3} \approx 42\%. \quad (210)$$

Таким образом, эффективность выделения энергии в аккреционном диске, окружающем вращающуюся черную дыру, может быть значительно больше по сравнению с шварцшильдовской черной дырой. С другой стороны, для обратного вращения получаем

$$e' = \sqrt{25/27}, \quad (211)$$

так что эффективность выделения энергии здесь не превосходит 4%.

С другой стороны, вращающаяся черная дыра обладает и рядом новых замечательных свойств, связанных с появлением гравитомагнитных сил. Качественно их можно объяснить эффектом "увлечением пространства". Иными словами, можно сказать, что само пространство, как воздух в циклоне, увлекается во вращение вместе с вращающейся черной дырой. Угловая скорость такого вращения ω (203) с хорошей точностью может быть записана в виде

$$\omega = \Omega_H \frac{r_g^3}{r^3}, \quad (212)$$

где величина $\Omega_H = \omega(r_g)$

$$\Omega_H = \frac{c^3}{2GM r_g} a \quad (213)$$

имеет смысл угловой скорости вращения черной дыры. Действительно, для наблюдателей, движущихся с угловой скоростью $d\varphi/dt = \omega$, метрический тензор становится диагональным. Эффект увлечения систем отсчета называют также эффектом Лензе-Тирринга.

На первый взгляд, существование выделенной системы отсчета противоречит самому духу общей теории относительности. Однако вращение по своей сути отличается от линейного движения. Если одна система отсчета движется равномерно и прямолинейно относительно другой, то наблюдатели (каждый в своей лаборатории) не смогут ответить на вопрос, какая из двух систем отсчета покоится, а какая находится в движении. Результаты опытов окажутся идентичными. Если же одна система отсчета покоится, а другая

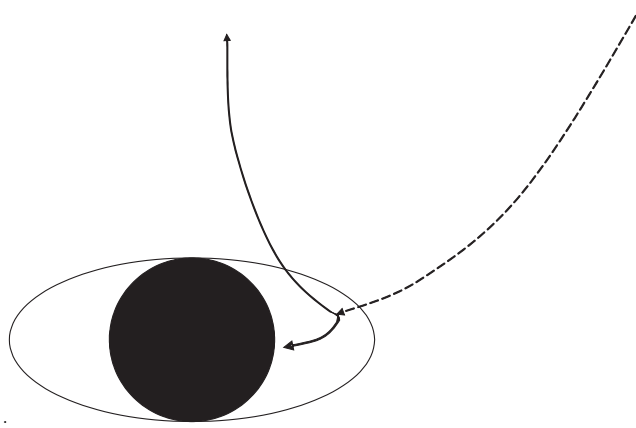


Рис. 27: Эргосфера черной дыры. Показан также процесс Пенроуза, при котором возможно отбирать энергию от вращающейся черной дыры

вращается относительно первой, то неинерциальность одной из систем приведет к появлению дополнительных сил, которые будут немедленно обнаружены. Таким образом, прямой эксперимент (наличие или отсутствие прецессии) позволяет определить те системы отсчета, в которых отсутствует гравитомагнитное поле, т.е. вращение гироскопов.

Но если "увлечение пространства" является физически наблюдаемым эффектом, то это должно привести к ряду новых необычных свойств, которыми будут обладать вращающиеся черные дыры. Действительно, оказывается, что вблизи вращающейся черной дыры никакое тело не может находиться в покое, т.е. в состоянии, при котором $r = \text{const}$, $\theta = \text{const}$, $\varphi = \text{const}$, поскольку в этом случае ему пришлось бы двигаться относительно пространства со скоростью, превышающей скорость света. Более точно, невозможность покоя должно иметь место в области $r_g < r < r_e$, где

$$r_e = \frac{GM}{c^2} + \sqrt{\frac{G^2 M^2}{c^4} - a^2 \cos^2 \theta}. \quad (214)$$

Область в пределах $r_g < r < r_e$ носит название эргосферы, а поверхность эргосферы $r = r_e$ называется пределом статичности (см. Рис. 27).

Действительно, рассмотрим мировую линию тела, соответствующую его покою в точке $r = \text{const}$, $\theta = \text{const}$, $\varphi = \text{const}$. Тогда интервал ds^2 запишется в виде

$$ds^2 = -(c^2 \alpha^2 - g_{\varphi\varphi} \omega^2) dt^2. \quad (215)$$

После несложных преобразований получаем

$$ds^2 = -\frac{S}{\Sigma^2 \rho_1^2} (r^2 + a^2 \cos^2 \theta - 2GM r/c^2) c^2 dt^2, \quad (216)$$

где величина

$$S = r^4 + a^2 r^2 (3 \cos^2 \theta - 1) + a^4 \cos^2 \theta (2 \cos^2 \theta - 1) - 2GM r a^2 \sin^2 \theta / c^2 \quad (217)$$

положительна при $r > r_g$. Поэтому координата r , соответствующая условию $ds^2 = 0$, должна быть найдена из уравнения

$$r^2 + a^2 \cos^2 \theta - 2GMr/c^2 = 0, \quad (218)$$

решение которого и приводит к выражению (214). В итоге, в области $r_g < r < r_e$ квадрат интервала ds^2 будет положительной величиной. С другой стороны, как мы видели, движению материальных тел возможно лишь для времени-подобных траекторий $ds^2 < 0$. А это и означает, что в пределах эргосферы никакое тело не может оставаться в состоянии покоя.

Еще одно удивительное свойство вращающихся черных дыр состоит в том, что вблизи вращающейся черной дыры становится возможным движение с отрицательной (релятивистской) энергией $\mathcal{E} < 0$, т.е. движение, при котором дефект массы больше 100%. Действительно, в метрике Керра энергия частиц \mathcal{E} имеет вид

$$\mathcal{E} = mc^2 \gamma (\alpha + \omega \sqrt{g_{\varphi\varphi}} v_\varphi / c^2). \quad (219)$$

Здесь все величины измеряются наблюдателями, вращающимися с угловой скоростью ω вокруг черной дыры. В результате для частиц, движущихся в направлении, противоположном направлению вращения черной дыры, когда $v_\varphi < 0$ (и вблизи горизонта, где $\alpha \rightarrow 0$) величина \mathcal{E} может стать отрицательной. При этом легко показать, что для частиц, которые движутся по окружности с ультрарелятивистской энергией (т.е. когда $|v_\varphi| \approx c$) условие $\mathcal{E} = 0$ совпадает с условием $c\alpha = \omega \sqrt{g_{\varphi\varphi}}$, которое, как видно из соотношения (215), определяет границу эргосферы (214). За пределами эргосферы энергия тел всегда будет положительной.

Отрицательная энергия тел, находящихся в эргосфере, приводит к тому, что угловая скорость вращения черной дыры может быть уменьшена. А это значит, что энергия вращения может быть передана окружающим телам¹². Действительно, если частица с отрицательной энергией пересечет горизонт событий, то энергия черной дыры должна уменьшиться. Ясно, однако, что тело с отрицательной энергией нужно приготовить в самой эргосфере, поскольку механическая релятивистская энергия тела \mathcal{E} вдали от черной дыры всегда положительна. Кроме того, лишь в этом случае вращательная энергия черной дыры может быть передана на бесконечность. Процесс выделения энергии вращающейся черной дыры называется процессом Пенроуза.

Простейшая реализация процесса Пенроуза состоит в следующем (см. Рис. 27). Пусть на большом расстоянии от черной дыры мы имеем систему двух тел, соединенных сжатой пружиной, полная энергия которой (включая и энергию сжатой пружины) есть \mathcal{E}_{in} .

¹²Именно поэтому эргосфера (т.е. область, где может быть совершена работа) и получила свое название.

Понятно, что вместо пружины может быть использован любой другой источник энергии, например, ракетный двигатель. Двигаясь в гравитационном поле вращающейся черной дыры, система попадает в эргосферу, где срабатывает спусковой механизм. В результате, тела начинают двигаться по разным траекториям. При определенных условиях, если энергия одно из тел, падающих на черную дыру, оказывается отрицательной, второе тело может приобрести достаточную энергию, чтобы уйти на бесконечность. При этом энергия тела \mathcal{E}_{out} , ушедшего на бесконечность, будет больше начальной энергии \mathcal{E}_{in} . Увеличение же энергии тела, ушедшего на бесконечность, компенсируется уменьшением энергии вращения черной дыры.

Правда, точные расчеты показывают, что для выделения энергии вращения черной дыры относительная скорость тел в момент их разлета в эргосфере должна превышать половину скорости света. Поэтому эффект Пенроуза вряд ли можно будет использовать в качестве реального источника энергии. С другой стороны, замедление вращения черной дыры может быть реализовано для электромагнитных волн. Иными словами, если электромагнитная волна рассеивается на черной дыре, так что часть, обладающая отрицательной энергией, пересекает горизонт событий, то энергия уходящей волны может превышать энергию волны, падающей на черную дыру. Как будет показано ниже, этот процесс может играть ключевую роль при выделении энергии в центральных областях активных галактических ядер.

5.3 Астрофизика черных дыр

В настоящее время уже мало кто сомневается в существовании черных дыр. На сегодняшний день обнаружены уже несколько сот космических источников, в которых предполагается существование этих необычных объектов, предсказываемых общей теорией относительности. Тем не менее, необходимо сразу отметить, что во всех случаях речь идет не о прямом наблюдении черных дыр, а об их влиянии на движение окружающих тел.

Действительно, поскольку черная дыра сама излучать не может, то мы можем узнать о ее существовании либо определяя свойства вещества, аккрецирующего на черную дыру, либо по движению звезд в ее окрестности. Поэтому до сих пор нет ни одного кандидата, представляющего собой одиночную черную дыру. Все кандидаты либо входят в состав двойных систем, в которых происходит перетекание вещества со звезды-компаньона, либо находятся в центрах галактик, активность которых как раз и обусловлена аккрецией окружающего газа. Но до сих пор, за исключением линии железа 6.4 кэВ, о которой говорилось выше, нет ни одного наблюдения, когда излучающее вещество находится в непосредственной близости от горизонта событий, где гравитационный потенциал φ_g близок к c^2 . Во всех случаях гравитационное поле интересующего нас объекта достаточно

слабо, так что оно вполне может быть описано в рамках ньютоновской теории Всемирного тяготения. Как мы увидим, в большинстве случаев доказательство того, что этот объект может являться черной дырой, фактически, проводится от противного (ибо нельзя придумать более разумного объяснения).

5.3.1 Галактические черные дыры солнечных масс

Черные дыры солнечных масс могут образовываться на последней стадии эволюции массивных звезд, например, как результат взрыва сверхновых. Если масса центрального ядра предсверхновой превышает $10M_{\odot}$, то силы гравитации не могут быть скомпенсированы давлением среды. Начинается катастрофическое сжатие внутренних областей звезды, что и приводит к образованию черных дыр. Их массы при этом должны быть больше $3M_{\odot}$; при меньших массах образуется нейтронная звезда, в которой силы гравитации уравновешиваются давлением нейтронного вещества. Другим каналом образования черных дыр являются тесные двойные системы, в которых в результате аккреции на нейтронную звезду ее масса становится больше предела их устойчивости $\sim 3M_{\odot}$. Понятно, что характерные массы таких черных дыр должны составлять $(3 - 30)M_{\odot}$. Именно такие массы и имеют галактические кандидаты в черные дыры.

Как уже подчеркивалось, все кандидаты в черные дыры солнечных масс входят в состав тесных двойных систем, в которых происходит перетекание вещества с нормальной звезды на релятивистский объект. Лишь в этом случае удастся достаточно точно оценить их массу. Нагрев аккрецирующего газа, связанный с выделением гравитационной энергии, приводит к тому, что все такие объекты являются мощными рентгеновскими источниками. Как уже говорилось, если сам релятивистский объект не регистрируется, а наблюдается только его компаньон, то можно определить лишь т.н. функцию масс

$$f(M) = \frac{M_X^3 \sin^3 i}{(M_X + M_V)^2}. \quad (220)$$

Поскольку $\sin i < 1$, а $M_X + M_V > M_X$, то функция масс определяет нижнюю границу массы релятивистского объекта. Дальнейший анализ сводится к попытке оценить угол i и массу видимого компаньона M_V из наблюдений (затмения, спектр видимого компонента, нагрев компаньона рентгеновским излучением компактного объекта). В таблице показано, что все кандидаты имеют массы, превосходящие $3M_{\odot}$, что и является основным критерием их отбора.

При этом важно, что все кандидаты обладают свойствами, которые можно ожидать лишь для объектов, содержащих черные дыры. Так, ни один из кандидатов не проявляет себя как рентгеновский пульсар – источник пульсирующего рентгеновского излучения. Для этого аккреция должна идти на нейтронную звезду с сильным дипольным магнитным полем. Вблизи нейтронной звезды, где плотность энергии магнитного поля стано-

Таблица 5: Кандидаты в галактические черные дыры солнечных масс

источник	P_b (сут.)	$f(M)$ (M_\odot)	M_X (M_\odot)	M_V (M_\odot)
Cyg X-1	5.5	0.24 ± 0.01	16 ± 5	33 ± 9
LMC X-3	1.7	2.3 ± 0.3	9 ± 2	6 ± 2
LMC X-1	4.2	0.14 ± 0.05	7 ± 3	22 ± 4
A0 620-00	0.3	2.91 ± 0.08	10 ± 5	0.6 ± 0.1
GS 2023+338	6.5	6.08 ± 0.06	12 ± 2	0.7 ± 0.1
GS 2000+25	0.3	4.97 ± 0.10	10 ± 4	0.5 ± 0.1
GRO J0422+32	0.2	1.13 ± 0.09	10 ± 5	0.4 ± 0.1
GRO J1655-40	2.6	2.73 ± 0.09	6.3 ± 0.5	2.4 ± 0.4
H 1705-250	0.5	4.86 ± 0.13	6 ± 1	0.4 ± 0.1
4U 1543-47	1.1	0.22 ± 0.02	5 ± 1	2.5 ± 1.0
GRS 1009-45	0.3	3.17 ± 0.12	4.1 ± 0.5	0.6 ± 0.1
SAX J1819.3-2525	2.8	2.74 ± 0.12	9.6 ± 1.1	6.5 ± 1.0
XTE J1118+480	0.175	6.1 ± 0.3	6.8 ± 0.9	0.3 ± 0.2
GRS 1915+105	33.5	9.5 ± 3.0	14 ± 4	1.2 ± 0.2

вится больше плотности энергии плазмы, аккрецирующее вещество начинает двигаться вдоль магнитных силовых линий, т.е. стремиться к магнитным полюсам, где и происходит основное энерговыделение. Вращение же нейтронной звезды приводит к тому, что удаленный наблюдатель будет регистрировать переменное рентгеновское излучение.

Иными словами, для того, чтобы компактный объект мог проявлять себя как рентгеновский пульсар, необходимо наличие твердой поверхности и сильного дипольного магнитного поля. Однако, как будет показано в следующем разделе, черная дыра не может обладать собственным дипольным магнитным полем. С другой стороны, черные дыры не имеют и твердой поверхности, так что они поглощают не только вещество, но и излучение. Поэтому отсутствие пульсирующего излучения в объекте, масса которого превышает $3M_\odot$, является "доказательством от противного" того, что здесь, по-видимому, находится черная дыра.

5.3.2 Сверхмассивные черные дыры в центрах галактик

Другим, более многочисленным классом кандидатов в черные дыры являются ядра активных галактик. Если постараться кратко сформулировать основные черты централь-

ной машины, ответственной за их активность, то в настоящее время большинство астрофизиков склоняется к следующей картине. В центре родительской галактики находится сверхмассивная черная дыра (ее масса достигает $10^6 - 10^9 M_{\odot}$), на которую происходит аккреция окружающего вещества. Далее, обычно предполагается, что аккреция вещества имеет дисковый характер. Только в этом случае удастся объяснить как чрезвычайно высокую эффективность переработки массы падающего вещества в наблюдаемое излучение, так и компактность центральной машины. При этом энергетическим источником активности галактических ядер может быть как энергия вращения черной дыры

$$E_{\text{tot}} = \frac{I_r \Omega^2}{2} \sim 10^{63} \left(\frac{M}{10^9 M_{\odot}} \right) \left(\frac{a}{M} \right)^2, \quad (221)$$

так и энергия аккрецирующего вещества.

Кроме того, в пространстве естественным образом возникает выделенное направление – ось вращения – вдоль которого и происходит формирование струйных выбросов, о которых пойдет речь ниже. Многочисленные данные наблюдений действительно указывают на присутствие в центрах галактических ядер сверхмассивного центрального объекта. В следующем разделе мы подробно обсудим механизм энерговыделения. Здесь же мы кратко обсудим механизмы образования сверхмассивных черных дыр, а также сформулируем их основные свойства и наблюдательные проявления.

В настоящее время существует несколько точек зрения на вопрос о том, образовались ли такие сверхмассивные черные дыры уже после возникновения галактик, в самом процессе их образования, или же черные дыры существовали задолго до появления галактик и сами способствовали их образованию. Согласно современным представлениям, массивные центральные объекты присутствуют в большинстве галактик и становятся активными, когда на них в достаточном количестве падает вещество (“горючее”). Вне зависимости от содержания межзвездного газа, вблизи центрального объекта всегда имеет место процесс, обогащающий веществом его окрестности — это разрыв звезд, пролетающих вблизи черной дыры, за счет приливного возмущения. Оценки показывают, что в обычной галактике такое событие случается раз в 10^4 лет. Помимо того, что этот процесс обогащает аккреционный диск вокруг центрального объекта, индивидуальное событие (длительностью около нескольких месяцев) может быть зарегистрировано, и такие наблюдения могли бы предоставить еще одну возможность исследования ближайших окрестностей центрального объекта. В настоящее время ведутся поиски таких событий, однако возможности этого метода ограничены из-за сложности возникающих при этом газо-динамических расчетов, поэтому вряд ли можно надеяться, даже при наличии соответствующих наблюдений, вычислить параметры центрального объекта (скорость вращения черной дыры, угол наклона и массу диска и т.п.).

Как и для галактических черных дыр солнечных масс, сверхмассивные черные дыры регистрируются по их гравитационному воздействию на окружающий газ и звезды. При

этом, как и в случае маломассивных черных дыр, угловое разрешение современных приемников позволяет строить лишь изображения, на много порядков превышающее размер черной дыры. Тем не менее, прогресс в угловом разрешении, достигнутый в последние годы, позволил существенно продвинуться в понимании процессов, происходящих в этих объектах.

Прежде всего, это касается черной дыры, находящейся в центре нашей Галактики. Ее масса оценивается как $2 \cdot 10^6 M_{\odot}$. Однако долгое время наблюдения показывали лишь то, что такая масса находится объеме, линейный размер которого составляет 1 парсек ($3 \cdot 10^{18}$ см). В принципе, такое же воздействие на окружающий газ и звезды могло оказывать компактное звездное скопление. Лишь в последние годы, когда удалось проследить трехмерное движение отдельных звезд, находящихся в непосредственной близости от черной дыры, стало ясно, что размер области, содержащей массу $2 \cdot 10^6 M_{\odot}$, не превышает 0.003 парсека. По современным представлениям время жизни звездного скопления, размер которого составляет 0.003 парсека, должен быть меньше времени жизни Галактики. По истечению этого времени размер звездного скопления за счет гравитационного взаимодействия звезд друг с другом неизбежно должен стать настолько малым (а его плотность, соответственно, настолько большой), что столкновение звезд все равно должно было привести к образованию сверхмассивной черной дыры.

Что же касается сверхмассивных черных дыр в ядрах других галактик, то компактное ядро с массами $(10^6 - 10^9) M_{\odot}$ обнаружено уже у несколько сотен объектов. Основными методами здесь также является измерение скоростей вращения звезд и газа в их центральных областях. В нескольких случаях в аккреционных дисках удалось зарегистрировать мазерные источники, чьи собственные частоты излучения известны с очень высокой точностью. Поэтому с помощью эффекта Доплера для этих галактик также стало возможным точно определить скорости вращения в аккреционном диске.

Еще раз подчеркнем, что речь здесь вновь идет лишь о косвенном обнаружении черных дыр. Как и в случае черной дыры в центре нашей Галактики, можно сказать лишь то, что масса центрального объекта заключена в достаточно малом объеме, в котором не может находиться компактное звездное скопление. Единственным указанием на существование сверхмассивных черных дыр (которые, к тому же, должны быстро вращаться) является красное смещение у линии железа 6.4 кэВ. Действительно, как мы видели, в случае невращающейся черной дыры гравитационное красное смещение на последней устойчивой орбите $1 - \alpha(r_0) \approx 0.06$ слишком мало, чтобы объяснить большие красные смещения, показанные на Рис. 19. Это можно сделать лишь в случае метрики Керра, когда последняя устойчивая орбита находится в более сильном гравитационном поле.

5.4 Электродинамическая модель "центральной машины" в активных галактических ядрах

5.4.1 Струйные выбросы

Наблюдения показывают, что существование струйных выбросов (джетов), т.е. узконаправленных потоков вещества, является общим свойством, характерным для многих астрономических источников (см. Рис. 28). Они наблюдаются как в активных ядрах удаленных галактик, так и в микрокварах и радиопульсарах, расположенных в нашей Галактике. Еще одним классом объектов, обладающим струйными выбросами, являются молодые звезды. Есть все основания полагать, что природа активности гамма-всплесков, находящихся на космологических расстояниях, также связана с направленными выбросами вещества.

Уже через несколько лет после открытия на рубеже шестидесятых годов радиогалактик – диффузионных областей, излучающих в радиодиапазоне (Рис. 28а) – стало ясно, что их активность связана со струйными выбросами, истекающими из квазаров (по-английски QSO – квазизвездный объект). Квазары представляют собой ядра активных галактик, столь далеких, что их оптическое изображение неотлично от обычных звезд. Именно джеты поставляют в эти области вещество и энергию, которые сам джет черпает из активного ядра галактики. Вещество джетов движется со скоростями v , близкими к скорости света c , так что лоренц-фактор джета как целого $\gamma = 1/(1 - v^2/c^2)^{1/2}$ достигает как минимум нескольких единиц. Характерные размеры струйных выбросов составляют сотни тысяч световых лет, что сравнимо с расстояниями между галактиками. Другим удивительным свойством джетов является их высокая коллимация – угол расхождения составляет всего несколько градусов. При этом потери энергии, связанные со струйными течениями, могут достигать 10^{45} эрг/с, что по порядку величины совпадает с полной светимостью галактики.

Сейчас уже мало кто сомневается в том, что источник активности квазаров связан с аккрецией (падением) вещества на сверхмассивные черные дыры (их массы, как мы видели, оценивают в миллиарды солнечных масс), поскольку лишь в этом случае удается объяснить высокую эффективность энерговыделения этих удивительных источников. Отметим, что время активной фазы квазаров не может превышать 10^7 лет. За это время черная дыра поглощает весь окружающий газ и звезды, после чего активность ядер галактик резко снижается. Именно такими "выключившимися" квазарами и являются центральные области ближайших к нам галактик, о которых шла речь в предыдущем разделе.

Напротив, недавно открытые микрокваразы – это объекты, расположенные в нашей Галактике (Рис. 28б). Формирование их струйных выбросов связывают со сверхкритиче-

ской аккрецией на черную дыру или нейтронную звезду солнечных масс, когда компактный объект не в состоянии поглотить все падающее на него вещество. В этом случае часть аккрецирующего вещества неизбежно должна быть отброшена в окружающее пространство. Характерный продольный размер таких струйных выбросов составляет обычно доли светового года. Общая же мощность излучения оценивается в 10^{37} эрг/с, что лишь в тысячу раз больше светимости Солнца. Несмотря на большое различие в размерах и энерговыделении, внешне струйные выбросы у квазаров и микроквазаров оказываются чрезвычайно похожими, что и объясняет возникновение самого термина микроквазар.

Наконец, молодые звезды были впервые обнаружены в начале пятидесятых годов, когда Дж. Хербиг и Дж. Аро открыли новый класс протяженных образований, существующих обычно парами и, как стало ясно значительно позже, также связанных струйными течениями с молодыми быстровращающимися звездами (Рис. 28б). У 60% источников видны оба струйных выброса, а в остальных случаях удаленный от нас выброс закрыт аккреционным диском. Протяженность джетов может достигать нескольких световых лет при характерной скорости движения вещества 600 км/с. Как мы видим, ситуация здесь вполне аналогична активным галактическим ядрам, когда вначале был обнаружен целый ряд разнородных источников (квазары, радиогалактики) и лишь затем стало ясно, что активность всех этих объектов имеет единую природу. Более того, подобие наблюдательных свойств наводит на мысль, что и физический механизм формирования направленных течений у этих объектов может быть одинаков. И это несмотря на то, что физические условия вблизи молодых звезд (их массы порядка 3 – 10 масс Солнца, а полная излучаемая мощность составляет $10^{31} - 10^{36}$ эрг/с) существенно отличаются от условий в центре квазаров. И одним из главных отличий является нерелятивистский характер истечения вещества из молодых звезд.

Если же говорить об источниках космологических гамма-всплесков, то здесь имеются лишь косвенные, хотя и достаточно надежные аргументы в пользу существования у них струйных выбросов, причем именно связанных с релятивистскими течениями, о которых пойдет речь в этой статье. Как известно, обнаружение оптического послесвечения (afterglow), позволившее по красному смещению определить расстояния до этих объектов, накладывает серьезные ограничения на их энергетику. Если считать, что излучение в гамма диапазоне происходит изотропно, то для расстояний порядка 10^9 световых лет, характерных для этих источников, приходится предположить, что их полная светимость достигает 10^{54} эрг/с. Однако в настоящее время процессы с таким огромным энерговыделением нам неизвестны. Если же считать, что излучение происходит в узкий конус углов θ порядка 1° , то излучаемая мощность может быть понижена до 10^{51} эрг/с, что по порядку величины близко к энергии, выделяющейся во время взрывов сверхновых. Но это означает, что характерный лоренц-фактор частиц, с которыми связано наблюдаемое

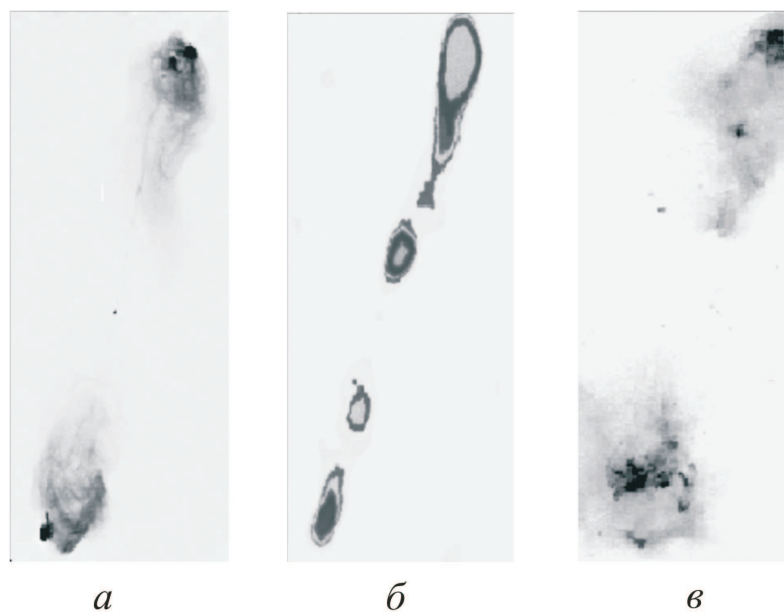


Рис. 28: Струйные выбросы, наблюдаемые в активных галактических ядрах (а), микрокварах (б) и молодых звездах (в)

излучение, должен составлять несколько сотен ($\gamma \sim 1/\theta \sim 100 - 300$).

Что же касается самой центральной машины, приводящей к образованию струйных выбросов, то здесь обычно обсуждается либо столкновение двух нейтронных звезд или нейтронной звезды и черной дыры, либо коллапс массивного ядра необычной сверхновой. Однако в большинстве моделей в итоге все равно образуется быстро вращающаяся черная дыра с массой порядка массы Солнца, окруженная аккреционным диском. Иными словами, модель вновь строится по той же схеме, что и для активных галактических ядер.

5.4.2 Источник активности – вращающаяся черная дыра

Таким образом, струйные выбросы существуют не только в релятивистских объектах, но и в молодых звездах, где скорости истекающего вещества не являются релятивистскими. Поэтому причина их образования должна быть достаточно универсальной. Как мы видели, во всех рассмотренных выше примерах таким общим свойством была аккреция вещества на компактный объект. При этом наличие аккреционных дисков однозначно указывало на то, что падающее вещество обладает большим угловым моментом (в случае молодых звезд, помимо диска, вращается и сама молодая звезда). Но хорошо известно, что существование углового момента препятствует падению на гравитирующий центр. Если потери энергии и углового момента отсутствуют, то вращение может продолжаться бесконечно долго. Именно такая ситуация имеет место в солнечной системе, когда пла-

неты в среднем находятся на постоянном расстоянии от Солнца. С другой стороны, было ясно, что единственным механизмом, способным объяснить высокую энергетику компактных объектов, является выделение гравитационной (т.е. попросту потенциальной) энергии при падении вещества на компактный объект. Поэтому для объяснения высокой активности компактных объектов (которая как раз и показывает, что темп аккреции вещества во всех случаях достаточно велик) необходимо было найти эффективный механизм отвода углового момента за пределы аккреционного диска. Как оказалось, именно эту роль и играют струйные выбросы. Для молодых звезд при этом более важно, что истечение вещества уменьшает также и угловой момент вращающейся молодой звезды, помогая ей сжаться до необходимых размеров.

Построение модели центральной машины распадается на несколько частей. Необходимо объяснить полные потери энергии в компактном объекте, механизм передачи углового момента от аккрецирующего вещества к струйным выбросам и саму коллимацию струйных выбросов. К сожалению, осветить здесь все эти вопросы не представляется возможным. Поэтому мы обсудим здесь лишь модель, в которой основное место отводится вращающейся черной дыре. Как мы видели, этот случай реализуется как в случае активных ядер, так и в микрокварах и источниках гамма-всплесков. При этом мы сосредоточимся лишь на механизме потерь энергии этих объектов, который связан с сильными электрическими токами, текущими в магнитосфере черной дыры.

Прежде всего, постараемся понять, почему для объяснения релятивистских струйных выбросов необходим столь необычный объект, каким является черная дыра. Как уже отмечалось, в перечисленных выше случаях скорости частиц в струйных выбросах оказываются близки к скорости света. В принципе, частицы могут быть ускорены до больших скоростей за счет газового давления или давления света. Но для этого светимость компактного объекта должна быть достаточно велика, что не наблюдается в действительности. При аккреции же релятивистские скорости могут появиться лишь в том случае, если гравитационный потенциал $\varphi_g = -GM/R$ сравнивается с квадратом скорости света c^2 . А это и означает, что радиус компактного объекта R должен быть близок к радиусу горизонта черной дыры r_g . Напомним, что для звезды солнечной массы гравитационный радиус составляет лишь 3 км, а для сверхмассивных черных дыр в центрах квазаров может достигать размеров солнечной системы.

Конечно же, для эффективного отвода энергии и углового момента необходимо, чтобы скорости частиц были направлены от черной дыры. Для этого и приходится привлекать сильные электромагнитные поля. Имеется также ряд косвенных аргументов, говорящих в пользу существования черной дыры у основания струйных выбросов. Как известно, черная дыра по определению не может быть источником вещества. Поэтому истечение может стать возможным лишь в том случае, если плазма сама будет рождаться в магнитосфере

черной дыры. Оказалось, что эффективный механизм рождения плазмы действительно существует. Дело в том, что благодаря сильному разогреву внутренних областей аккреционных дисков они должны генерировать большое количество жестких гамма-квантов, энергии которых превышают массу покоя электрона. Столкновение таких энергичных фотонов и приводит к рождению электронно-позитронных пар, заполняющих магнитосферу черной дыры. Таким образом, мы приходим к заключению, что струйные выбросы, связанные не с аккреционным диском, а с черной дырой, должны состоять не из обычной, а из электронно-позитронной плазмы. В пользу этой модели говорит, например, тот факт, что в релятивистских выбросах микроквазаров не наблюдаются эмиссионные линии, которые излучает обычное вещество. Ультрарелятивистский характер течения в источниках гамма-всплесков также накладывает ограничение на состав частиц в струйном выбросе, т.к. существование протонов с такой энергией противоречило бы полному энерговоделению гамма-всплеска. Как показывают оценки, вклад протонов должен составлять лишь 10^{-5} от полного числа частиц, так что и здесь речь может идти именно об электронно-позитронных джетах.

Далее, аккрецирующее вещество, обладающее собственным угловым моментом, будет увеличивать скорость вращения черной дыры. Поэтому не исключено, что черная дыра сможет раскрутиться до предельных значений $\Omega_{\text{Hr}_g} \sim c$. Как полагают, именно за счет аккреции были раскручены миллисекундные радиопульсары, так что этот процесс действительно может быть очень эффективным. В результате, в обсуждаемой модели источником активности будет не только гравитационная энергия падающего вещества, но и энергия вращения черной дыры. Тогда для случая предельного вращения запасы энергии будут сравнимы с ее энергией покоя: $\mathcal{E}_{\text{tot}} \sim Mc^2$. Этой энергии вполне достаточно, чтобы объяснить наблюдаемое энерговоделение.

Наконец, как мы видели, выделение энергии вращения не противоречит тому факту, что ни вещество, ни излучение не могут покинуть черную дыру. При определенных условиях энергия вещества и/или излучения, пересекающих горизонт событий, может быть отрицательной, что должно привести к уменьшению энергии черной дыры.

5.4.3 Униполярный индуктор

Чтобы понять, почему существование магнитного поля может приводить к потерям энергии вращающейся черной дыры, обратимся к известному примеру униполярного индуктора, представляющего собой вращающийся намагниченный шар, к которому двумя концами присоединен проводник (см. Рис. 29). При этом обязательно должно быть выполнено два условия.

1. Проводник не должен вращаться вместе с шаром (точнее, он не должен вращаться с той же угловой скоростью).

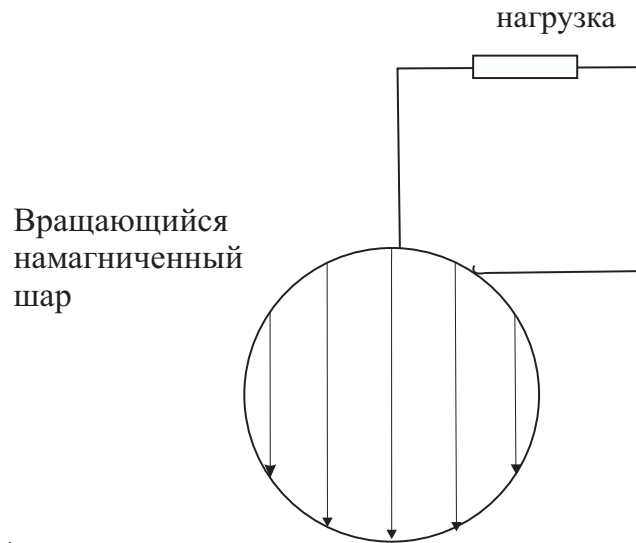


Рис. 29: Униполярный индуктор как модель магнитосферы компактного объекта. Вращающийся намагниченный шар играет роль нейтронной звезды или черной дыры, тогда как невращающийся проводник моделирует магнитные силовые линии

2. Два конца проводника должны подсоединяться к шару на разных широтах (наиболее выгодный способ – полюс и экватор).

В результате, между концами проводника появляется разность потенциалов U , которая и является источником электрического тока. Ее величину легко оценить, зная, что проводящая среда, движущаяся со скоростью \mathbf{v} в магнитном поле \mathbf{B} , создает электрическое поле $\mathbf{E} = -\mathbf{v} \times \mathbf{B}/c$. Поэтому по порядку величины разность потенциалов $U \sim ER$ на концах проводника будет равна

$$U \sim \frac{\Omega R}{c} RB, \quad (222)$$

где Ω – угловая скорость вращения шара, R – его радиус, а B – магнитное поле на его поверхности. При этом электрический ток I , текущий вдоль провода, будет отводить энергию от батареи (т.е. от вращающегося шара) к внешней нагрузке, где и происходит выделение тепла. В результате, кинетическая энергия вращения за счет действия электромагнитных сил будет уноситься на бесконечность. Механизм же замедления вращения будет связан с поверхностными токами, которые неизбежно должны появиться для замыкания токов, текущих за пределами шара. Сила Ампера $\mathbf{F}_A = \mathbf{J} \times \mathbf{B}/c$, действующая со стороны тока на поверхность шара, и приводит к его торможению.

Подчеркнем, что здесь появление электродвижущей силы не связано с эффектом Фарадея. Как видно из Рис. 29, магнитный поток через внешний контур не изменяется со временем. Электрический ток становится возможным благодаря появлению сторонних сил, действующих на свободные электроны со стороны вещества вращающегося шара.

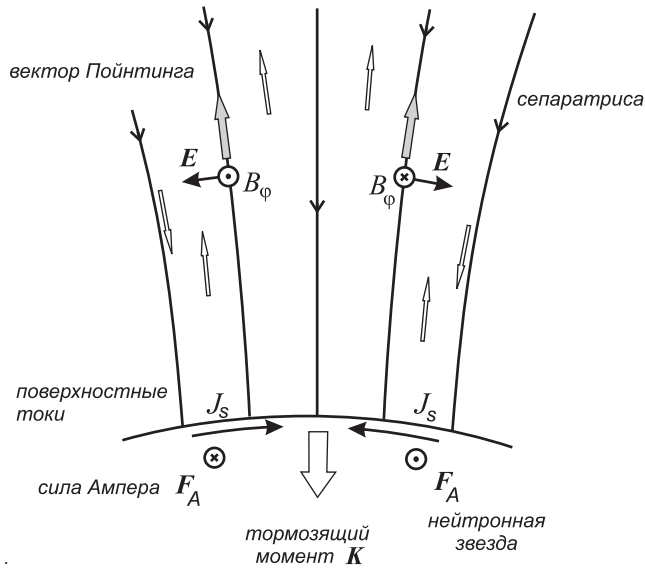


Рис. 30: Структура электрических токов вблизи поверхности нейтронной звезды (контурные стрелки). Токи, текущие в магнитосфере пульсара, должны замыкаться на его поверхности. Сила Ампера, F_A связанная с поверхностными токами J_s , приводит к замедлению вращения нейтронной звезды. При этом токи текут лишь в ограниченной области магнитосферы, которая не может вращаться вместе со звездой

Действительно, появление электрического поля $\mathbf{E} = -\mathbf{v} \times \mathbf{B}/c$ в объеме шара требует определенного перераспределения в последнем электрических зарядов. Распределение зарядов нарушается, когда электроны начинают течь вдоль проводника от одной части шара к другой. В результате в шаре возникают дополнительные силы, которые и перемещают электроны по поверхности шара против действия электрического поля, замыкая ток, текущий по проводу.

В настоящее время точно установлено, что именно таким образом происходит выделение энергии у радиопульсаров – вращающихся намагниченных нейтронных звезд (см. Рис. 30). При этом механизм замедления вращения связан с поверхностными токами, которые неизбежно должны появиться для замыкания токов, текущих вне звезды. Сила Ампера $\mathbf{F}_A = \mathbf{J} \times \mathbf{B}/c$, действующая со стороны тока на поверхность звезды, и приводит к его торможению. С другой стороны, от поверхности звезды энергия будет уноситься за счет потока вектора Пойнтинга $\mathbf{S} = (c/4\pi)\mathbf{E} \times \mathbf{B}$. Действительно, вращение звезды приводит к возникновению электрического поля \mathbf{E} , тогда как продольный ток генерирует тороидальное магнитное поле \mathbf{B}_ϕ . Как показано на Рис. 30, эти поля и формируют поток электромагнитного поля. Подчеркнем, что эта энергия переносится на нулевой частоте и поэтому уносящее эту энергию электромагнитное поле не является электромагнитной волной в обычном понимании этого слова.

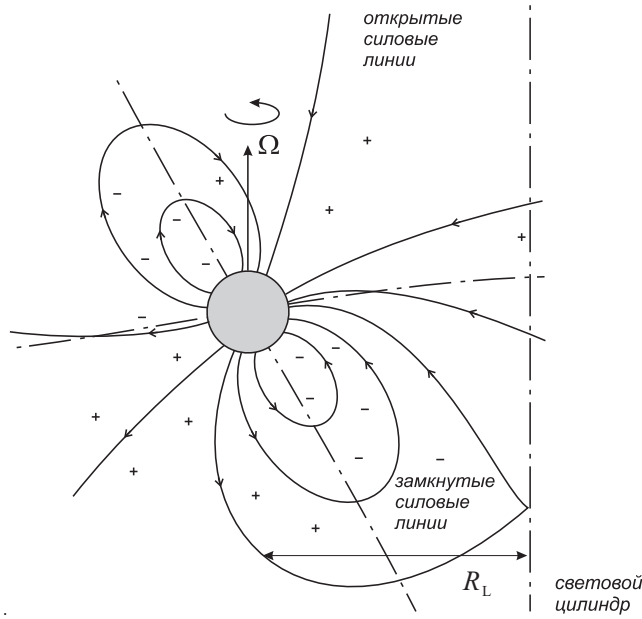


Рис. 31: Открытые и замкнутые силовые линии в магнитосфере радиопульсара

Что же касается величины электрического тока I , то она определяется из следующих соображений. Плотность электрического заряда ρ_e , необходимая для генерации индукционного электрического поля $E \sim (v/c)B$, может быть оценена как $\rho_e \sim E/L$ (см. уравнение Максвелла (128)), где L – характерный размер области с электрическим полем E . Для вращающегося тела $v \sim \Omega L$. Поэтому по порядку величины $\rho_e \sim \Omega B/c$. Точное же выражение для плотности заряда имеет вид

$$\rho_e = -\frac{(\boldsymbol{\Omega} \cdot \mathbf{B})}{2\pi c}. \quad (223)$$

Умножая теперь плотность заряда ρ_e на скорость света c , а также на площадь πR_0^2 , где R_0 – поперечный размер области, в котором течет ток I , получаем

$$I \sim \pi R_0^2 \rho_e c. \quad (224)$$

Полные же потери энергии W могут быть оценены как $W = IU$.

Для радиопульсаров радиус R_0 определяется размером полярной шапки, из которой магнитные силовые линии дипольного магнитного поля уходят за пределы т.н. светового цилиндра

$$R_L = \frac{c}{\Omega}. \quad (225)$$

Дело в том, что плазма, заполняющая магнитосферу нейтронной звезды, стремится заэкранировать продольное электрическое поле. При этом она начинает вращаться твердотельно вместе со звездой. Такая коротация хорошо известна в магнитосфере Земли и

больших планет. Ясно, однако, что коротация становится невозможной за пределами светового цилиндра, т.к. здесь скорость вращения превышала бы скорость света. Поэтому, как показано на Рис. 31, в магнитосфере выделяются две группы магнитных силовых линий. Одна группа, выходящая из пределов полярной шапки, пересекает световой цилиндр и уходит на бесконечность, а другая, находящаяся вдали от магнитной оси, замыкается в пределах светового цилиндра. При этом плазма, находящаяся на замкнутых магнитных силовых линиях, будет вращаться вместе с нейтронной звездой, и, следовательно, не сможет участвовать в работе униполярного индуктора. Продольные токи смогут течь лишь в области открытых силовых линий.

Воспользовавшись теперь оценкой $R_0 \sim (\Omega R/c)^{1/2} R$, справедливой для дипольного магнитного поля (это соотношение можно получить, например, сравнивая магнитные потоки, проходящие через полярную шапку и поверхность светового цилиндра), получаем окончательно хорошо известное выражение для потерь энергии радиопульсаров

$$W \sim (\Omega R/c)^4 B_0^2 R^2 c. \quad (226)$$

Для стандартных параметров нейтронных звезд (магнитное поле $B \sim 10^{12}$ Гаусс, радиус $R \sim 10$ км, $\Omega R/c \sim 10^{-2} - 10^{-4}$) эта формула полностью соответствует наблюдаемым потерям энергии радиопульсаров $W = I_r \Omega \dot{\Omega} \sim 10^{31} - 10^{38}$ эрг/с.

Итак, магнитное поле действительно может играть ключевую роль в процессе выделения энергии из быстро вращающихся компактных объектов. Понятно, что одновременно компактный объект будет терять и момент количества движения. Мы уже подчеркивали универсальность этого механизма. Следовательно, магнитное торможение может иметь место у самых различных астрофизических источников. В частности, подобным же образом должны терять энергию и угловой момент нерелятивистские объекты, т.е. обычные звезды.

И действительно, надежно установлено, что "магнитные" звезды, у которых магнитное поле на поверхности достигает нескольких килогауссов, вращаются гораздо медленнее, чем звезды, не обладающие такими сильными магнитными полями. Естественно связать это различие именно со значительно более эффективным замедлением звезды в течение ее жизни. Не исключено, что именно магнитное торможение объясняет и хорошо известный факт, что момент количества движения в солнечной системе содержится не в Солнце, а в больших планетах.

5.4.4 Черная дыра в магнитном поле

Простота и естественность, с которой удалось объяснить активность радиопульсаров, привела к тому, что этот процесс стал рассматриваться как универсальный механизм,

применимый и для случая вращающейся черной дыры. Для этого, как мы видели, необходимо существование

1. регулярного магнитного поля, играющего роль проводника,
2. вращения центрального тела, создающего разность потенциалов U , и
3. электрического тока I , уносящего энергию от вращающегося тела и уменьшающего скорость его вращения.

А основной задачей, которую нужно было решить при определении природы активности компактного объекта, стала задача нахождения структуры электрических токов, текущих в его магнитосфере.

Здесь, следует сделать два важных замечания. Во-первых, необходимо объяснить, почему магнитное поле может играть роль проводника (этот вопрос относится и к магнитосфере радиопульсаров). Действительно, в земных условиях внешнее магнитное поле, например, собственное магнитное поле Земли, никак не влияет на направление распространения токов. Однако если магнитное поле достаточно велико а плотность среды достаточно мала (так что длина свободного пробега частиц становится больше радиуса окружности $r_L = mc\nu/eB$, по которой заряженная частица движется в магнитном поле), и при этом радиус r_L будет много меньше размера системы, то в этом случае проводимость среды поперек магнитного поля должна быть существенно подавлена. В результате, с хорошей точностью можно считать, что заряженная частица, а, значит, и электрический ток, будет двигаться лишь вдоль внешнего магнитного поля.

Во-вторых, согласно т.н. "теореме об отсутствии волос" черная дыра не может иметь своего собственного магнитного поля. Магнитное поле любого источника, например, контура с током, приближающегося к поверхности черной дыры, будет искажаться ее гравитационным полем, в результате чего удаленный наблюдатель увидит, что магнитное поле в его окрестности будет быстро уменьшаться с течением времени. В итоге, в момент прохождения источником горизонта событий магнитное поле вне черной дыры окажется равным нулю. Единственным собственным полем, которым может обладать черная дыра, будет электрическое поле точечного заряда, расположенного в ее центре.

Чтобы понять причину этого явления, удобнее рассмотреть не магнитный, а электрический диполь, т.е. два близко расположенных разноименных электрических заряда. На Рис. 32 показано, как искажается электрическое поле одного заряда q , находящегося вблизи поверхности черной дыры. Несмотря на то, что заряд q находится вблизи горизонта, удаленному наблюдателю будет казаться, что его положение совпадает с центром черной дыры. А это значит, что при приближении электрического диполя к горизонту событий изображения зарядов по мнению удаленного наблюдателя будут приближаться к центру черной дыры, так что величина дипольного момента $\mathbf{d} = q\mathbf{l}$ за счет уменьшения

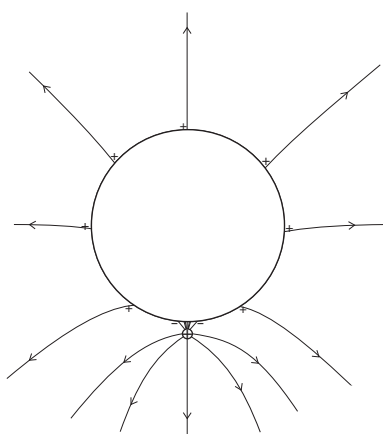


Рис. 32: Силовые линии электрического поля заряда q , находящегося вблизи горизонта черной дыры. Перемещение изображения заряда в центр черной дыры связан с кривизной пространства-времени, но он может быть формально объяснен и появлением поверхностных зарядов. Однако в действительности никаких реальных зарядов на поверхности горизонта не существует, а необходимость их введения связана с тем, что удаленный наблюдатель не может получать информацию о событиях, происходящих внутри горизонта событий. Поэтому мы вынуждены считать, что силовые линии, проходящие через горизонт, оканчиваются или начинаются на поверхностных зарядах

расстояния l между зарядами будет стремиться к нулю. Следовательно, будет стремиться к нулю и электрическое поле диполя, регистрируемое вдали от черной дыры. Точно также, будет уменьшаться и магнитное поле магнитного диполя по мере его приближения к горизонту событий.

Таким образом, собственное магнитное поле черной дыры должно быть равно нулю. Однако "теорема об отсутствии волос" не запрещает черной дыре быть погруженной во внешнее магнитное поле. Такое магнитное поле вполне может быть создано в аккреционном диске, который способен играть роль динамо-машины. Поэтому возможность существования сильного магнитного поля в рассматриваемых системах не вызывает сомнения.

5.4.5 Процесс Блэндфорда-Знайека

Мы видим, что идея униполярного индуктора может быть напрямую применена и к магнитосфере черной дыры. В частности, потери энергии можно оценивать также, как это было сделано для магнитосферы пульсаров. Отличие состоит лишь в том, что радиус полярной шапки R_0 следует заменить на гравитационный радиус r_g , поскольку, в отличие от радиопульсаров, вращение плазмы вблизи поверхности черной дыры не обязательно происходит с той же угловой скоростью, что и угловая скорость черной дыры. Следова-

тельно, роль униполярного индуктора может играть вся поверхность горизонта событий. В результате, оценка для потерь энергии по формуле $W = IU$ немедленно приводит к выражению

$$W_{\text{BZ}} \sim (\Omega r_g/c)^2 B^2 r_g^2 c, \quad (227)$$

полученному в 1977 году Р. Блэндфордом и Р. Знайеком при анализе строения магнитосферы вращающейся черной дыры. Для параметров, характерных для активных ядер (магнитное поле $B \sim 10^4$ Гаусс, радиус черной дыры $r_g \sim 10^{14}$ см, $\Omega r_g/c \sim 1$), эта формула хорошо согласуется с наблюдаемыми потерями энергии $W \sim 10^{45}$ эрг/с. Именно Блэндфорд и Знайек первыми показали, что вращающаяся черная дыра, погруженная во внешнее магнитное поле, может работать как униполярный индуктор, и поэтому этот процесс носит их имя.

Вместе с тем, несмотря на повсеместное использование механизма Блэндфорда-Знайека для объяснения природы активности разных астрофизических объектов, сама возможность электромагнитного выделения энергии вращающейся черной дырой уже много лет подвергается сомнению. При этом вопрос оказался столь серьезным, что в некоторых работах был даже сделан вывод о том, что черная дыра вообще не может работать как униполярный индуктор. Дело в том, что в 1982 году К. Торном и Д. Макдональдом был сформулирован подход, который полностью закреплял аналогию между магнитосферой нейтронной звезды и черной дыры. Для этого было введено понятие проводящей мембраны, расположенной у горизонта черной дыры, свойства которой были полностью эквивалентны свойствам поверхности нейтронной звезды (см. Рис. 33). В частности, по поверхности мембраны могли течь электрические токи \mathbf{J}_H , замыкающие токи, текущие в магнитосфере. При этом было показано, что тормозящее действие силы Ампера, связанной с поверхностными токами, действительно приводит к выражению (227) для потерь энергии вращающейся черной дыры.

Казалось бы, такое совпадение можно было рассматривать как прямое подтверждение существования процесса Блэндфорда-Знайека. Однако вскоре стало ясно, что многие элементы теории, связанной с поверхностными токами, не выдерживают критики. Например, как мы видели, поверхность черной дыры по определению причинно не связана с внешним пространством, так что информация о поверхностном токе (а, значит, и о величине потерь энергии) не может быть передана удаленному наблюдателю. Кроме того, черная дыра не имеет твердой поверхности. Поэтому совершенно непонятно, какой смысл имеют поверхностные токи, и на что действует сила Ампера, которая должна тормозить черную дыру.

Лишь в последнее время стало ясно, что критика касалась лишь мембранной трактовки, и не затрагивала самих основ процесса Блэндфорда-Знайека. Прежде всего, отметим, что в случае черной дыры для замыкания тока совсем не нужно, чтобы за-

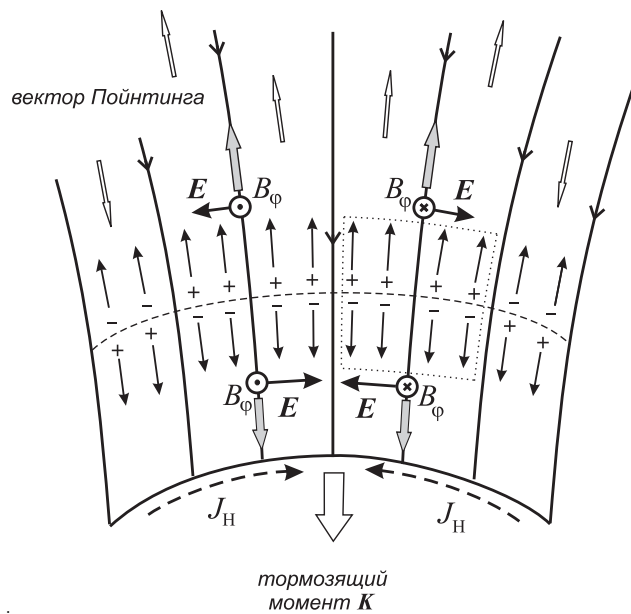


Рис. 33: Структура электрических токов вблизи поверхности черной дыры (контурные стрелки). Источником тока является область рождения электронно-позитронных пар (штриховая линия). Замыкание тока связано с тем, что в разных частях горизонта событий его пересекают частицы разного знака заряда. Показан также фиктивный ток \mathbf{J}_H , текущий по поверхности мембраны. Выделяемая энергия уносится от звезды электромагнитным полем (вектором Пойнтинга \mathbf{S}). Падающая на черную дыру электромагнитная энергия отрицательна. Направления электромагнитных полей показаны для наблюдателей, вращающихся с угловой скоростью ω .

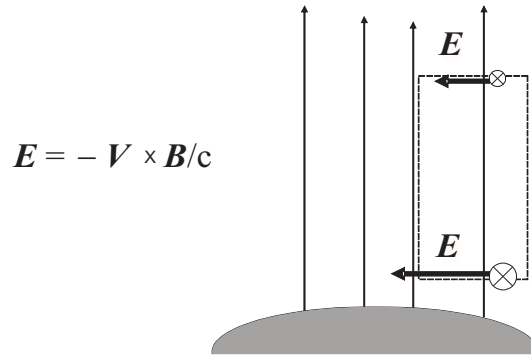


Рис. 34: Появление электродвижущей силы в контуре, покоящимся относительно вращающейся черной дыры. Направления электрического поля показаны для покоящихся наблюдателей.

ряженные частицы покидали черную дыру (см. Рис. 33). Для этого достаточно, чтобы суммарный заряд, пересекающий горизонт событий, был равен нулю. Более того, для замыкания тока заряды, фактически, не должны пересекать и магнитные силовые линии, как это имело место на поверхности нейтронной звезды. Как уже отмечалось, удаленному наблюдателю будет казаться, что два разноименных заряда, приблизившихся к горизонту, будут находиться в центре черной дыры, где, с его точки зрения, и произойдет замыкание тока. Этот пример показывает, что поверхностные токи вообще не должны играть никакой роли при торможении черной дыры.

Но тогда в электрическом контуре, содержащем черную дыру, следует найти другой источник электродвижущей силы, т.к. теперь черная дыра уже не может быть источником сторонних сил, и, следовательно, играть роль батареи постоянного тока. Оказалось, что появление ЭДС связано с эффектом Лензе-Тирринга (эффектом увлечения систем отсчета), т.е. с появлением гравитомагнитных сил. Попросту говоря, можно считать, что тело, находящееся в покое относительно вращающейся черной дыры, на самом деле движется относительно выделенной системы отсчета, вращающейся с угловой скоростью Лензе-Тирринга ω . Поэтому, как и в случае любых тел, движущихся в магнитном поле, относительное движение приводит к появлению электрического поля

$$\mathbf{E} = -\mathbf{v} \times \mathbf{B}/c, \quad (228)$$

где \mathbf{v} – линейная скорость тела относительно выделенной системы отсчета. При этом важно, что угловая скорость вращения будет различной на разных расстояниях от черной дыры. Поэтому, как показано на Рис. 34, циркуляция электрического поля в контуре будет отлична от нуля. В результате, даже в том случае, если электрический контур покоится относительно черной дыры, так что магнитный поток через него остается постоянным, ”движение пространства” приведет к появлению электродвижущей силы в

этом контуре. При этом, как мы видим, "электрическая батарея" будет находиться над поверхностью горизонта черной дыры.

Наконец, как уже подчеркивалось, для того, чтобы в магнитосфере черной дыры могли течь электрические токи (и, значит, она могла работать как униполярный индуктор), нужно, чтобы существовал механизм рождения плазмы. При этом, как и в эффекте Пенроуза, одна частица должна упасть на черную дыру, а вторая уйти на бесконечность. Таким образом, процесс Блэндфорда-Знайека является просто электромагнитной реализацией процесса Пенроуза. Только он касается не самих заряженных частиц, а создаваемого ими электромагнитного поля. То есть замедление вращения черной дыры связано не с электрическими токами, текущими по поверхности горизонта, а с тем, что плотность энергии электромагнитного поля, пересекающего поверхность черной дыры, оказывается отрицательной. Эта интерпретация представляется сейчас наиболее правдоподобной, так что к ней склоняется большинство исследователей, занимающихся этим кругом вопросов.

Таким образом, можно надеяться, что мы сейчас правильно понимаем механизм выделения энергии релятивистских компактных объектов. С другой стороны, мы видели, как следует быть осторожным при попытке перенести наше понимание электродинамических процессов в плоском пространстве на случай магнитосферы черной дыры. Несмотря на то, что вращающаяся черная дыра, погруженная во внешнее магнитное поле, действительно может работать как униполярный индуктор, сам механизм замедления вращения оказывается не связанным с электрическими токами, текущими по поверхности черной дыры.

6 Заключение

Таким образом, общая теория относительности, построенная как чисто дедуктивная теория, позволила сделать многочисленные предсказания, многие из которых были впоследствии успешно подтверждены. Более того, можно с уверенностью сказать, что сейчас не существует никаких указаний на какие-либо расхождения между предсказаниями общей теории относительности и наблюдениями. Вместе с тем, не надо забывать, что пока удалось проверить предсказания ОТО лишь в слабом гравитационном поле. Поэтому говорить об окончательной победе еще преждевременно. Это прекрасно понимал и Эйнштейн, который всегда подчеркивал, что эксперимент никогда не говорит "да", а лишь "нет" или "может быть". Однако сама возможность сформулировать четкие предсказания (и, значит, впоследствии опровергнуть теорию, если она будет противоречить наблюдаемым данным) является важнейшим свойством ОТО, характеризующим ее как истинно научную теорию. Этим теория относительности выгодно отличается от многих современных теорий фундаментальных взаимодействий, которые, несмотря на внутреннюю стройность и математическое совершенство, до сих пор не сформулировали ни одного наблюдательного теста, способного проверить их справедливость.

Что же дальше? Что ожидает нас в ближайшем будущем? Прежде всего, анализ прихода импульсов от двойных пульсаров позволит сравнить предсказания общей теории относительности в следующем порядке по величине r_g/r , т.е. до величин $r_g^2/r^2 \sim c^{-4}$. Это заведомо должно случиться уже через несколько лет, поскольку процедура обработки моментов прихода радиоимпульсов четко фиксирует время, которое необходимо для получения необходимой точности. Далее, есть реальная надежда, что в ближайшее время будут открыты гравитационные волны. Если это произойдет, то теория относительности будет проверена вплоть до величин $\sim c^{-5}$. Однако окончательная проверка станет возможной лишь при обнаружении эффектов вблизи поверхности черных дыр, таких как отсутствие твердой поверхности и бесконечно сильного замедления времени вблизи горизонта. Для этого необходимо увеличить угловое разрешение приемной аппаратуры до 10^{-6} угловых секунд. В этом случае станет возможным наблюдать процессы вблизи сверхмассивных черных дыр в ядрах галактик (такое угловое разрешение позволит разрешать и изображения микролинзирования на черных дырах). В настоящее время такие эксперименты также планируются, так что не исключено, что в ближайшие пять-десять лет мы сможем непосредственно наблюдать процессы, происходящие в центральной машине активных галактических ядер¹³. Тем самым, будет сделан еще один шаг к дьящемуся уже без малого сто лет спору о том, справедлива ли теория гравитации, сформулированная Альбертом Эйнштейном в ноябре 1915 года.

¹³Что же касается возможности наблюдать внутренние области аккреционных дисков у черных дыр солнечных масс, то для этого потребовалось бы угловое разрешение 10^{-10} угловых секунд.

7 Приложения

Как показывает опыт, семинарские занятия (т.е. попросту решение задач) являются совершенно необходимым элементом преподавания. Поэтому мы приведем здесь вывод некоторых ключевых соотношений, которые были использованы нами выше. При этом иногда мы не будем интересоваться численными коэффициентами, а проследим лишь зависимость от физических параметров, характеризующих задачу. Как мы увидим, большинство таких соотношений могут быть получены "на пальцах", если правильно понимать физическую картину явлений. С другой стороны, в некоторых случаях мы сможем получить и точные выражения.

7.1 Закон преобразования координат "из первых принципов"

Как мы видели, обычно общие принципы (симметрия, лоренц-инвариантность) могут помочь ограничить теорию, но не определяют ее до конца. Однако в некоторых случаях сам математический вид теории может быть получен "из первых принципов". Мы покажем это на примере формулы (9) для преобразования компонент вектора при поворотах системы координат, которая в двумерном виде записывается как

$$\begin{aligned}A_{x'} &= A_x \cos \varphi + A_y \sin \varphi, \\A_{y'} &= -A_x \sin \varphi + A_y \cos \varphi.\end{aligned}\tag{229}$$

Предположим теперь, что вид преобразования нам неизвестен. Поэтому запишем его в самом общем виде

$$A_{x'} = f_1(\varphi)A_x + f_2(\varphi)A_y,\tag{230}$$

$$A_{y'} = f_3(\varphi)A_x + f_4(\varphi)A_y,\tag{231}$$

которое может иметь место при линейном преобразовании координат. Покажем теперь, что все четыре неизвестных функции могут быть найдены, если мы потребуем, что

1. при поворотах остается инвариантной длина любого вектора \mathbf{A} (т.е. что всегда имеет место соотношение $A_x^2 + A_y^2 = A_{x'}^2 + A_{y'}^2$),
2. два последовательных поворота вокруг оси z на углы φ и φ' есть также поворот на некоторый угол φ_s (групповое свойство).

Рассмотрим прежде всего первое условие. Возводя в квадрат выражения (229), получаем

$$\begin{aligned}A_{x'}^2 + A_{y'}^2 &= [f_1^2(\varphi) + f_3^2(\varphi)]A_x^2 + 2[f_1(\varphi)f_2(\varphi) + f_3(\varphi)f_4(\varphi)]A_xA_y + [f_2^2(\varphi) + f_4^2(\varphi)]A_y^2 \\ &= A_x^2 + A_y^2.\end{aligned}\tag{232}$$

При этом важно, что соотношение (232) должно выполняться для любых векторов \mathbf{A} . А это значит, что коэффициенты $f_1(\varphi)$, $f_2(\varphi)$, $f_3(\varphi)$ и $f_4(\varphi)$ должны удовлетворять условиям

$$\begin{aligned} f_1^2(\varphi) + f_3^2(\varphi) &= 1, \\ f_1(\varphi)f_2(\varphi) + f_3(\varphi)f_4(\varphi) &= 0, \end{aligned} \quad (233)$$

$$f_2^2(\varphi) + f_4^2(\varphi) = 1, \quad (234)$$

для любых углов поворота φ . Можно показать, что это становится возможным лишь при условии

$$\begin{aligned} f_2(\varphi) &= \sqrt{1 - f_1^2(\varphi)}, \\ f_3(\varphi) &= -\sqrt{1 - f_1^2(\varphi)}, \end{aligned} \quad (235)$$

$$f_4(\varphi) = f_1(\varphi). \quad (236)$$

Таким образом, преобразования координат векторов при поворотах координатной сетки вокруг оси z полностью определяются лишь одной функцией $f_1(\varphi)$, которую ниже мы будем обозначать просто $f(\varphi)$.

Используем теперь второе (групповое) предположение. Для этого совершим еще один поворот на угол φ' . Тогда должны быть выполнены условия

$$A_{x''} = f_1(\varphi')A_{x'} + f_2(\varphi')A_{y'}, \quad (237)$$

$$A_{y''} = f_3(\varphi')A_{x'} + f_4(\varphi')A_{y'}. \quad (238)$$

В результате, получаем, например, для $A_{x''}$

$$\begin{aligned} A_{x''} &= [f(\varphi)f(\varphi') - \sqrt{1 - f^2(\varphi)}\sqrt{1 - f^2(\varphi')}]A_x \\ &\quad + [\sqrt{1 - f^2(\varphi)}f(\varphi') + f(\varphi)\sqrt{1 - f^2(\varphi')}]A_y \end{aligned} \quad (239)$$

(уравнение для $A_{y''}$ дополнительной информации не дает). Но по второму условию два последовательных поворота также должны быть поворотом на некоторый угол φ_s . Поэтому для любых углов φ и φ' должны быть выполнены соотношения

$$f(\varphi_s) = f(\varphi)f(\varphi') - \sqrt{1 - f^2(\varphi)}\sqrt{1 - f^2(\varphi')}, \quad (240)$$

$$\sqrt{1 - f^2(\varphi_s)} = \sqrt{1 - f^2(\varphi)}f(\varphi') + f(\varphi)\sqrt{1 - f^2(\varphi')}. \quad (241)$$

Единственное решение этой системы имеет вид

$$f(\varphi) = \cos \varphi, \quad (242)$$

$$\varphi_s = \varphi + \varphi', \quad (243)$$

в результате чего соотношения (240) – (241) превращаются в хорошо известные тригонометрические тождества

$$\cos(\varphi + \varphi') = \cos(\varphi) \cos(\varphi') - \sin(\varphi) \sin(\varphi'), \quad (244)$$

$$\sin(\varphi + \varphi') = \sin(\varphi) \cos(\varphi') + \cos(\varphi) \sin(\varphi'). \quad (245)$$

Таким образом, для поворотов систем координат сама математическая формула преобразования координат векторов может быть получена из общих свойств, которым должны удовлетворять эти преобразования.

7.2 Понятие о ковариантной производной

Как уже говорилось, в интегральной форме законы физики могут быть сформулированы независимо от выбора координатной сетки. Например, отсутствие источников у магнитного поля (магнитные силовые линии либо замкнуты, либо уходят на бесконечность) можно сформулировать как условие того, что поток вектора магнитного поля через любую замкнутую поверхность равен нулю; здесь уточнения *замкнутую* и *любую* являются совершенно необходимыми. С другой стороны, понятно, что само условие отсутствия источников не зависит от выбора координат. Например, если мы нарисуем замкнутые, нигде не пересекающиеся линии на плоской поверхности, а затем станем эту поверхность произвольным образом деформировать, то и на искривленной поверхности линии нигде пересекаться не будут.

Напомним, что в декартовых координатах уравнение Максвелла $\operatorname{div} \mathbf{B} = 0$, описывающее в дифференциальной форме условие отсутствия источников у магнитного поля \mathbf{B} , записывается как

$$\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} = 0. \quad (246)$$

Возникает вопрос, можно ли записать это дифференциальное уравнение таким образом, чтобы результат не зависел ни от кривизны поверхности, ни от выбора системы координат. Оказывается, можно. В частности, для произвольных ортогональных координат x_1 , x_2 и x_3 оно имеет вид

$$\frac{1}{\sqrt{g}} \frac{\partial}{\partial x_1} \left(\frac{\sqrt{g}}{\sqrt{g_{11}}} B_1 \right) + \frac{1}{\sqrt{g}} \frac{\partial}{\partial x_2} \left(\frac{\sqrt{g}}{\sqrt{g_{22}}} B_2 \right) + \frac{1}{\sqrt{g}} \frac{\partial}{\partial x_3} \left(\frac{\sqrt{g}}{\sqrt{g_{33}}} B_3 \right) = 0, \quad (247)$$

где B_1 , B_2 и B_3 – проекции вектора \mathbf{B} на координатные оси, а $g = g_{11}g_{22}g_{33}$ – детерминант матрицы $\langle g \rangle$ ¹⁴. Уравнение (247) и есть пример ковариантной записи дифференциального оператора.

¹⁴Ковариантная производная может быть записана и для произвольных (не ортогональных) координат, но выражение для нее слишком сложно, и мы его здесь не приводим.

Соответственно, компоненты метрического тензора $\langle g \rangle$ входят и в инвариантные определения следа и квадрата произвольного тензора. Так, например, ковариантное определение следа диагональной матрицы 3×3 выглядит как

$$\text{Sp} \langle A \rangle = \frac{A_{11}}{g_{11}} + \frac{A_{22}}{g_{22}} + \frac{A_{33}}{g_{33}}, \quad (248)$$

и, аналогично, для тензора 4×4 . Поэтому, независимо от выбора системы ортогональных координат, получаем

$$\text{Sp} \langle g \rangle = d, \quad (249)$$

где d – размерность пространства. Соответственно, квадрат симметричной матрицы 3×3 запишется как

$$\langle A \rangle^2 = \frac{A_{11}^2}{g_{11}^2} + 2\frac{A_{12}^2}{g_{11}g_{22}} + \frac{A_{22}^2}{g_{22}^2} + 2\frac{A_{13}^2}{g_{11}g_{33}} + 2\frac{A_{23}^2}{g_{22}g_{33}} + \frac{A_{33}^2}{g_{33}^2} \quad (250)$$

Как мы видим, для диагонального метрического тензора $\langle g \rangle$ квадрат и след совпадают друг с другом.

7.3 Плотность энергии электромагнитного и слабого гравитационного поля

Как мы видели, плотность энергии электромагнитного поля записывается в виде

$$\varepsilon_{\text{em}} = \frac{E^2 + B^2}{8\pi}. \quad (251)$$

Покажем, как для электрического поля оно может быть получено из элементарных соображений. Для этого рассмотрим конденсатор, плотность заряда на обкладках которого есть $\pm\sigma_e$ (= заряд/площадь). Используя теперь теорему Гаусса, легко получить, что электрическое поле между обкладками будет равно

$$E = 4\pi\sigma_e, \quad (252)$$

а вне обкладок равно нулю.

Переместим теперь одну из пластин на расстояние δh . В результате, объем, занятый электрическим полем, изменится на величину $\delta V = S\delta h$, так что энергия поля изменится на величину $\delta\mathcal{E} = \varepsilon_{\text{em}}S\delta h$, где ε_{em} – плотность энергии электрического поля. Понятно, что эта энергия должна быть равна работе, которую нужно затратить, чтобы переместить пластину в новое положение. Эта работа равна $\delta A = F\delta h$, где сила $F = QE$ может быть записана в виде

$$F = \frac{1}{2}\sigma_e SE. \quad (253)$$

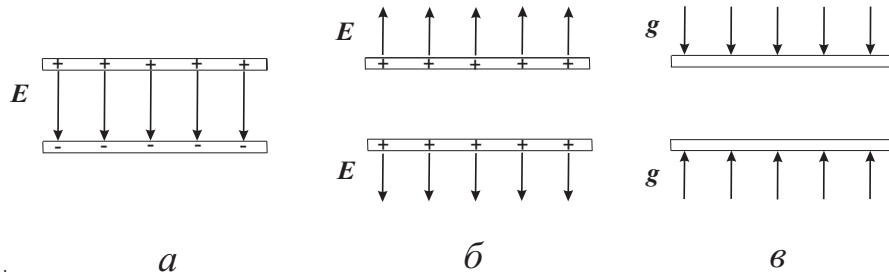


Рис. 35: Электрическое поле в обычном конденсаторе (а), поле двух одинаково заряженных пластин (б), а также гравитационное поле двух массивных пластин (в). Перемещение пластин приводит к изменению объема, занятого полем, причем изменение энергии поля равно совершаемой работе

Здесь дополнительный фактор $1/2$ связан с тем, что пластины никогда не бывают идеально тонкими, так что в объеме пластины электрическое поле спадает от величины $4\pi\sigma_e$ до нуля. Сравнивая теперь выражения для δA и $\delta\mathcal{E}$, получаем окончательно

$$\varepsilon_{\text{em}} = \frac{E^2}{8\pi}. \quad (254)$$

Для нас важно, что мы, фактически, нигде не использовали то обстоятельство, что обкладки конденсатора имели заряды противоположного знака. Те же самые выкладки можно было бы провести для двух одинаково заряженных пластин (см. Рис. 35б). В этом случае электрическое поле будет иметь место лишь за пределами обкладок, однако ее величина вновь будет определяться выражением $E = 4\pi\sigma_e$. Конечно, в этом случае энергия поля будет сосредоточена теперь лишь за пределами обкладок. Однако изменение энергии $\delta\mathcal{E} = \varepsilon_{\text{em}}S\delta h$ остается без изменений. В результате, мы вновь возвращаемся к выражению (254).

Рассмотрим теперь вопрос о плотности энергии гравитационного поля. Как было показано, в рамках общей теории относительности такая постановка задачи некорректна, поскольку гравитационное поле может быть устранено переходом в соответствующую систему отсчета. Однако для слабых гравитационных полей, удовлетворяющих "электродинамическим" уравнениям (132)–(135), понятие плотности энергии оказывается физически корректным. Поэтому для определения плотности энергии гравитационного поля мы можем рассмотреть поле двух массивных пластин (см. Рис. 35в). При этом подобие уравнений для гравитационных и электромагнитных полей немедленно приводит к выражению $g = 4\pi G\sigma_m$, где σ_m – плотность массы. Вновь рассматривая теперь изменение энергии гравитационного поля вне двух массивных пластин при перемещении одной из них на расстояние δh , получаем

$$\varepsilon_g = \frac{g^2}{8\pi G}. \quad (255)$$

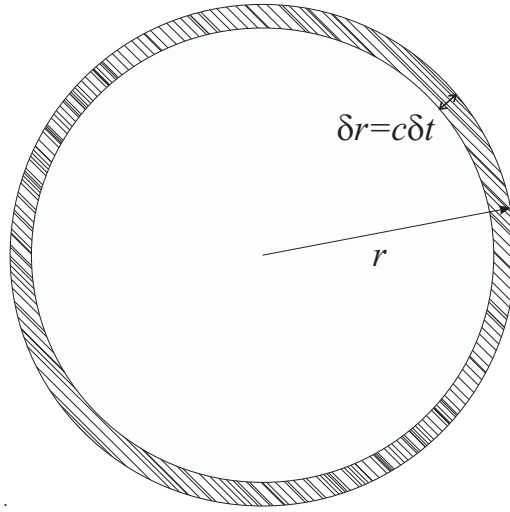


Рис. 36: Электромагнитная волна распространяется со скоростью c . Поэтому через время δt вся энергия $4\pi r^2 \epsilon_{\text{em}} c \delta t$, сосредоточенная в заштрихованной области, пересечет поверхность сферы радиуса r

7.4 Магнитодипольное излучение

Как уже упоминалось во введении, магнитодипольное излучение долго рассматривалось как основной механизм замедления вращения радиопульсаров. Лишь позже стало ясно, что основные потери энергии связаны с токовыми потерями (226). Тем не менее, до сих пор в большинстве работ, посвященных радиопульсарам, упоминается именно магнитодипольный механизм потерь энергии. Причина этого будет ясна из дальнейшего.

Рассмотрим намагниченный шар (радиус R , магнитное поле на полюсе B_0), вращающийся с угловой скоростью Ω . Если ось намагниченности параллельна оси вращения, то вне шара магнитное поле будет совпадать с полем точечного диполя, т.е. спадать с радиусом r как r^{-3}

$$B(r) \approx B_0 \left(\frac{R}{r} \right)^3. \quad (256)$$

Если же ось магнитного диполя наклонена к оси вращения, то выражение (256) будет справедливо лишь на малых расстояниях $r \ll c/\Omega$, когда за время $\delta t = r/c$ шар не успеет повернуться на достаточно большой угол. Действительно, благодаря конечности скорости распространения электромагнитного сигнала, о котором говорилось выше (см. Рис. 16), электромагнитное поле на расстоянии r от шара будет соответствовать не моменту времени t , а моменту времени $t - r/c$. Поэтому если мы отойдем от источника на расстояние $r \approx c/\Omega$, то направление поля должно существенно отличаться от статического.

Рассмотрим теперь т.н. волновую зону $r \gg c/\Omega$. В этой области электромагнитное поле должно быть близко к полю плоской электромагнитной волны, в которой амплитуды $|\mathbf{E}| = |\mathbf{B}|$ ($I_1 = 0$), причем электрическое и магнитные поля перпендикулярны друг

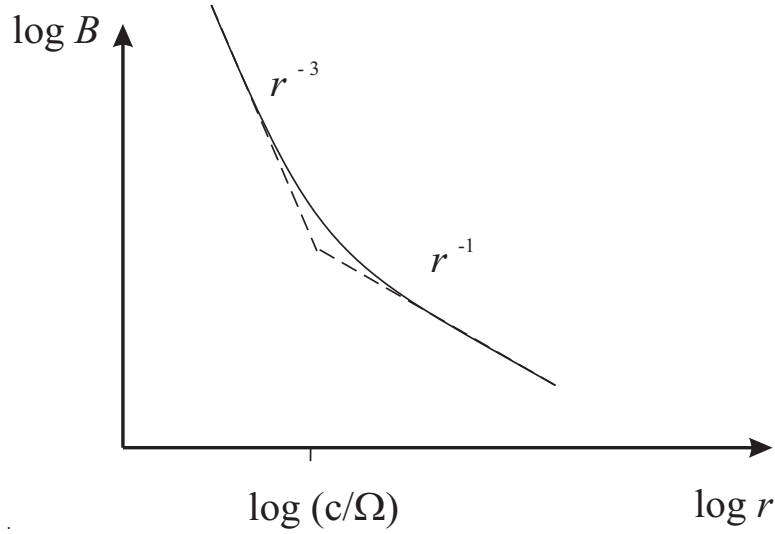


Рис. 37: Зависимость амплитуды переменного магнитного поля B от радиуса r для наклонного ротатора

другу ($I_2 = 0$). Поэтому по порядку величины здесь $\varepsilon_{\text{em}} \approx B^2/8\pi$, где здесь B – амплитуда магнитного поля. С другой стороны, электромагнитная волна распространяется со скоростью света, так что за время δt вся энергия $4\pi r^2 \varepsilon_{\text{em}} c \delta t$, сосредоточенная в заштрихованной области, показанной на Рис. 36, пересечет поверхность сферы радиуса r . Следовательно, потери энергии W (энергия/время) может быть оценена как

$$W_{\text{md}} \approx 4\pi r^2 \varepsilon_{\text{em}} c. \quad (257)$$

Конечно, при строгом анализе необходимо также учитывать зависимость плотности энергии от полярного угла θ . Однако поскольку нас здесь интересует лишь оценка по порядку величины, этой зависимостью можно пренебречь.

Понятно, что величина потерь энергии не должна зависеть от радиуса r , на котором мы оцениваем плотность энергии волны. Поэтому можно сделать вывод, что в волновой зоне плотность энергии должна вести себя как r^{-2} , так что магнитное поле будет спадать с расстоянием как r^{-1}

$$B(r) \sim \frac{A}{r}, \quad (258)$$

где A – некоторая постоянная. С другой стороны, при $r \sim c/\Omega$ поля (256) и (258) должны быть одного порядка (см. Рис. 37). Отсюда легко получить, что

$$A \sim B_0 \frac{\Omega^2 R^3}{c^2}. \quad (259)$$

Воспользовавшись теперь соотношениями (257) и (258), получаем окончательно

$$W_{\text{md}} \sim \frac{B_0^2 \Omega^4 R^6}{c^3}. \quad (260)$$

Как мы видим, оценка для мощности магнитодипольных потерь (260) в буквенном выражении оказалась в точности такой же, как и для потерь (226), связанных с продольными токами, текущими в магнитосфере радиопульсаров. По видимому, с этим и связан тот факт, что магнитодипольный механизм до сих пор используется для оценки потерь энергии нейтронных звезд.

7.5 Излучение гравитационных волн

Покажем теперь, что способ, которым мы воспользовались выше при выводе формулы (260) для потерь энергии при магнитодипольном излучении, позволяет получить и оценку формулы Эйнштейна (116) для мощности излучения гравитационных волн. Для этого нам нужно будет определить характер спада напряжения гравитационного поля (ускорения свободного падения) \mathbf{g} с расстоянием r .

Для простоты рассмотрим двойную систему, состоящую из звезд одинаковой массы M , вращающихся по круговым орбитам радиуса a (см. Рис. 38). Тогда $\Omega^2 a = GM/(2a)^2$, так что угловая скорость вращения есть

$$\Omega^2 = \frac{GM}{4a^3}. \quad (261)$$

Посчитаем ускорение свободного падения в ближней зоне $r \ll c/\Omega$ в направлении оси y . Поскольку в этом случае с достаточной точностью можно пользоваться выражением $g = GM/l^2$, где l – расстояние от тела до точки наблюдения, получаем

$$g_y = -\frac{GMa \sin \Omega t}{[(r - a \cos \Omega t)^2 + a^2 \sin^2 \Omega t]^{3/2}} + \frac{GMa \sin \Omega t}{[(r + a \cos \Omega t)^2 + a^2 \sin^2 \Omega t]^{3/2}}. \quad (262)$$

Проводя несложные вычисления, имеем в пределе $r \gg a$

$$g_y = -3 \frac{GMa^2 \sin(2\Omega t)}{r^4}. \quad (263)$$

Отметим два важных момента, которые следуют из соотношения (263). Прежде всего, мы видим, что несмотря на то, что двойная система вращается с угловой скоростью Ω , переменная часть величины g изменяется с удвоенной частотой 2Ω , поскольку уже через пол периода система перейдет точно в такое же состояние, как и в момент времени $t = 0$. При этом, в отличие от дипольного излучения, спадание гравитационного поля оказывается более быстрым: $g \propto r^{-4}$. Эти свойства, связанные с равенством гравитационной и инертной масс, показывают, что дипольное гравитационное излучение невозможно. Излучение гравитационных волн имеет квадрупольный характер.

Что же касается дальнейшего расчета, то он проводится точно также, как и для магнитодипольного излучения. В волновой зоне $r \gg c/\Omega$ мощность потерь запишется как

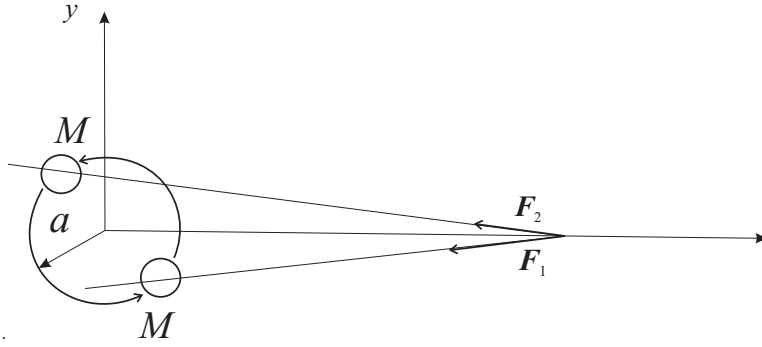


Рис. 38: Определение величины ускорения свободного падения \mathbf{g} от двойной системы. Сумма двух сил, \mathbf{F}_1 и \mathbf{F}_2 , имеет составляющую вдоль оси y , зависящую от времени

$W_g \approx 4\pi r^2 \varepsilon_g c$. Поскольку выражение для плотности энергии ε_g (255) пропорционально g^2 , мы приходим к заключению, что в волновой зоне

$$g(r) \sim \frac{A_g}{r}, \quad (264)$$

где A_g – некоторая постоянная. С другой стороны, при $r \sim c/\Omega$ поля (263) и (264) должны быть одного порядка (см. Рис. 39). Отсюда легко получить, что

$$A_g \sim \frac{GMa^2\Omega^3}{c^3}. \quad (265)$$

Воспользовавшись теперь соотношениями (255) и (264), а также выражением (261) для Ω , получаем окончательно

$$W_g \sim \frac{G^4 M^5}{c^5 a^5}, \quad (266)$$

в полном согласии с формулой Эйнштейна.

7.6 Функция масс

Как уже говорилось, функция масс двойной системы определяется с помощью эффекта Доплера, позволяющего найти радиальную (параллельную лучу зрения) компоненту скорости одной из звезд $K_1(t)$ (измеряется в км/с). Для нерелятивистских скоростей $V \ll c$ связь между частотой источника ν_0 и частотой ν , которую будет регистрировать наблюдатель, имеет вид

$$\nu = \frac{\nu_0}{1 - V_{\parallel}/c}, \quad (267)$$

где V_{\parallel} – проекция скорости источника на луч зрения. Соотношение (267) является чисто кинематическим, поскольку при $V_{\parallel} \neq 0$ свету необходимо разное время, чтобы достигнуть наблюдателя от разных частей траектории источника. Действительно, как показано на

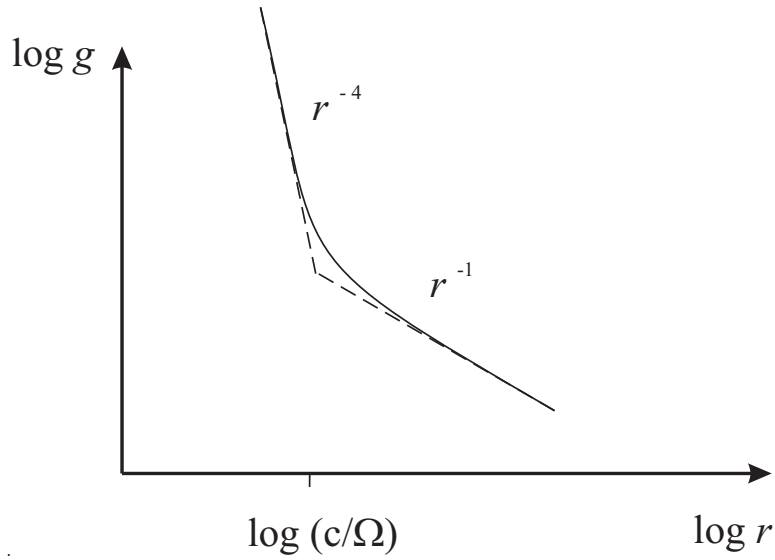


Рис. 39: Зависимость амплитуды напряженности переменного гравитационного поля g_y от радиуса r

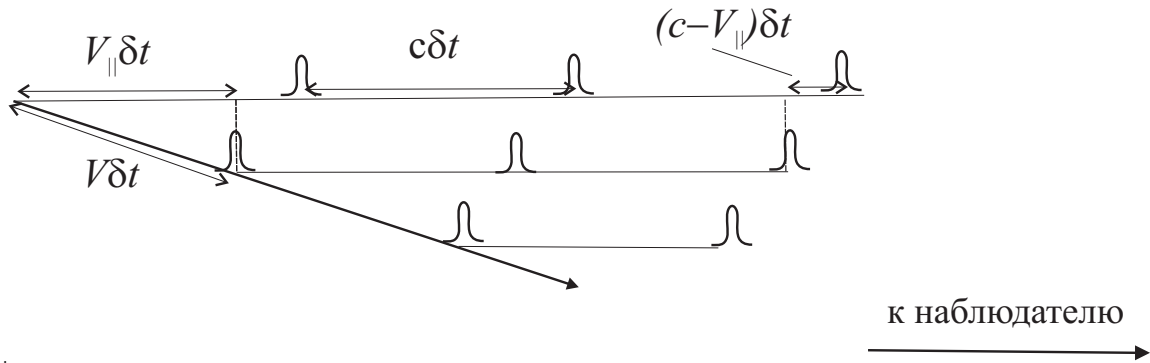


Рис. 40: Эффект Доплера

Рис. 40, если источник излучает импульсы через промежутки времени δt , то к наблюдателю они будут приходить с интервалами $\delta t - V_{||}\delta t/c$, что и отображено в соотношении (267).

Как уже подчеркивалось, для радиопульсаров мы можем определять не только его частоту вращения, но и время прихода каждого импульса. С учетом соотношения (267) можно записать выражение для разности времени прихода импульсов ΔP как

$$\Delta P = P \left(1 - \frac{V_{||}}{c} \right). \quad (268)$$

В результате, для эллиптической орбиты

$$\Delta P = P \left[\frac{K_1}{c} \cos \omega (\cos E + e) - \frac{K_1}{c} \sin \omega \sin E \right]. \quad (269)$$

Здесь ω – угол, определяющий положение периастра, а

$$K_1 = \frac{(2\pi G)^{1/3} M_2 \sin i}{P_b^{1/3} (M_1 + M_2)^{2/3} (1 - e^2)^{1/2}} \quad (270)$$

есть амплитуда наблюдаемой лучевой скорости (i – угол наклона плоскости орбиты к картинной плоскости). Зная зависимость экцентричной аномалии E от времени t , можно по форме зависимости лучевой скорости $K_1(t) = (\Delta P/P)c$ от времени определить и эксцентриситет орбиты e . Наконец, широко используемый в астрономии угол E есть т.н. экцентрисическая аномалия, фиксирующая положение тела на орбите. Угол E связан с углом φ , измеряемым относительно фокуса эллипса, соотношениями

$$r \cos \varphi = a(\cos E - e), \quad (271)$$

$$r \sin \varphi = a\sqrt{1 - e^2} \sin E, \quad (272)$$

так что $r = a(1 - e \cos E)$, где r – расстояние тела до фокуса орбиты, а a – большая полуось. Как мы видим, помимо известных величин (орбитальный период P_b , эксцентриситет e), величина K_1 зависит лишь от функции масс двойной системы

$$f(M) = \frac{M_2^3 \sin^3 i}{(M_1 + M_2)^2}. \quad (273)$$

Следовательно, величина $f(M)$ может быть непосредственно определена из наблюдений.

Полезно получить выражение для величины K_1 ”на пальцах”. Для этого рассмотрим двойную систему, в которой звезды движутся по круговым орбитам с радиусами a_1 и a_2 вокруг общего центра масс. Тогда закон Ньютона для наблюдаемой звезды с массой M_1 , движущейся по окружности радиуса a_1 , запишется в виде

$$\frac{M_1 V_1^2}{a_1} = \frac{G M_1 M_2}{(a_1 + a_2)^2}. \quad (274)$$

Здесь мы учли, что расстояние между звездами есть $a_1 + a_2$. Далее, очевидно, что

$$a_1 M_1 = a_2 M_2. \quad (275)$$

Наконец, еще одно уравнение следует из определения орбитального периода

$$P_b = \frac{2\pi a_1}{V_1}. \quad (276)$$

Учитывая теперь, что в выражении для эффекта Доплера входит проекция орбитальной скорости на луч зрения, получаем для амплитуды лучевой скорости $K_1^A = V_1 \sin i$

$$K_1^A = \frac{(2\pi G)^{1/3} M_2 \sin i}{P_b^{1/3} (M_1 + M_2)^{2/3}}, \quad (277)$$

что для круговых орбит совпадает с (270).

7.7 Гравитационное красное смещение в двойных системах

Посмотрим теперь, каково будет время задержки импульсов, если мы учтем гравитационное красное смещение. Здесь необходимо отметить, что наряду с гравитационным красным смещением необходимо учитывать также и т.н. квадратичный эффект Доплера, поскольку при скоростях v , сравнимых со скоростью света, необходимо использовать точную формулу для доплеровского сдвига частоты

$$\nu = \frac{\nu_0 \sqrt{1 - V^2/c^2}}{1 - V_{\parallel}/c}. \quad (278)$$

Как мы видим, здесь дополнительная поправка к частоте ν также оказывается пропорциональна c^{-2} .

В результате, изменение периода пульсаций следует записать в виде

$$\Delta P = P \left(1 - \frac{V_{\parallel}}{c} + \frac{V^2}{2c^2} + \frac{GM_2}{c^2 r_{12}} \right), \quad (279)$$

что при движении по эллиптической орбите дает

$$\Delta P = P \left[\left(\frac{K_1}{c} \cos \omega + \frac{eA}{c^2} \right) (\cos E + e) - \frac{K_1}{c} \sin \omega \sin E \right]. \quad (280)$$

Здесь

$$A = \frac{(2\pi G)^{2/3} M_2 (M_1 + 2M_2)}{P_b^{2/3} (M_1 + M_2)^{4/3} (1 - e^2)}. \quad (281)$$

Величина ePA/c^2 и дает амплитуду задержки γ_p (157).

Выражение (280) показывает, что задержка, связанная с гравитационным красным смещением, может быть выделена лишь в том случае, если эксцентриситет орбиты e отличен от нуля. В противном случае радиопульсар находился бы в постоянном гравитационном поле компаньона (а квадрат орбитальной скорости также был бы постоянным). Менее тривиальным обстоятельством является то, что для определения величины A необходимо также, чтобы имело место изменение долготы периастра. Дело в том, что при постоянном значении ω появление дополнительного слагаемого eA/c^2 в уравнении (280) приведет лишь к переопределению величины K_1 , определяющей, как мы видели, функцию масс системы.

7.8 Движение периастра

Прежде чем переходить к выводу формулы (150) для смещения периастра, рассмотрим колебание массы m на пружине вдоль оси x вокруг положения равновесия в точке x_0 . Закон сохранения энергии может быть записан в виде

$$\frac{m}{2} \left(\frac{dx}{dt} \right)^2 + \frac{k}{2} (x - x_0)^2 = \mathcal{E}, \quad (282)$$

где

$$U(x) = \frac{k}{2}(x - x_0)^2 \quad (283)$$

есть потенциальная энергия сжатой пружины а k – жесткость пружины. Для малых гармонических колебаний

$$x = x_0 + A_0 \sin(\Omega t), \quad (284)$$

и поэтому

$$\frac{dx}{dt} = A_0 \Omega \cos(\Omega t). \quad (285)$$

Поскольку в правой части уравнения (282) стоит величина, независящая от времени, то коэффициенты при $\sin^2(\Omega t)$ и $\cos^2(\Omega t)$ в левой части уравнения (282) должны быть одинаковыми:

$$\frac{mA_0^2\Omega^2}{2} = \frac{kA_0^2}{2}, \quad (286)$$

что дает известное соотношение $\Omega^2 = k/m$. С другой стороны, величина k есть вторая производная от потенциала U , взятая в точке равновесия $x = x_0$: $k = U''(x_0)$. В итоге, получаем

$$\Omega^2 = \frac{U''(x_0)}{m}. \quad (287)$$

Таким образом, для определения частоты малых колебаний нужно знать лишь поведение потенциала вблизи положения равновесия $x = x_0$.

Покажем теперь, как с помощью соотношения (287) можно получить выражение (150) для угла смещения периастра за один период. Для этого воспользуемся явным выражением эффективного потенциала $U_{\text{эфф}}(r)$ (288)

$$U_{\text{эфф}}(r) = -\frac{mc^2}{2(e')^2} \left(1 - \frac{r_g}{r}\right)^2 \left[(e')^2 - \left(1 - \frac{r_g}{r}\right) \left(1 + \frac{L^2}{c^2 r^2}\right) \right] + \mathcal{E} - mc^2. \quad (288)$$

При этом необходимо учесть тот факт, что нас интересуют лишь малые поправки к нерелятивистскому случаю, когда отношение r_g/r_0 можно считать малым. Поэтому радиус круговой орбиты r_c , определяемый из условия $dU_{\text{эфф}}/dr = 0$, может быть записан как

$$r_c = r_0 \left(1 - \frac{3}{2} \frac{r_g}{r_0}\right), \quad (289)$$

где

$$r_0 = \frac{2L^2}{c^2 r_g} = \frac{L^2}{GM} \quad (290)$$

– радиус круговой орбиты без учета релятивистских эффектов. После элементарных, хотя и громоздких вычислений, получаем

$$U''(r_c) = \frac{mL^2}{(e')^2 r_0^4} \left(1 + \frac{r_g}{r_0}\right). \quad (291)$$

Поэтому частота радиальных колебаний Ω_r для случая $r_g/r_0 \ll 1$ будет равна

$$\Omega_r = \frac{L}{e'r_0^2} \left(1 + \frac{1}{2} \frac{r_g}{r_0} \right). \quad (292)$$

С другой стороны, частота вращения по орбите Ω_φ легко может быть найдена из закона сохранения углового момента $L = rv_\varphi\gamma$ (188). Здесь нужно лишь не забыть, что скорость $v_\varphi = r d\varphi/d\tau$ измеряется по часам наблюдателя, находящегося в гравитационном поле. В итоге, получаем для угловой частоты вращения $\Omega_\varphi = d\varphi/dt$

$$\Omega_\varphi = \frac{L}{e'r_0^2} \left(1 + 2 \frac{r_g}{r_0} \right). \quad (293)$$

Здесь мы вновь воспользовались условием $\alpha\gamma = e'$, а также малостью отношения r_g/r .

Как мы видим, невозмущенные величины Ω_φ и Ω_r совпадают друг с другом. Это и означает, что в ньютоновской теории траектория движения является замкнутой. Учет же релятивистских поправок приводит к тому, что периоды орбитального ($P_\varphi = 2\pi/\Omega_\varphi$) и радиального ($P_r = 2\pi/\Omega_r$) движения уже не будут в точности равны друг другу. В результате, за один период орбита повернется на дополнительный угол $\Delta\varphi = \Omega_\varphi\Delta t$, где $\Delta t = 2\pi/\Omega_r - 2\pi/\Omega_\varphi$ – разность периодов колебаний. В итоге, получаем

$$\Delta\varphi = 2\pi \frac{\Omega_\varphi - \Omega_r}{\Omega_r} = 2\pi \frac{3GM}{c^2 r_0}, \quad (294)$$

что совпадает с точным значением (150).

8 Упражнения

1. Покажите, что для определения вида преобразований Лоренца достаточно использовать следующие условия:

- (а) при поворотах остается инвариантным интервал $ds^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2$,
- (б) два последовательных преобразования Лоренца вдоль оси x со скоростями V_1 и V_2 есть также преобразование Лоренца с некоторой скоростью V_3 .

2. Найдите кривизну поверхности параболоида вращения

$$x^2 + y^2 - z^2 = R^2. \quad (295)$$

3. Найдите коэффициент λ в выражении

$$L_0/R - 2\pi = \lambda k \delta S. \quad (296)$$

4. Найдите расстояние между точками с координатами $r = r_g$ и $r = 2r_g$ для шварцшильдовской и керровской черной дыры.

5. Покажите, что в рамках нерелятивистской теории Всемирного тяготения для потенциала

$$\varphi_g = -\frac{GM}{r - r_g} \quad (297)$$

(т.н. потенциал Пачинского-Вииты) радиус последней устойчивой орбиты, как и в метрике Шварцшильда, задается соотношением $r_0 = 3r_g$, а скорость движения периастра составляет $2/3$ от истинного значения (294).

6. Переходя при интегрировании к углу E (271)–(272), покажите, что среднее по времени расстояние планеты от фокуса есть

$$r_t = a \left(1 + \frac{e^2}{2} \right). \quad (298)$$

7. Получите формулу Эйнштейна для мощности излучения гравитационных волн при условии $M_1 \ll M_2$. Покажите, что и в этом случае излучение идет на удвоенной частоте.

8. Покажите прямой подстановкой, что поверхность эргосферы черной дыры $r = r_e$ (214), в пределах которой отсутствие вращения для любого тела становится невозможным, задается условием $\alpha^2 c^2 = \omega^2 g_{\varphi\varphi}$.

9. Определите уравнение кривой на Рис. 24.

10. О какой планете идет речь на Рис. 20?