

Федеральное агентство по образованию
Московский инженерно-физический институт
(государственный университет)

К.А. Бронников С.Г. Рубин

Лекции по гравитации и космологии

Рекомендовано УМО “Ядерная физика и технологии”
в качестве учебного пособия
для студентов высших учебных заведений

Москва 2008

УДК 524.8(075) + 530.12:531.51(075)

ББК 22.313.7

Б88

Бронников К.А., Рубин С.Г. Лекции по гравитации и космологии. Учебное пособие. М.: МИФИ, 2008. 460 с.

В книге излагаются основы теории гравитации и космологии. Рассматриваются основные понятия физики черных дыр и кротовых нор. Обсуждаются проблемы, вытекающие из предположения о существовании дополнительных измерений, и способы их решения.

Для студентов физических специальностей.

Пособие подготовлено в рамках Инновационной общеобразовательной программы.

Рецензент А.В. Берков

ISBN 978-5-7262-1021-6

© Московский инженерно-физический институт
(государственный университет), 2008

Оглавление

Глава 1. Современные представления о гравитации и космологии — краткий очерк	12
Глава 2. Основные положения ОТО	37
2.1. СТО. Геометрия Минковского	38
2.1.1. Геометрия	38
2.1.2. Преобразования координат	40
2.1.3. Кинематические эффекты	42
2.1.4. Элементы релятивистской механики точки	45
2.2. Риманово пространство-время. Системы координат и системы отсчета	47
2.2.1. Ковариантность, карты и атласы	48
2.2.2. Системы отсчета и относительность	50
2.2.3. Системы отсчета и хронометрические инварианты	51
2.2.4. Ковариантность и относительность	53
2.3. Риманово пространство-время	55
2.4. Действие для гравитационного поля и динамические уравнения	59
2.4.1. Уравнения Эйнштейна	59
2.4.2. Уравнения геодезических	61
2.4.3. Принцип соответствия	62
2.5. Макроскопическая материя и негравитационные поля в ОТО	63

2.5.1.	Идеальная жидкость	64
2.5.2.	Скалярные поля	65
2.5.3.	Электромагнитное поле	66
2.6.	Наиболее симметричные пространства	67
2.6.1.	Группы изометрий и векторы Киллинга	67
2.6.2.	Изотропная космология. Пространства dS и AdS	69
Глава 3. Черные дыры		71
3.1.	Сферически-симметричные гравитационные поля	72
3.2.	Решение Райснера — Нордстрема — (анти)-де Ситтера	77
3.2.1.	Решение уравнений Эйнштейна	77
3.2.2.	Частные случаи	79
3.3.	Горизонты и геодезические в статической сферически- симметричной метрике	84
3.3.1.	Общий вид уравнений геодезических	84
3.3.2.	Горизонты, геодезические и квазиглобаль- ная координата	86
3.3.3.	Горизонты, R- и T-области	90
3.4.	Черная дыра Шварцшильда. Геодезические и гло- бальное описание	92
3.4.1.	R- и T-области	92
3.4.2.	Геодезические в R-области	94
3.4.3.	Захват частиц черной дырой	96
3.4.4.	Глобальное описание — метрика Крускала	98
3.4.5.	От диаграммы Крускала к диаграмме Картера- Пенроуза для поля Шварцшильда	100
3.5.	Глобальная причинная структура пространств с горизонтами	103
3.5.1.	Переход через горизонт в общем случае	103
3.5.2.	Построение диаграмм Пенроуза	106
3.6.	Черная дыра как результат гравитационного кол- лапса	111

3.7.	Черные дыры и безмассовые скалярные поля . . .	116
3.7.1.	Общая СТТ и преобразование Вагонера . . .	116
3.7.2.	Скалярное поле с минимальной связью . . .	118
3.7.3.	Скалярное поле с конформной связью . . .	122
3.7.4.	Аномальные поля. Решение анти-Фишера .	126
3.8.	Скалярные поля с произвольными потенциалами. Теоремы запрета	129
3.8.1.	Для чего нужны теоремы запрета	129
3.8.2.	Основные уравнения	132
3.8.3.	Теоремы о глобальной структуре	133
3.8.4.	Теоремы “no-hair”	137
3.9.	Черные дыры с вращением	139
3.10.	Термодинамика черных дыр	142
3.10.1.	Четыре закона термодинамики ЧД	142
3.10.2.	Испарение черных дыр	144
3.11.	Регулярные черные дыры и черные вселенные	145
3.11.1.	Различные виды регулярных черных дыр .	145
3.11.2.	Черные вселенные с минимально-связанным скалярным полем	151
Глава 4.	Кротовые норы	157
4.1.	Понятие кротовой норы	157
4.2.	Кротовая нора как машина времени	162
4.3.	Кротовые норы как решения уравнений гравитации	166
4.3.1.	Сферически-симметричные кротовые норы. Общие свойства	167
4.3.2.	Кротовые норы с безмассовыми скалярными полями	170
4.3.3.	Построение кротовых нор с помощью решения следа уравнений Эйнштейна . . .	173
4.3.4.	Альтернативная гравитация и вакуум как источники кротовых нор	177

4.4. Возможные наблюдательные проявления кротовых нор	181
Глава 5. Этапы развития Вселенной	184
5.1. Космологический принцип и уравнения Эйнштейна	184
5.2. Пространство де Ситтера	192
5.3. Инфляция	197
5.4. Послеинфляционные стадии	203
5.4.1. Постинфляционный разогрев Вселенной . .	203
5.4.2. Стадия доминантности излучения	207
5.4.3. Стадия доминантности вещества	210
5.4.4. Современная стадия ускоренного расширения (вторичной инфляции)	212
5.4.5. Будущее Вселенной: ожидается ли “Большой разрыв”?	214
5.5. Масштабный фактор в общем случае	217
5.6. Зачем нужен инфляционный период?	220
5.6.1. Проблема плоскостности	221
5.6.2. Начальный размер Вселенной	223
5.6.3. Причинные связи во время инфляции и после нее	223
5.7. Основные свойства расширяющегося пространства	225
5.7.1. Красное смещение	225
5.7.2. Фотометрическое расстояние	227
5.7.3. Скорость частиц в пространстве ФРУ . . .	229
Глава 6. Инфляционные модели	231
6.1. Квадратичная инфляция	232
6.2. Гибридная инфляция	234
6.2.1. Образование черных дыр в модели гибридной инфляции	239
6.2.2. Гибридная инфляция с наклоном	241
6.3. Влияние массивных полей на процесс инфляции .	246
6.4. Фазовые переходы первого рода	254

6.4.1.	Квантовые поправки к вероятности распада вакуума	257
6.4.2.	Подавление распада вакуума виртуальными частицами	269
Глава 7.	Барионная асимметрия Вселенной	275
7.1.	Механизм спонтанного бариогенезиса	276
7.2.	Крупномасштабные флуктуации барионного заряда	281
7.3.	Обсуждение	286
Глава 8.	Крупномасштабная структура	289
8.1.	Космический микроволновый фон	289
8.2.	Квантовые флуктуации во время инфляции	292
8.2.1.	Предварительное рассмотрение	292
8.2.2.	Подробное рассмотрение	295
8.2.3.	Классическая эволюция квантовых флуктуаций	305
8.3.	Развитие возмущений плотности	308
8.3.1.	Возмущения плотности в пространстве Минковского	308
8.3.2.	Возмущения плотности в расширяющейся Вселенной	311
8.4.	Массивные первичные черные дыры	312
Глава 9.	Многомерие	320
9.1.	Введение	320
9.2.	Компактные дополнительные пространства	323
9.2.1.	Модель Калуцы — Клейна с одним дополнительным измерением	326
9.2.2.	Модель Калуцы-Клейна. Общий случай.	334
9.3.	Нелинейная гравитация в моделях с одним дополнительным пространством	340
9.3.1.	$F(R)$ -теория. Основные уравнения	340
9.3.2.	Приближение медленных изменений. Картина Эйнштейна	342

9.3.3.	Некоторые оценки	346
9.3.4.	Квадратичная гравитация с космологической постоянной	349
9.3.5.	Кубическая гравитация	351
9.3.6.	Обобщения	352
9.3.7.	Интенсивное рождение частиц в постинфляционный период	355
9.3.8.	Поведение поля вблизи $K = 0$	359
9.3.9.	Выводы	362
9.4.	Нелинейная D -мерная гравитация с несколькими дополнительными пространствами	363
9.4.1.	Приближение медленных изменений. Редукция к d_0 измерениям	365
9.4.2.	Эффекты структуры дополнительных факторпространств	368
9.5.	Асимметричные браны в $d_0 = d + 2$ измерениях	373
9.5.1.	Геометрия	375
9.5.2.	Уравнения поля	379
9.5.3.	Асимптотическое поведение решений	380
9.5.4.	Симметричные и несимметричные браны	382
9.5.5.	Конфигурации с возрастающими масштабными факторами	384
9.5.6.	Удержание материи	387
9.5.7.	Гравитация на бране	390

Глава 10. Тонкая настройка параметров Вселенной 396

10.1.	Зачем нужно слабое взаимодействие?	396
10.1.1.	Масса электрона и свойства Вселенной	396
10.1.2.	Углеродный уровень	397
10.1.3.	Медленные реакции в звездах	398
10.1.4.	Примеры тонкой настройки Вселенной	400
10.1.5.	Какой должна быть Окончательная Теория?	402
10.2.	Механизмы тонкой настройки	405

10.2.1. Каскадная редукция	405
10.2.2. Случайный потенциал	424
Заключение	433
Спор между сторонниками множественности вселенных (M) и Единой теории (E)	433
Литература	437

Обозначения

Во всей книге $\hbar = c = 1$ (кроме раздела 2.1);

Сигнатура метрики: $(1, -1, -1, -1)$.

Космологическая метрика с синхронным временем:

$$ds^2 = dt^2 - \gamma_{\alpha\beta} dx^\alpha dx^\beta.$$

Единицы измерения

1 световой год $\sim 10^{18}$ см.

1 парсек ~ 3 световых года.

Масса Солнца $M_\odot \sim 2 \cdot 10^{33}$ г.

Светимость Солнца $L_\odot \sim 4 \cdot 10^{33}$ эрг/с.

Планковские единицы:

$$M_{\text{Pl}}^2 = 1/G \text{ или } M_{\text{Pl}}^2 = 1/(8\pi G).$$

$$M_{\text{Pl}} \sim 10^{-5} \text{ г} \sim 10^{19} \text{ ГэВ.}$$

$$l_{\text{Pl}} = 1/M_{\text{Pl}} \sim 10^{-33} \text{ см.}$$

$$t_{\text{Pl}} = 1/M_{\text{Pl}} \sim 10^{-44} \text{ с.}$$

Основные параметры нашей Вселенной

Размер Вселенной $\sim 10^{28}$ см или ~ 6000 Мпк.

Возраст Вселенной ~ 14 млрд лет.

Параметр Хаббла $H_0 = 100h$ км с $^{-1}$ Мпк $^{-1}$; $h \simeq 0.72$.

Состав Вселенной: $\Omega_\Lambda \simeq 0.72$; $\Omega_{\text{DM}} \simeq 0.24$; $\Omega_b \simeq 0.04$.

$\sim 10^{11}$ галактик, сосредоточенных в скоплениях с пустотами (voids) между ними.

Галактика

Размер Галактики ~ 50 кпк.

$\sim 10^{11}$ звезд, нейтронные звезды, белые карлики, темные (коричневые) карлики, черные дыры, скрытая масса (темная материя).

История Вселенной, основные вехи

Время с момента Большого взрыва	Температура среды, К	Примечания
$10^{-43} - 10^{-37}$ с	$> 10^{26}$	Инфляция
10^{-6} с	$> 10^{12}$	Плазма из кварков, глюонов, электронов, нейтрино и их античастиц
$3 \cdot 10^{-5}$ с	10^{12}	Образование нуклонов
10^{-4} - 3 мин	$10^{12} - 10^9$	Образование дейтерия, гелия и лития
15 тыс. лет	10^4	Равенство плотностей энергии материи и излучения
300 тыс. лет	4000	Образование атомов, прозрачная Вселенная
15 млн лет	300	Комнатная температура
1-3 млрд лет	20	Рождение первых звезд
3 млрд лет	10	Образование тяжелых ядер в результате взрыва сверхновых
3-14 млрд лет	3	Появление разумной жизни
10^{14} лет	-	Прекратился процесс звездообразования
10^{37} лет	-	Погасли последние звезды
10^{100} лет	-	Все черные дыры испарились

Глава 1

Современные представления о гравитации и космологии — краткий очерк

В начале XX века были известны всего два физических поля — электромагнитное и гравитационное. Созданная в 1902–1905 гг. специальная теория относительности (СТО) отлично справлялась с описанием механических и электромагнитных явлений при любых скоростях вплоть до световой, чего не получалось в классическом подходе Галилея и Ньютона. Но ньютоновская теория тяготения, служившая практически идеальной основой небесной механики и земной физики, формулировалась в старых понятиях абсолютного пространства и абсолютного времени и не вписывалась в новые представления.

После создания СТО были многочисленные попытки описать гравитационное поле в пространстве Минковского, надеясь таким образом включить в описание быстро движущиеся тяготеющие объекты, получив ньютоновскую теорию гравитации в пределе малых скоростей движения тяготеющих тел.

Великий французский математик Анри Пуанкаре, фактически одновременно с Эйнштейном “открывший” СТО, первым по-

пытался распространить ее и на гравитацию, предположив конечную скорость распространения гравитационного поля. Мысли о том, что гравитация передается со скоростью света, высказывались, конечно, и раньше. Подобные теории, будучи внутренне непротиворечивыми, встретились тем не менее с рядом серьезных трудностей. Одна из них — неспособность объяснить аномальное вековое смещение перигелия орбиты Меркурия, около $43''$ в столетие, необъяснимое и в теории Ньютона. Уже тогда это смещение вполне уверенно обнаруживалось из астрономических наблюдений.

По законам Кеплера, вытекающим из закона всемирного тяготения Ньютона, планеты в Солнечной системе движутся по замкнутым эллиптическим орбитам. Впрочем, эта идеальная картина была бы верна, если бы каждая планета взаимодействовала только с Солнцем; влияние планет друг на друга слегка искажает орбиты и, в частности, их эллипсы немного поворачиваются со временем, превращаясь в незамкнутые спиралеобразные траектории. Такие эффекты хорошо рассчитываются в ньютоновской небесной механике. Аномальный сдвиг перигелия Меркурия остается уже после вычитания всех “нормальных” сдвигов, в сумме составляющих несколько тысяч угловых секунд в столетие.

Меркурий, кстати, — самая быстрая из планет Солнечной системы. Для искомой “гравитации высоких скоростей” объяснение особенностей его движения должно было стать первым пробным камнем.

Еще одно обстоятельство теоретического характера делало попытки описать гравитацию в рамках СТО малопривлекательными. Со времен Галилея было известно, что если исключить сопротивление воздуха, то самые разные тела — пушинка, кусок дерева, камень, слиток свинца, сосуд с водой и так далее — падают на Землю с совершенно одинаковым ускорением. Универсальность ускорения свободного падения была подтверждена с высокой точностью (до 10^{-9}) в конце XIX века опытом Этве-

ша с крутильными весами: фактически, была установлена эквивалентность между силой земного притяжения и инерционным центробежным ускорением за счет суточного вращения Земли. В уравнениях Ньютона это проявляется как равенство между инертной и гравитационной массами — так называемый принцип эквивалентности (ПЭ). Сама теория Ньютона объяснить это равенство неспособна, как и все ее обобщения в рамках СТО.

Инертная масса, фигурирующая во втором законе механики Ньютона $\vec{a} = \vec{F}/m$ (ускорение равно силе, деленной на массу), и гравитационная, входящая в закон тяготения — величины, по существу, совершенно разной физической природы. Эйнштейну было ясно, что равенство инертной и гравитационной масс не может быть случайным совпадением и должно иметь глубокие причины. Универсальность действия гравитации на тела привела его к идее, ставшей основой общей теории относительности (ОТО): гравитационное поле есть свойство самого пространства, причем свойство, меняющееся от точки к точке, так как поле тяготения, вообще говоря, неоднородно. Следовательно, пространство Минковского — плоское, однородное (одинаковое во всех точках) и изотропное (одинаковое во всех направлениях) — не годится, гравитация должна его искажать и искривлять. Так возникает идея кривизны физического пространства-времени.

У главной идеи ОТО, как, наверное, у всякой фундаментальной идеи, разумеется, были и предтечи, и провозвестники. Уже первооткрыватель неевклидовой геометрии Н.И. Лобачевский в 1826 году говорил об экспериментальном определении геометрии мира. Зависимость кривизны пространства от свойств заполняющей его материи предполагали Риман (1854) и Клиффорд (1876), причем у Клиффорда можно найти и мысль о кривизне, распространяющейся волнами. Идеи, как говорится, витали в воздухе. Однако именно Эйнштейн (во взаимодействии с Гильбертом, Пуанкаре и другими крупными математиками и физиками того времени) оформил их в стройную, логически непротиворечивую теорию.

Плод, что называется, созрел к 1915 году. ОТО стала еще одним шагом в сторону от простых и наглядных представлений классической физики. В ней четырехмерное пространство-время (часто для краткости говорят просто “пространство”) стало искривленным. Уже существовавшая к тому времени геометрия Римана стала математическим аппаратом, языком новой физической теории.

В римановой геометрии, а следовательно, и в ОТО, основная характеристика пространства — симметричный метрический тензор $g_{\mu\nu}(x^\alpha)$ (метрика), зависящий от четырехмерных координат x^μ , следовательно, меняющийся от точки к точке. Метрика несет информацию об интервалах между точками-событиями, через нее выражаются величины, характеризующие искривление пространства-времени — тензоры Римана и Риччи, и именно компоненты метрического тензора являются неизвестными в динамических уравнениях ОТО — уравнениях Эйнштейна (или Гильберта — Эйнштейна, как иногда говорят, желая подчеркнуть роль великого немецкого математика в создании ОТО):

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = -\frac{8\pi G}{c^4}T_{\mu\nu} \quad (1.1)$$

В общем случае это система из десяти нелинейных дифференциальных уравнений в частных производных относительно десяти неизвестных функций четырех пространственно-временных координат. Основной их смысл заключается в непосредственной связи кривизны пространства-времени (левая часть уравнений) с распределением и движением материи (правая часть, тензор энергии-импульса). Таким образом, “материя говорит пространству, как ему искривляться”. Всякое решение уравнений Эйнштейна описывает какую-то мыслимую конфигурацию гравитационного поля.

Как СТО не отменила механику Ньютона (пригодную на малых скоростях), так и ОТО не отменила СТО, которая справедлива на любом небольшом участке искривленного, но гладкого пространства-времени. Чем меньше размеры такого участ-

ка по сравнению с радиусом кривизны пространства, тем точнее такой участок совпадает с некоторой областью касательного к нему плоского пространства Минковского и с тем большей точностью выполняются СТО и ее многочисленные следствия. Ньютоновская теория тяготения тоже следует из ОТО в соответствующих условиях: ньютоновские уравнения получаются из уравнений ОТО в пределе малой кривизны (то есть в слабых гравитационных полях) и малых относительных скоростей тяготеющих тел. Большинство наблюдаемых явлений попадает как раз в такой вот “слабый” режим. Правда, в ОТО совершенно иначе интерпретируются гравитационные силы: теперь это не силы, а некоторые геометрические характеристики мировых линий, по которым движутся тела в четырехмерном пространстве-времени. С точки зрения ОТО, тело, свободно падающее в поле тяготения, движется вообще без внешних сил, и его мировая линия — геодезическая (то есть кратчайшая) в кривом четырехмерном мире, аналог прямой линии в плоском пространстве.

Важнейшее новое явление, которого нет в теории Ньютона, но предсказывается ОТО — гравитационные волны. Их существование непосредственно следует из волнового характера уравнений ОТО и подтверждается (пока, правда, лишь теоретически) многочисленными их волновыми решениями. Имеются лишь косвенные экспериментальные подтверждения факта их существования на основе анализа динамики пульсаров.

ОТО охотно приняла вызов со стороны наблюдений и с удивительной точностью объяснила упомянутую выше аномалию в движении Меркурия, бывшую до того камнем преткновения всех теорий тяготения. Другой классический эффект ОТО, поддающийся проверке, — действие гравитации на свет, приводящее к искривлению светового луча в поле небесного тела. По расчетам Эйнштейна, проходя рядом с Солнцем, световой луч должен отклониться на угол в $1,75''$. Аналогичный эффект можно получить и в ньютоновской теории, представляя свет потоком частиц, летящих — естественно — со скоростью света, но тогда расчет-

ное отклонение будет вдвое меньше: около $0,87''$ при пролете у самого края светила.

Полное солнечное затмение 29 мая 1919 года дало возможность измерить этот эффект, фотографируя изображения звезд рядом с закрытым Луной солнечным диском и сличая полученные кадры с обычными ночными снимками того же участка звездного неба. Как и ожидалось, на картинках с затмением звезды оказались чуть-чуть отодвинуты от края диска по сравнению с их ночными положениями. Угол отклонения варьировался в данных разных наблюдателей в пределах от $1,61''$ до $1,98''$ возле края диска, постепенно уменьшаясь по мере удаления от него, при ошибке в пределах $0,30''$. Можно сказать, небо подтвердило правоту Эйнштейна.

Это был подлинный триумф: теория, рожденная на кончике пера, стремительно завоевала место под Солнцем. И в общем сохраняет лидирующее положение и в наши дни, выдержав проверку временем и всевозможными экспериментами.

Эйнштейн после Эйнштейна

Но не будем забегать вперед, вернемся к 1920–30-м годам. Это было время активного проникновения физики в микромир и формирования языка, адекватного свойствам микромира — квантовой механики, а позднее квантовой электродинамики и — шире — квантовой теории поля. Квантовая теория поначалу строилась в рамках старых, ньютоновских понятий абсолютного пространства и абсолютного времени (нерелятивистская квантовая механика) и с немалыми усилиями осваивала мир высоких скоростей и высоких энергий, обретая содержание в четырехмерном пространстве-времени Минковского.

Понимание гравитации как кривизны пространства придавало ОТО исключительный характер по сравнению со всей остальной физикой, а это противоречило важному как для философов, так и для физиков ощущению единства материального мира. С другой стороны, в самой ОТО возникает ряд важных про-

блем, одна из них известна как проблема энергии. Как известно, понятия энергии и других сохраняющихся величин играют весьма существенную роль в построении квантовой теории. В плоском пространстве без затруднений формулируются законы сохранения энергии, импульса, момента импульса, которые связаны с симметрией пространства Минковского относительно временных и пространственных сдвигов и поворотов, образующих 10-параметрическую группу Пуанкаре. В искривленном пространстве подобных симметрий в общем случае нет совсем, и поэтому определить энергию и импульс гравитационного поля в ОТО непротиворечивым образом затруднительно.

По этой и некоторым другим причинам не все физики согласились с ОТО, и по сей день продолжаются попытки построения теории гравитации в пространстве Минковского. В отличие от первых подобных попыток, новые авторы научились объяснять классические наблюдаемые эффекты ОТО, а гравитация в них представляется полем с нормальными законами сохранения и с надеждами на квантование наравне с другими физическими полями. Согласно книге Уилла [308], уже к 1960 году таких теорий насчитывалось не менее 25. Но ни тогда, ни впоследствии такие теории не вызвали по-настоящему серьезного интереса (правда, их приверженцы, скорее всего, с этим не согласятся).

А вот противоположная тенденция “свести всю физику к геометрии” породила целый ряд новых идей, которые и поныне остаются актуальными в теоретической физике. В этой связи ОТО рассматривалась (и рассматривается поныне) как основа для обобщения, которое достигалось за счет введения более сложных видов геометрии, чем риманова (Вейль, Эддингтон, Картан), повышения размерности пространства-времени путем введения дополнительных невидимых координат (Калуца, Клейн), расширения требований к симметрии исходной формулировки теории (принцип калибровочной симметрии Вейля). Ставилась амбициозная задача, выходящая за рамки простого объединения электромагнитного и гравитационного полей: получить из еди-

ного поля заодно и характеристики тех немногих элементарных частиц, которые к тому времени были уже известны. Эйнштейн не только не остался в стороне от этих усилий, но и был лидером программы построения единой теории поля на основе ОТО и оставался им до конца жизни.

Описание этих попыток увело бы нас далеко от основной темы — гравитации. Поэтому приведем лишь слова, сказанные Вернером Гейзенбергом в начале 60-х: “Это великолепная в своей основе попытка... Но в то самое время, когда Эйнштейн занимался единой теорией поля, непрерывно открывались новые элементарные частицы, а с ними — сопоставленные им новые поля. Вследствие этого для проведения эйнштейновской программы еще не существовало твердой эмпирической основы, и попытка Эйнштейна не привела к каким-либо убедительным результатам”.

Задача построения единой “теории всего на свете” (“theory of everything”) остается и сегодня центральной задачей теоретической физики.

Технологический прорыв

К концу 1950-х физика знала уже не два, а четыре фундаментальных взаимодействия — гравитационное, электромагнитное, сильное ядерное (благодаря которому протоны и нейтроны объединены в атомные ядра) и слабое ядерное (отвечает за многие превращения частиц и ядерные реакции, из которых наиболее известен бета-распад). Среди них гравитационное взаимодействие представлялось чем-то малосущественным: будучи, применительно к частицам, на много порядков слабее даже слабого взаимодействия, оно казалось совершенно неважным в физике микромира. Все новые и новые экспериментальные данные о трех остальных взаимодействиях добывались на ускорителях. Бурно развивалась квантовая теория поля в плоском пространстве-времени на основе СТО, ставя и решая проблемы физики элементарных частиц. На этом фоне занятия гравитаци-

ей казались как бы чудачеством. Признавалась важность ОТО как фундаментальной мировоззренческой, почти философской теории, но ее экспериментальный базис был невелик: один эффект — в движении Меркурия проверялся с точностью около 1% и один — отклонение света вблизи Солнца — с точностью около 30%. Космологические наблюдения могли лишь свидетельствовать о нетривиальности геометрии Вселенной, но не о справедливости тех или иных уравнений гравитации... Кипу Торну, в то время студенту, а сегодня одному из корифеев гравитационной физики, профессора советовали не заниматься ОТО, теорией, по их мнению, очень слабо связанной с остальной физикой и астрономией. Он, как известно, не послушался и, надо полагать, не жалеет.

Ситуация стала меняться только в конце 50-х — начале 60-х годов. Развитие экспериментальной техники позволило запланировать и осуществить ряд новых проверок теории гравитации, а астрономические наблюдения все больше свидетельствовали о реальности источников сильных гравитационных полей в космическом пространстве. Возросло число альтернативных теорий гравитации. Были предсказаны десятки новых эффектов, не меньше было и предложений по их проверке.

Тем более поразительно, что именно ОТО подтверждается со все возрастающей точностью. Вот как, в общих чертах, выглядит ее экспериментальный статус на сегодняшний день.

Одна из первооснов теории — принцип эквивалентности — сегодня проверен с применением крутильных весов с различными пробными телами с впечатляющей точностью до 10^{-12} [2]. Похоже, возможности эксперимента на поверхности Земли этим исчерпаны: мешают многочисленные атмосферные, сейсмические и техногенные шумы. Планируемый спутниковый эксперимент STEP (Satellite Test of the Equivalence Principle) [312] позволит повысить точность до 10^{-17} – 10^{-18} . Принцип эквивалентности предсказывается всеми метрическими теориями гравитации — обобщениями ОТО, в которых гравитация отождествляется с

кривизной.

Другой, тоже универсальный эффект, одинаковый в большом классе теорий — так называемое гравитационное красное смещение. Его суть проста: фотон, удаляясь от тяготеющего центра, теряет энергию, перемещаясь в более длинноволновую часть спектра (“краснеет”), а приближаясь — “голубеет”. Примерно как камень, подброшенный вверх, теряет скорость, а падающий вниз ее увеличивает. В ОТО этот эффект связан с замедлением хода часов: чем ближе к источнику гравитационного поля, тем их ход медленнее. Он проверен как для фотонов (опыты Паунда, Ребки и Снайдера с резонансным поглощением гамма-фотонов атомными ядрами) [248, 249], так и непосредственно для часов (сдвиги в показаниях прецизионных атомных часов при кругосветных полетах на самолетах) [159].

Это, кстати, тот самый эффект, который превращает ОТО из абстрактной теории в реально работающий инструмент. И, надо сказать, очень неплохо работающий. Глобальные спутниковые навигационные системы типа GPS, активно используемые моряками, военными, спасателями и туристами, включают в себя сверхточные часы. На точность их хода влияют и скорость спутника (эффект СТО), и гравитационное поле Земли (эффект ОТО), поэтому поправки на все эти эффекты закладываются в программы обсчета сигналов, и летающие часы периодически и дозированно “замедляют”, согласуя их показания с часами на земной поверхности. За один оборот спутника вокруг Земли набирается такая разность хода часов, пренебрежение которой ведет к ошибке в 50-100 метров при определении координат наземного приемника, при том что точность навигации — до метра или чуть больше.

Эффект отклонения света (и радиоволн) многократно и с большой точностью перепроверен и стал основой теории гравитационного линзирования — главного метода обнаружения сгустков скрытого вещества во Вселенной.

Еще одно подтверждение ОТО — измерение задержки элек-

тромагнитного сигнала в поле Солнца (эффекта Шаapiro, иногда называемого четвертым классическим эффектом ОТО, в дополнение к смещению перигелия, отклонению света и красному смещению). Задержка обусловлена не замедлением скорости сигналов (скорость света всюду одинакова, так как в малой окрестности каждой точки выполнена СТО), а удлинением их пути по сравнению с расчетной траекторией в плоском пространстве. Опыты состояли в активной радиолокации космических аппаратов, а наиболее точные данные получены с помощью орбитального и спускаемого аппаратов “Викинг”, запущенных на Марс (1977) [275]. Эффект составил около 250 микросекунд, тогда как сами сигналы путешествовали в межпланетном пространстве около часа.

Эффекты ОТО в Солнечной системе представляют собой малые поправки к предсказаниям классической физики, проверены они с точностью до десятых долей процента, и в ряде современных проектов планируется эту точность улучшить на несколько порядков, достигая второго постньютоновского приближения теории гравитации.

Гораздо более интересных явлений следует ожидать в сильных гравитационных полях. Астрономические наблюдения вместе с техническим прогрессом несут все больше новой информации о проявлениях сильных полей тяготения во Вселенной. Так, нейтронные звезды, существование которых предсказано еще в 30-х годах Оппенгеймером и Волковым [241] (кстати, на основе уравнений ОТО), были открыты в 1967 году в виде радиопульсаров [166]. Бурное развитие астрономии и физики пульсаров привело, помимо прочего, и к новым проверкам ОТО. Пульсары — сверхплотные объекты с массами порядка солнечной и размерами в несколько километров — нередко находятся в двойных системах, порой довольно тесных. Там “работают” гравитационные поля в сотни и тысячи раз сильнее, чем в Солнечной системе! Высокая стабильность пульсарных “часов” дает возможность неплохо отслеживать небесную механику такой двойной

системы. Так пришло подтверждение выводов ОТО о вековых сдвигах перицентров (аналогов перигелия) в двойных системах.

Еще более весомый вклад наблюдений за пульсарами в теорию гравитации связан с гравитационными волнами. Они, как известно, пока непосредственно не обнаружены, но наблюдения за двойными пульсарами недвусмысленно наводят на их след: медленное убывание периода орбитального движения двойной системы говорит о потере энергии, и, как показывает расчет по формулам ОТО, такая потеря с разумной точностью объяснима затратами энергии на излучение именно гравитационных волн. Сегодня мало кто сомневается в их реальности — наоборот, говорят о скором становлении гравитационно-волновой астрономии, важного источника информации о бурных событиях, происходящих во Вселенной. С другой стороны, полученный результат сразу отсекает много альтернативных теорий, которые предсказывают те же, что и ОТО, эффекты в Солнечной системе, но более быструю потерю энергии системой двойной звезды на излучение гравитационных волн.

Самое известное и самое экзотическое предсказание ОТО — это, несомненно, черные дыры. Представление о гигантских ненасытных пастях, которые пожирают все подряд и ничего не возвращают назад, стало, без преувеличения, частью человеческой культуры, от фантастической литературы до фольклора. Современная астрофизика рассматривает черные дыры как вполне реальные космические объекты, возникающие в результате гравитационного коллапса тяжелых звезд, а существованием сверхмассивных черных дыр (до миллиардов солнечных масс) в центрах галактик и квазаров объясняются многие наблюдаемые в них явления.

Интересно, что уже первое точное решение уравнений Эйнштейна — решение Шварцшильда [272], характеризующее статическое поле тяготеющего центра, — содержит в себе описание простейшей черной дыры. Впрочем, полное понимание свойств решения Шварцшильда было достигнуто лишь в 50-х годах [204,

287], а некоторые и сегодня продолжают спорить о его особенностях.

С начала 60-х физика черных дыр развивается как самостоятельное научное направление, которое уже привело к ряду интересных и неожиданных результатов. Так, выяснилось, что, используя вращающиеся черные дыры, можно добывать энергию, запуская в их окрестность пробные тела по определенным траекториям (вращение при этом, естественно, замедляется) [245]; что черные дыры можно с полным правом рассматривать как термодинамические объекты с определенной температурой и энтропией [62]; что черные дыры “испаряются”, излучая энергию в окружающее пространство точно так же, как обычные тела, нагретые до соответствующей температуры [162]. Процесс хокинговского испарения связан с квантовым рождением частиц в классическом гравитационном поле дыры. Испарение дыр с массами порядка звездной и выше идет крайне медленно и не влияет ни на какие наблюдаемые процессы (в реальных условиях дыры, наоборот, наращивают свою массу за счет падающего на них вещества). Но этот процесс крайне важен для мелких черных дыр, которые могли сохраниться с начальных этапов расширения Вселенной. У них интенсивность испарения постепенно нарастает (так как температура черной дыры растет с уменьшением массы) и заканчивается взрывом.

В главе 3 этой книги мы вводим основные понятия и кратко обсуждаем некоторые вопросы физики черных дыр. Подробное и систематическое изложение этой области физики можно найти, например, в книгах [26, 299].

Есть и другое, не менее интригующее следствие описания гравитационного поля с помощью кривизны пространства-времени. А именно, если оно в принципе кривое, естественно предположить, что при каких-то условиях искривление будет очень сильным — например, приведет к чему-то, похожему на узкие перемычки между различными слабо искривленными вселенными или на “ручки”, соединяющие удаленные области одной и той

же вселенной. Такие образования получили название кротовых нор. Если они реально существуют, то, по крайней мере в принципе, возможны построение машины времени и сверхбыстрые межзвездные путешествия. Всч это с завидной регулярностью проделывают герои фантастических произведений, но, что более удивительно, свойства кротовых нор и условия их существования широко обсуждаются в специальной физической литературе — см., напр., книгу [297] и обзор [206]. Мы кратко коснемся этого круга вопросов в главе 4.

Осторожно, не квантовать!

Как мы уже отмечали, эксперимент пока полностью на стороне ОТО. В теории, однако, картина не столь безоблачная. О проблеме энергии гравитационного поля уже упоминалось. Другая известная трудность ОТО — существование сингулярностей, которые возникают в большинстве точных решений уравнений ОТО и, в частности, скрываются за горизонтами черных дыр, возникают в начале и, в ряде моделей, в конце эволюции Вселенной, а также при построения моделей изолированных тел. Это, упрощенно говоря, точки, линии или поверхности, в которых пространство-время теряет гладкость, а величины, характеризующие его кривизну, обращаются в бесконечность. Сингулярности могут быть связаны с бесконечными плотностями и давлениями материи, но встречаются и чисто геометрические сингулярности, например, в решениях уравнений Эйнштейна в вакууме — при полном отсутствии материи. Неизбежность сингулярностей в решениях ОТО при очень общих условиях доказана в целом ряде теорем, а это недвусмысленно указывает, что ОТО, по-видимому, не совсем точна при описании сверхсильных гравитационных полей.

В отличие, например, от горизонта событий, то есть границы черной дыры (выделенной в пространстве, но вполне регулярной поверхности, работающей по замечательному принципу “всех впускать, никого не выпускать”), сингулярности представ-

ляют для теории реальную проблему, так как, исходя из самой теории, указывают границы ее применимости, места, где она перестает работать. Так сама ОТО подсказывает необходимость выхода за ее собственные рамки. Как именно это надо делать — важнейший вопрос, предмет многих исследований и споров, вопрос, пожалуй, выходящий за рамки не только теории гравитации, но и физики в целом.

Кажется естественным, например, попытаться учесть квантовые явления.

Взаимоотношения гравитации и квантовой теории — отдельная, длинная и крайне запутанная история. С одной стороны, как любое волновое поле, гравитация должна на малых масштабах проявлять квантовые свойства. С другой — выясняется, что математические процедуры квантования, хорошо проявившие себя в теории частиц, с кривизной пространства справляются плохо. Есть несколько способов получения квантовых версий ОТО, которые приводят к принципиально разным результатам. По этой причине многие полагают, что квантовая гравитация должна строиться не на основе ОТО, а на основе более общей и более глубокой теории, объединяющей гравитацию с другими взаимодействиями. Альберт Эйнштейн, как известно, был многолетним оппонентом Нильса Бора по части интерпретации квантовой механики, и, вполне возможно, неприятие Эйнштейном вероятностной картины мира опиралось и на несовместимость ее с ОТО.

Некоторые качественные соображения подсказывают, чего следует ждать от квантовой гравитации. Так, из трех фундаментальных констант — постоянной Планка \hbar (квант действия, неотъемлемый атрибут квантовой теории), скорости света c (столь же фундаментальная величина СТО и ОТО, максимальная скорость материальных тел и передачи взаимодействий) и гравитационной постоянной G (она одна и та же в теории Ньютона и в ОТО) можно составить величину размерности длины. Она называется планковской длиной (l_P) и равна приблизительно 10^{-33} см — именно это, по-видимому, характерный масштаб,

на котором пространство-время начинает проявлять квантовые свойства. Это необычайно малая величина: достаточно сказать, что она меньше размеров атомного ядра примерно во столько же раз (порядок величины 10^{-13} см), во сколько само ядро меньше земного шара ($\sim 10^7$ см).

На таких длинах должны быть значительные флуктуации метрики и даже топологии — соответственно, ожидается пенообразная структура пространства, которая непрерывно меняется, “дышит” в темпе планковского времени l_{Pl}/c . При еще меньших длинах само понятие гладкого пространства-времени становится неприменимым, вместо него, возможно, имеет место какая-то дискретная (точечная) структура. Все это, разумеется, не более чем гипотезы, хотя есть многочисленные попытки математических расчетов и наблюдательных предсказаний на их основе.

Зоопарк теорий

Чтобы в общих чертах понять, чем сегодня заняты гравитационисты, достаточно полистать программу какой-нибудь крупной гравитационной конференции последних лет. Пожалуй, около трети представленных работ окажется в области классической ОТО, ее астрофизических и космологических приложений. Совершенствуется математический аппарат, включая методы поиска решений уравнений Эйнштейна, находят новые решения и анализируются старые, обсуждаются принципиальные вопросы и рассчитываются наблюдаемые эффекты. В экспериментальном разделе много работ по попыткам регистрации гравитационных волн и предложений о проведении измерений в космосе. Есть раздел альтернативных теорий гравитации, среди которых на почетном месте многомерные теории и теории объединения взаимодействий, включая гравитацию. (Заметим, что уже само слово “альтернативные” — имеется в виду по отношению к ОТО — свидетельствует об особом месте, которое по-прежнему занимает ОТО среди гравитационных теорий.) И непременно есть квантовая секция.

Разработчики обобщений ОТО преследуют достаточно разнообразные цели. Это и попытки преодолеть ее трудности, сохранив или усилив достоинства; это и стремление учесть принципы и явления, в ОТО не представленные. Но, пожалуй, главное во всех новых теориях — это подход к гравитации как к составной части будущей “теории всего на свете”. Объединенные модели гравитации, как правило, используют более сложные геометрические структуры, чем четырехмерная риманова геометрия, и новые физические поля помимо метрики. Многие из них основаны на идеях, выдвинутых еще в начале двадцатых годов. Каждая из таких теорий при наложении некоторых ограничений сводится к ОТО. Как и в ОТО, в них ведется поиск решений, представляющих физический интерес (черные дыры, космологические модели и так далее), и наблюдательных предсказаний.

Приведем несколько примеров, далеко не исчерпывающих все разнообразие подходов.

Скалярно-тензорные теории (СТТ). В соответствии с названием, гравитация в них характеризуется, кроме метрического тензора, от которого зависит кривизна пространства, одним или несколькими скалярными полями. Приведем лагранжиан общей СТТ Бергмана — Вагонера — Нордтведта с одним скалярным полем [66, 235, 298]:

$$L_{\text{STT}} = f(\phi)R + h(\phi)g^{\mu\nu}\phi_{,\mu}\phi_{,\nu} - 2U(\phi) + L_m, \quad (1.2)$$

где R — скалярная кривизна, ϕ — скалярное поле, f , h , U — произвольные функции ϕ , L_m — лагранжиан остальной материи. СТТ — самое простое с математической точки зрения обобщение ОТО, предсказывающее в общем случае зависимость гравитационной постоянной от скалярного поля и, следовательно, от положения в пространстве и времени, отличные от ОТО величины классических эффектов и большее разнообразие гравитационных волн. Лагранжиан (1.2) сводится к лагранжиану ОТО $L_{\text{GR}} = (R - 2\Lambda)/(16\pi G)$, если поле ϕ постоянно во всем

пространстве-времени. (Здесь G — гравитационная постоянная, Λ — космологическая постоянная.)

Данные наблюдений сильно ограничивают выбор допустимых СТТ. Не касаясь мотивации появления первых СТТ в работах Йордана, Бранса, Дикке в 50-х - 60-х годах [73, 176], заметим, что многочисленные скалярные поля естественно появляются при редукции различных многомерных теорий, следующих из современных моделей объединения, к четырехмерной формулировке.

Теории, нелинейные по кривизне. Другой важный класс обобщений ОТО — теории гравитации, лагранжианы которых включают, помимо скалярной кривизны R , какие-либо функции от R (так называемые $f(R)$ -теории), а также другие инварианты кривизны. Отличительный признак таких теорий — то, что в них уравнения гравитации включают производные от метрического тензора до четвертого порядка, тогда как в уравнениях Эйнштейна этот порядок не выше второго. В результате заметно увеличивается множество решений уравнений гравитации (правда, вместе с трудностью их нахождения).

Следует заметить, что появление нелинейных по кривизне членов в гравитационном лагранжиане является прямым следствием квантовой теории поля в искривленном пространстве-времени, так как такие члены неизбежно возникают в результате процедур регуляризации и перенормировки, необходимых для придания смысла результатам расчетов [7, 68]. Иначе говоря, любые теории, не учитывающие вклада нелинейных по кривизне членов, следует считать приближенными и способными описывать реальность лишь при малых значениях кривизны — разумеется, в той мере, в которой заслуживают доверия методы квантовой теории поля в искривленном пространстве-времени.

Калибровочные теории. Главная идея калибровочных теорий гравитации восходит к работам Г. Вейля 1918-1922 годов, в которых предлагалось использовать уравнения гравитации и электромагнетизма с более богатым набором симметрий, чем у урав-

нений Эйнштейна. Дополнительная симметрия включает некоторые преобразования самих полей. Начиная с 50-х годов (работы Янга и Миллса), подобные симметрии, называемые локальными калибровочными, широко используются для описания взаимодействий частиц. На этом пути достигнуто объединение слабого ядерного взаимодействия, ответственного за бета-распад, с хорошо нам знакомым электромагнитным (Вайнберг, Салам, Глэшоу). Есть перспективные варианты, объединяющие полученное “электрослабое” взаимодействие с сильным — в этом случае говорят о “теориях великого объединения”. Но это все пока без гравитации. Важно, что калибровочные симметрии могут описываться в терминах геометрии некоторых особых (расслоенных) пространств, продолжая таким образом геометризацию физики — см., напр., книги [18, 127].

Попытки включить гравитацию в общую схему калибровочных теорий привели к различным вариантам теории гравитации с кручением, которое, наряду с кривизной, является одним из видов “искажения” геометрии плоского пространства (теория Эйнштейна — Картана, пуанкаре-калибровочная теория и др. [14, 28, 34]). В таких теориях, как выясняется, можно избавиться от многих видов сингулярностей, имеющихся в решениях ОТО, и по-новому поставить проблемы энергии и квантования.

Многомерные теории, модели суперструнного происхождения. Еще одна идея геометризации физики опирается на возможность существования дополнительных измерений пространства-времени. Она берет начало от пионерских работ Т. Калуцы и О. Клейна (1921), в которых предпринимались попытки объединить гравитацию и электромагнетизм, а в последние десятилетия стала необходимой составляющей практически всех попыток объединения всех четырех физических взаимодействий.

Среди претендентов на роль “теории всего на свете” (Theory of Everything) в настоящий момент считаются наиболее перспективными так называемые теории суперструн [15, 153]. Струны —

одномерные микрообъекты, которые, как и, скажем, гитарные струны, могут испытывать колебания с определенным спектром частот. Этим частотам сопоставляются энергии различных частиц. Приставка “супер” в данном случае означает присутствие так называемой суперсимметрии — симметрии между бозонами (частицами с целым спином) и фермионами (частицами с полуцелым спином). В силу суперсимметрии у каждого бозона должен быть “симметричный” ему фермион. Суперструны “живут” в искривленных пространствах 10 или 11 измерений (в зависимости от конкретного варианта теории) и при определенных условиях приводят к некоему 10- или 11-мерному подобию ОТО. Соответственно, классическая ОТО получается при выделении четырех “обычных” координат в таких пространствах и наследует от своего многомерного прототипа солидный набор разнообразных полей.

То яма, то канава. Скажем несколько слов об одном “модном” направлении в теории объединения суперструнного происхождения — концепции “мира на бране”.

Все теории, сформулированные в пространстве-времени размерности выше четырех, вынуждены отвечать на вопрос о том, почему дополнительные измерения до сих пор не наблюдаются. В большинстве случаев, начиная с работ Калуцы и Клейна, ответ звучит так: дополнительные измерения замкнуты, или компактны, и имеют крайне малые размеры. Трубку (двумерную поверхность) диаметром в микрометр мы без микроскопа не отличим от нити, одномерного объекта. Так и с пятым или десятым измерением — пока наши приборы не добрались до достаточно малых размеров, или, с точки зрения физики частиц, до достаточно высоких энергий, мир для нас четырехмерен.

Но возможен и другой ответ: пятое измерение не мало, может быть, даже бесконечно, но мы сами и все наблюдаемые частицы и поля “заперты” на одной четырехмерной поверхности или в тонком слое, а для выхода во внешний объем нужна огромная энергия. То есть мы как бы сидим на дне глубокой ямы и

не можем вылезти. Яма протяженная — три пространственных измерения, так что скорее это вовсе не яма, а канава: вдоль — сколько угодно, а вверх и вбок — силенок не хватает.

Такая выделенная поверхность получила название браны (brane), по аналогии с всем известными мембранами, а вся концепция известна как “мир на бране” (brane world). Первые подобные модели были предложены в 80-е годы [45, 46, 243, 262] и стали весьма популярными в последние несколько лет. Выяснилось, что они приводят к новым подходам в целом ряде проблем физики элементарных частиц и космологии. С точки зрения теории гравитации интересно то, что уравнения поля тяготения на бране (где мы только и можем его наблюдать) сложнее уравнений Эйнштейна и приводят к несколько иным предсказаниям. Так, закон тяготения Ньютона должен нарушаться на расстояниях менее долей миллиметра: в знаменателе вместо квадрата радиуса появится куб или более высокая степень (см. обзор [92]). Открытие такой модификации закона тяготения стало бы весомым аргументом в пользу реальности дополнительных измерений. В мире на бране могут быть несингулярные черные дыры [83], кротовые норы без экзотической материи [102] и другие нестандартные явления.

Кажется, пора остановиться и резюмировать.

Итак, несмотря на блестящий экспериментальный статус ОТО, большинство специалистов сегодня рассматривают ее не как последнее слово в этой области физики, а как низкоэнергетический предел пока неизвестной фундаментальной теории, скорее всего многомерной, объединяющей все взаимодействия и свободной от таких трудностей, как сингулярности, проблема энергии и неоднозначность квантования.

Среди экспериментальных предсказаний альтернативных теорий гравитации уже упоминалось различие свойств гравитационных волн в разных теориях. Есть целый ряд других “новшества” по сравнению с ОТО, и, пожалуй, наиболее важное из них, общее для большинства новых теорий — предсказание пе-

ременности тех величин, которые пока считаются фундаментальными физическими константами (ФФК). В первую очередь это гравитационная постоянная G , которая входит как в закон всемирного тяготения Ньютона, так и в уравнения Эйнштейна. Переменными могут быть заряд электрона, массы элементарных частиц и многие другие константы, правда характерные времена их изменения должны быть сравнимы с временем существования Вселенной или значительно его превышать — иначе такая переменность была бы давно обнаружена. В настоящее время эксперимент и наблюдения дают лишь верхние пределы возможных вариаций ФФК. Есть, пожалуй, лишь два исключения. Во-первых, начиная с 70-х годов, время от времени появляются сообщения об обнаружении переменности G на уровне порядка 10^{-12} в год, но достоверных результатов, подтвержденных различными группами, пока нет. Во-вторых, наблюдения оптических спектров квазаров, проведенные в последние годы, говорят с некоторой вероятностью о медленном изменении постоянной тонкой структуры $e^2/(\hbar c)$, где e — заряд электрона [91]. Подробно о ситуации с переменностью ФФК можно прочесть в книгах и обзорах [16, 17, 32, 293, 308].

Гравитация и Вселенная

Важнейшая область приложения гравитационной теории — космология, наука о Вселенной в целом или о ее части, доступной для наблюдений. Современная космология — бурно развивающаяся область знаний: стремительно растет объем наблюдательных данных, обсуждается множество моделей. Некоторых из них мы будем подробно обсуждать в этой книге.

Ранее Вселенная рассматривалась как некий “сосуд”, содержащий разнообразные объекты — частицы, планеты, звезды и т.д. Казалось, нет никакой связи между свойствами этого “сосуда” и входящими в него объектами. Ситуация стала меняться после открытия А.А. Фридмана, обнаружившего, что стационарное состояние Вселенной неустойчиво, и она должна либо рас-

ширяться, либо сжиматься. Это был недвусмысленный вывод из анализа уравнений общей теории относительности Эйнштейна. Скорость расширения или сжатия Вселенной оказалась зависящей от средней плотности вещества и от его свойств. Свойства “ящика” оказались зависящими от его содержимого. Дальнейшие исследования привели к выводу о том, что наблюдаемая сейчас область Вселенной около 14 млрд лет назад имела размер порядка 10^{-27} см или меньше. А это на 19 порядков меньше размера атома. Понятно, что в такой маленькой области не могло существовать все обилие частиц, составляющих звезды. Следовательно, Вселенная и частицы рождались одновременно или почти одновременно. И, конечно же, влияли друг на друга. Сейчас уже ясно, что наша Вселенная — это не сосуд, в котором могло бы находиться любое содержимое, а сложный организм, все части которого переплетены и взаимообусловлены. Ведь здесь важны и свойства элементарных частиц, и общая теория относительности (ОТО), и статистическая физика, причем зачастую все одновременно.

Принципиальной особенностью современной эпохи является становление космологии как науки. Сейчас мы можем аргументированно отвечать на множество вопросов, объяснять наблюдаемые данные и теоретически предсказывать новые эффекты. Становится понятным общий ход эволюции Вселенной, начиная с момента ее рождения и далее в будущее. Конечно, существенные детали еще не ясны, но практически все возможные варианты эволюции нашей Вселенной уже названы.

Гравитация играет ключевую роль не только в объяснении свойств локальных объектов — звезд, галактик, черных дыр — но и Вселенной в целом. Так, размер нашей Вселенной, а точнее размер ее видимой части, однозначно зависит от гравитационной постоянной. Постепенно устанавливаются контуры Стандартной космологической модели (СКМ), которая по своей значимости и фундаментальности начинает приближаться к Стандартной модели физики элементарных частиц.

СКМ на базе ОТО объясняет прежде всего следующие основные феномены:

1. Однородность и изотропию Вселенной.
2. Флуктуации температуры реликтового излучения.
3. Первичный нуклеосинтез.

Можно считать, что история космологии, науки — изучающей Вселенную в целом, началась в 1920 г. с открытия Э. Хаббла. Он обнаружил, что многочисленные туманности, выглядевшие на фотографиях как маленькие невнятные пятна, представляют собой галактики, расположенные в миллионах световых лет от нашей Галактики Млечный путь. К тому же оказалось, что эти галактики разлетаются, свидетельствуя о некоем моменте рождения Вселенной. Дальнейшие исследования подтвердили эту догадку. Избыток гелия во Вселенной и наличие микроволнового реликтового излучения, открытого в 1965 г., указывало также на то, что первоначально Вселенная состояла из сильно нагретой плазмы. В этом суть идеи Большого взрыва.

В среднем Вселенная в высокой степени однородна, что труднообъяснимо само по себе, но крупномасштабные структуры типа галактик могли образоваться лишь при наличии первичных флуктуаций плотности энергии. Детальный анализ показал, что одних только флуктуаций барионов недостаточно, чтобы образовалась современная структура галактик. Должна существовать “невидимая” составляющая плотности энергии, чтобы гравитационные силы были в состоянии сжать область повышенной плотности — так и началось звездообразование. Существование темной материи действительно практически доказано, но до сих пор остается неясным ее состав. Ясно только, что по массе ее примерно в 5 раз больше нежели барионов — частиц, образующих звезды.

В 1998 г. было сделано фундаментальное открытие. Было обнаружено, что расширение Вселенной происходит не с замедлением, как полагали раньше, а с ускорением. В рамках ОТО это можно объяснить только тем, что основная плотность энер-

гии заключена не в барионах и даже не в темной материи, а в некоторой странной материи, обладающей большим отрицательным давлением. Ее носителем вообще не могут быть частицы — иначе она бы образовывала неоднородные структуры, что не соответствует наблюдениям. Этот странный вид материи, получивший название “темной энергии”, распределен в пространстве удивительно однородно, но ее присутствие сказывается на скорости разбегания галактик. Происхождение темной энергии неясно. Самое простое объяснение — минимум потенциальной энергии некоего гипотетического поля не равен строго нулю, а представляет собой малую положительную постоянную. Само по себе это не очень удивительно, странно то, что этот минимум чрезвычайно близок к нулю. Пока не существует общепринятой теории, объясняющей подобную малость.

На данный момент остается много открытых вопросов, но есть и значительные продвижения. Особенно впечатляющий шаг вперед произошел с появлением идеи инфляции — сверхбыстрого расширения пространства на начальной стадии эволюции Вселенной. Согласно инфляционному сценарию, сразу после инфляции произошел резкий нагрев вещества, воспроизводя известную картину горячей Вселенной, которая рассматривалась в доинфляционных моделях как начальный этап. Конкретных реализаций инфляции, моделей, существует великое множество, и дело будущего — выбрать правильную модель. Объединяет их лишь то, что все они предсказывают сверхбыстрое расширение пространства в рамках той или иной теории гравитации.

Следует иметь в виду, что космологические модели самым существенным образом зависят от выбора теории гравитации, и поэтому можно надеяться, что по мере дальнейшего накопления знаний космология поможет нам не только уяснить судьбу Вселенной, но и лучше понять механизмы гравитации.

Глава 2

Основные положения ОТО

Основы ОТО прекрасно изложены в целом ряде известных учебников и монографий (например, [23, 227, 288, 299, 300]). Предполагая знакомство читателя с основными положениями теории, в данной главе мы, прежде всего в справочных целях, приведем лишь некоторые ее основные факты и соотношения. Многие геометрические понятия (например, понятия вектора и тензора, ко- и контравариантных компонент векторов и тензоров, свертки тензоров и т.д.) будем считать известными и использовать без объяснений.

Пространство-время в ОТО представляет собой четырехмерное дифференцируемое многообразие с псевдоримановой метрикой (часто опускают приставку “псевдо”, говоря просто “риманово пространство-время”). Гравитационное поле в ОТО описывается в терминах кривизны пространства-времени, которая выражается через метрический тензор и его производные по координатам. Таким образом, ОТО относится к классу метрических теорий гравитации [308]; ОТО — исторически первая, самая простая и наиболее разработанная теория из этого класса.

В малой окрестности произвольной точки (события) пространство ОТО — псевдориманово пространство-время — почти (на уровне первых производных метрического тензора) совпадает со своим касательным пространством — плоским простран-

ством Минковского, и, следовательно, в каждой такой малой области приближенно справедлива специальная теория относительности (СТО). Поэтому начнем изложение с напоминания некоторых основных фактов из СТО.

2.1. СТО. Геометрия Минковского

2.1.1. Геометрия

В пространстве Минковского основной геометрический инвариант — интервал между двумя произвольными событиями. Существует класс привилегированных координат (координат Минковского) x^μ , каждый выбор которых соответствует некоторой инерциальной системе отсчета (СО). Суммарная сила, действующая на тело, покоящееся в какой-либо инерциальной СО (ИСО), равна нулю, иными словами, такое тело движется по инерции, равномерно и прямолинейно относительно любой другой ИСО. В любой ИСО квадрат четырехмерного “расстояния” (интервала) между событиями 1 [$x_1^\mu = (ct_1, \vec{x}_1)$] и 2 [$x_2^\mu = (ct_2, \vec{x}_2)$] записывается в виде

$$s^2(1,2) = c^2(t_2 - t_1)^2 - (\vec{x}_2 - \vec{x}_1)^2, \quad (2.1)$$

где c — универсальная постоянная, совпадающая со скоростью распространения электромагнитных волн (в частности, света) в вакууме и называемая скоростью света. Для близких событий 1 и 2 интервал (2.1) можно записать в виде

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu, \quad \mu = 0, 1, 2, 3, \quad (2.2)$$

где тензор с ковариантными компонентами

$$\eta_{\mu\nu} = \text{diag}(1, -1, -1, -1) \quad (2.3)$$

называется метрическим тензором (метрикой) Минковского. Матрица (2.3) вместе с обратной матрицей

$$\eta^{\mu\nu} = \text{diag}(1, -1, -1, -1) \quad (2.4)$$

контравариантных компонент тензора Минковского используются для поднятия и опускания векторных и тензорных индексов, так что, например, для произвольного вектора $a = (a^\mu)$ $a^\mu = \eta^{\mu\nu} a_\nu$, $a_\mu = \eta_{\mu\nu} a^\nu$. Тензор Минковского определяет скалярное произведение (ab) двух произвольных 4-векторов a^μ и b^μ по формуле

$$(ab) = \eta_{\mu\nu} a^\mu b^\nu = \eta^{\mu\nu} a_\mu b_\nu = a_\mu b^\mu = a^\mu b_\mu. \quad (2.5)$$

В частности, квадрат интервала есть скалярный квадрат вектора $x_1^\mu - x_2^\mu$.

Очевидно, скалярный квадрат $a^2 = a_\mu a^\mu = g_{\mu\nu} a^\mu a^\nu$ любого 4-вектора a может принимать положительное, отрицательное или нулевое значение. В случае $a^2 > 0$ 4-вектор a называется временноподобным (английский термин — timelike), в случае $a^2 < 0$ — пространственноподобным (spacelike), в случае $a^2 = 0$ — светоподобным или световым (null, light, lightlike). Эти понятия относятся и к векторам $x_1 - x_2$, связывающим события 1 и 2, и к интервалам между событиями.¹

Зафиксируем произвольную точку O и поместим в нее начало координат некоторой инерциальной СО. Тогда совокупность

¹Несколько замечаний о терминологии. Вместо “временноподобный” иногда говорят и пишут “временноподобный”. Это не вносит путаницы, но нарушает единство терминологии (тогда почему не говорить “пространствуподобный”?).

Хуже обстоит дело со светоподобными векторами: с чьей-то нелегкой руки такие векторы (а также соответствующие направления, кривые, поверхности и т.д.) в русскоязычной литературе нередко называют “изотропными”. Это, во-первых, противостоит естественному (любой ненулевой вектор выделяет некоторое направление, а термин “изотропный” в своем общепринятом основном значении определяется как “не зависящий от направления” или “одинаковый во всех направлениях”): “изотропный вектор” — это примерно как “горячий лед” или “неимуций богач”. Во-вторых, такое употребление может просто сбивать с толку: так, изотропная поверхность в нормальном понимании есть поверхность, свойства которой одинаковы во всех направлениях. Но, по описанной выше “кривой” терминологии это может быть и светоподобная поверхность, т.е. поверхность, содержащая световые направления, значит, уж точно анизотропная... Предлагаем всемерно избегать подобной путаницы и, в частности, неправомерного употребления слова “изотропный”.

концов всех световых векторов x^μ с началом в O образуют коническую поверхность, называемую *световым конусом* (точнее — световым конусом точки O). Его уравнение следует из (2.1):

$$(x^0)^2 - \vec{x}^2 \equiv c^2 t^2 - x^2 - y^2 - z^2 = 0 \quad (2.6)$$

(где, как обычно, $x^0 = ct$, $x^1 = x$, $x^2 = y$, $x^3 = z$, $\vec{x}^2 = x^2 + y^2 + z^2$ — трехмерный евклидов скалярный квадрат). Все временноподобные прямые, проходящие через точку O , находятся внутри светового конуса (2.6), пространственноподобные — вне его.

Траектории частиц ненулевой массы покоя временноподобны на всем своем протяжении. Так как скорость света в СТО — максимально возможная скорость движения материальных частиц и распространения сигналов, два события могут быть причинно связаны только если их разделяет временноподобный или световой интервал, для них имеют абсолютный смысл понятия “раньше” и “позже”. При этом одно из событий находится либо на световом конусе другого, либо внутри него. Световой конус любого события естественно делится на световой конус прошлого и световой конус будущего.

2.1.2. Преобразования координат

Переход от одной ИСО к другой выглядит проще всего, если скорость \vec{v} системы S' относительно исходной системы S направлена вдоль одной из координатных осей, например, вдоль оси

Иногда, калькируя английское “null”, о световом векторе говорят “нулевой”. Это тоже, пожалуй, не совсем удачно, так как по общепринятым геометрическим понятиям нулевой (zero) вектор есть вектор, у которого все компоненты равны нулю.

Еще один пример неудачной терминологии — названия “галилеева система координат” и даже “галилеева метрика”, иногда используемые для обозначения, соответственно, координат и метрики Минковского. Дело в том, что в пространстве Минковского справедлив принцип относительности Эйнштейна (частный), связанный с независимостью законов физики от выбора ИСО, а принцип относительности Галилея выполняется лишь в пределе малых скоростей в фиксированной ИСО.

x , то есть в системе S начало координат системы S' движется по закону $x = vt$. В этом случае преобразование координат (специальное преобразование Лоренца), оставляющее инвариантным интервал (2.1), имеет вид

$$\begin{aligned}x' &= \frac{x - vt}{\sqrt{1 - v^2/c^2}}, & y' &= y, & z' &= z, \\t' &= \frac{t - vx/c^2}{\sqrt{1 - v^2/c^2}},\end{aligned}\quad (2.7)$$

где штрихованные координаты относятся к ИСО S' .

Общее преобразование Лоренца, связывающее две произвольные ИСО, включает в себя буст (переход типа (2.7) от одной СО к другой, но со скоростью \vec{v} в произвольном направлении) и произвольный поворот пространственных координатных осей. Все преобразования Лоренца образуют 6-параметрическую группу, называемую группой Лоренца. Кроме общих преобразований Лоренца, интервал (2.1) инвариантен относительно пространственных и временных сдвигов (трансляций)

$$x'^{\mu} = x^{\mu} + a^{\mu}, \quad a^{\mu} = \text{const}, \quad (2.8)$$

образующих 4-параметрическую группу трансляций. Таким образом, полная группа *изометрий* (преобразований координат, оставляющих инвариантным вид метрического тензора) содержит 10 параметров. Она носит название группы Пуанкаре.

Матрица произвольного преобразования Лоренца $A = (A_{\mu}^{\nu})$ обладает определяющим свойством псевдоортогональности, которое и выражает инвариантность метрики Минковского относительно этих преобразований. А именно, пусть координаты x^{μ} системы S и координаты y^{μ} системы S' связаны линейным преобразованием

$$x^{\mu} = A_{\alpha}^{\mu} y^{\alpha} + a^{\mu}, \quad A_{\alpha}^{\mu}, a^{\mu} = \text{const}. \quad (2.9)$$

Согласно (2.9), $A_{\alpha}^{\mu} = \partial x^{\mu} / \partial y^{\alpha}$. Подставляя (2.9) в выражение для интервала (2.2), получаем

$$ds^2 = \eta_{\mu\nu} A_{\alpha}^{\mu} A_{\beta}^{\nu} dy^{\alpha} dy^{\beta}. \quad (2.10)$$

Инвариантность метрики $\eta_{\mu\nu}$ означает, что в правой части этого равенства стоит выражение $\eta_{\alpha\beta}dy^\alpha dy^\beta$, где матрица $\eta_{\alpha\beta}$ — та же, что и $\eta_{\mu\nu}$ и имеет вид $\text{diag}(1, -1, -1, -1)$. Условие

$$\eta_{\mu\nu}A^\mu_\alpha A^\nu_\beta = \eta_{\alpha\beta} \quad (2.11)$$

и есть условие псевдоортогональности матрицы A . Произвольные константы a^μ в преобразовании (2.9), образующие вектор трансляции, отсутствуют в условии (2.11) и играют роль констант интегрирования для условия (2.11), рассматриваемого как система уравнений относительно функций $x^\mu(y^\alpha)$.

2.1.3. Кинематические эффекты

Анализ преобразований Лоренца приводит к важнейшим кинематическим эффектам СТО. Так, произвольное движение точечной частицы в каждый фиксированный момент времени можно приближенно считать инерциальным и ввести ИСО, в которой частица в этот момент покоится. Предполагая без потери общности, что движение в этот момент происходит вдоль оси x , легко установить, что приращение времени dt' по часам, связанным с частицей (и равное ds/c), связано с приращением времени dt по часам неподвижной СО формулой

$$dt' = dt\sqrt{1 - v^2/c^2}. \quad (2.12)$$

В силу произвольности выбора осей, эта формула справедлива для произвольно направленной 3-скорости $\vec{v} = d\vec{x}/dt$. Следовательно, для произвольной траектории движения $\vec{x}(t)$ интервал собственного времени τ частицы (т.е. время, прошедшее по часам, связанным с частицей или любым объектом, размеры которого несущественны) определяется формулой

$$\tau(t_1, t_2) = \frac{1}{c} \int_{t_1}^{t_2} ds = \int_{t_1}^{t_2} dt\sqrt{1 - v^2(t)/c^2}, \quad (2.13)$$

если по неподвижным часам прошло время от t_1 до t_2 .

Из формул (2.12) и (2.13) следует, что собственное время движущегося объекта всегда меньше, чем время, прошедшее с точки зрения неподвижного наблюдателя. Это так называемое *лоренцево замедление времени*; оно приводит к известному *парадоксу близнецов*. Если один близнец неподвижен (или ограничивается медленными перемещениями) в некоторой ИСО, а второй совершает путешествие с релятивистскими скоростями и, описав замкнутую пространственную траекторию, воссоединяется с братом, то при встрече они окажутся разного возраста: путешественник будет моложе домоседа. Казалось бы, если рассматривать ситуацию в СО, в которой неподвижен путешественник, результат должен быть противоположным. Но внимательный анализ, учитывающий то, что путешественник должен был как минимум трижды переходить из одной ИСО в другую (при разгоне, повороте обратно и окончательном торможении), показывает, что и в этом случае его расчетный возраст будет меньше, чем у сидевшего на месте брата. Парадокс объясняется несимметричностью ситуации: интеграл (2.13) оказывается меньше для того объекта (или в данном случае субъекта), который совершал неинерциальное движение.

Другой эффект — *лоренцево сокращение длин*: длина линейки или стержня имеет разные значения при измерении из разных ИСО. Она максимальна в СО, в которой стержень покоится, и сокращается в отношении $\sqrt{1 - v^2/c^2}$ в СО, движущейся со скоростью \vec{v} параллельно стержню. Этот результат выводится нахождением координат концов стержня в один и тот же момент времени по часам движущейся СО. Поперечные размеры тела при движении не меняются, следовательно, объем меняется в том же отношении, что и продольная длина. Тело, движущееся с релятивистскими скоростями, уплощается, а в ультрарелятивистском пределе $v \rightarrow c$ сферическое тело приобретает форму диска или блина.

Третий эффект — относительность одновременности: в различных ИСО пространственноподобные гиперповерхности $t =$

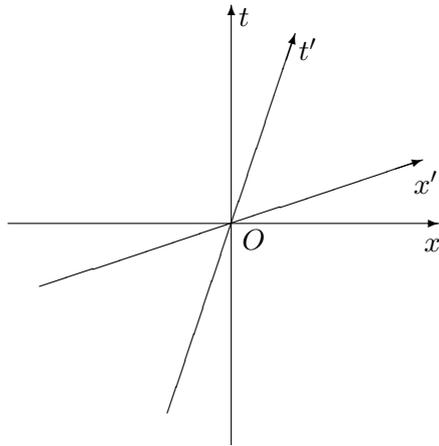


Рис. 2.1. Оси координат x и t в различных ИСО

const различны. Более того, любые два события, разделенные пространственноподобным интервалом, можно сделать одновременными путем выбора подходящей ИСО. Это обстоятельство лучше всего иллюстрируется рисунком 2.1.3, на котором показано двумерное сечение пространства Минковского по плоскости (t, x) . В координатах x, t , принадлежащих ИСО S , показаны оси x', t' другой ИСО S' ; линии $t = \text{const}$ представляют пространственные сечения (поверхности одновременности) ИСО S , линии $t' = \text{const}$ (ось x' и параллельные ей прямые) — аналогичные поверхности ИСО S' .

Есть и другие кинематические эффекты СТО, например, абберация света (несовпадение направлений на удаленный источник света при наблюдении из различных ИСО) и отличие релятивистских формул для сложения скоростей и эффекта Доплера от нерелятивистских. Их подробный анализ можно найти в многочисленных учебниках по СТО.

2.1.4. Элементы релятивистской механики точки

Для формулировки СТО в терминах четырехмерной геометрии Минковского вместо трехмерных величин, фигурирующих в нерелятивистской механике, вводятся их четырехмерные аналоги, что приводит к экономной и прозрачной записи многих уравнений. В частности, для движения материальной точки или элемента среды вводится 4-скорость ($\gamma \equiv 1/\sqrt{1 - v^2/c^2}$):

$$u^\mu = \frac{dx^\mu}{ds} = \left(\gamma, \frac{\gamma v^i}{c} \right), \quad (2.14)$$

Здесь $ds = c dt \sqrt{1 - v^2/c^2}$ — элемент интервала вдоль траектории частицы. Нетрудно убедиться, что так определенный вектор 4-скорости нормирован:

$$u_\mu u^\mu = 1. \quad (2.15)$$

4-ускорение материальной точки или элемента среды имеет вид

$$a^\mu = \frac{du^\mu}{ds} = \frac{\gamma}{c} \frac{du^\mu}{dt}. \quad (2.16)$$

В силу условия нормировки (2.15) 4-скорость и 4-ускорение взаимно ортогональны: $a_\mu u^\mu = 0$.

4-импульс материальной точки (частицы) равен, по определению,

$$p^\mu = m u^\mu = (E/c, \vec{p}), \quad m = \text{const}, \quad (2.17)$$

где m — масса покоя частицы, характеризующая ее инертные свойства, $\vec{p} = (p^i)$ — пространственный импульс и $E = \gamma m c^2$ — энергия частицы в заданной ИСО. По определению, квадрат импульса есть $p^2 = p_\mu p^\mu = m^2 c^2$.

При малых скоростях, $v \ll c$, разлагая выражение для энергии по степеням v/c , получаем

$$E \approx m c^2 + \frac{m v^2}{2} + O(v^2/c^2). \quad (2.18)$$

Отсюда следует важнейший вывод, что полная энергия частицы включает кроме классической кинетической энергии энергию покоя mc^2 .

Роль второго закона Ньютона играет равенство, которое можно рассматривать как определение 4-вектора силы f^μ :

$$f^\mu = \frac{dp^\mu}{ds} = mca^\mu = mc \frac{du^\mu}{ds}, \quad (2.19)$$

где компоненты f^μ связаны с обычным трехмерным вектором силы \vec{F} соотношениями

$$f^0 = \frac{\gamma}{c^2}(\vec{F}\vec{v}), \quad f^i = \frac{\gamma}{c}F^i. \quad (2.20)$$

Действие свободной частицы записывается в виде [23]

$$S = -mc \int ds = -mc^2 \int \sqrt{1 - \frac{v^2}{c^2}} dt, \quad (2.21)$$

и так как по определению $S = \int L dt$, лагранжиан L имеет вид

$$L = -mc^2 \sqrt{1 - \frac{v^2}{c^2}}. \quad (2.22)$$

Для частицы, движущейся под влиянием внешних сил, к этому действию должны добавляться соответствующие слагаемые, включающие потенциалы взаимодействия между частицами и действия различных физических полей.

Для свободной частицы вариация действия (2.21) по функциям $x^i(t)$, описывающим мировую линию частицы, равна нулю, откуда следует лагранжево уравнение

$$\frac{d}{dt} \frac{v^i}{\sqrt{1 - v^2/c^2}} = 0 \Rightarrow v^i = \text{const}, \quad (2.23)$$

т.е. получается, как и должно быть, равномерное прямолинейное движение.

Фотоны и другие безмассовые частицы движутся по световым прямолинейным траекториям и могут обладать произвольной энергией E и пространственным импульсом $p^i = (E/c^2)v^i$, так что $|\vec{p}| = E/c$. Связь между энергией фотона и частотой, $E = h\nu$, как известно, возникает не в СТО, а в квантовой механике.

В релятивистской механике справедлив принцип соответствия: механика Ньютона для массивных тел восстанавливается при скоростях $v \ll c$.

2.2. Риманово пространство-время. Системы координат и системы отсчета

До сих пор мы применяли в пространстве Минковского координаты Минковского, в которых метрика имеет вид (2.2), и связанные с этими координатами инерциальные СО. Ничто не мешает, однако, описывать физические процессы в рамках СТО, пользуясь другими системами координат, например, даже оставаясь в рамках фиксированной СО в том же пространстве Минковского, вводить сферические или цилиндрические пространственные координаты.

В самом общем случае — как в пространстве Минковского, так и в любых римановых пространствах и, шире, в любых дифференцируемых многообразиях допустимы координатные преобразования $x^\mu \mapsto y^\mu$ с произвольными функциями

$$x^\mu = x^\mu(y^0, y^1, y^2, y^3). \quad (2.24)$$

В физическом пространстве координатные преобразования (2.24) в общем случае приводят к изменению СО.

Следует заметить, что связь понятий системы координат и системы отсчета — достаточно тонкий вопрос, который нередко становится предметом путаницы и заблуждений. Поэтому целесообразно пояснить, в каком смысле мы будем употреблять эти понятия.

Обсуждение в данном разделе в основном следует книгам [4, 12, 13]. Мы будем рассматривать произвольное (псевдо)риманово пространство-время с интервалом

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu, \quad \mu, \nu = 0, 1, 2, 3, \quad (2.25)$$

с симметричным метрическим тензором $g_{\mu\nu}$, имеющим сигнатуру $(+ - - -)$. Фиксируя произвольную точку, мы можем привести в ней метрический тензор к диагональному виду с помощью линейных преобразований координат, и тогда, очевидными дополнительными преобразованиями координат упорядочивая получившиеся диагональные элементы и нормируя их на единицу, мы в силу сделанного выбора сигнатуры всегда получим тензор Минковского (2.3). Это говорит о том, что в малой окрестности каждой точки геометрия любого риманова пространства-времени совпадает с геометрией Минковского и все явления, в которых кривизна несущественна, можно описывать в рамках СТО.

2.2.1. Ковариантность, карты и атласы

Приведем общее определение системы координат в дифференцируемом многообразии. Определения всех соответствующих понятий и их строгое и подробное обсуждение можно найти в учебниках и монографиях по дифференциальной геометрии в целом, например [29, 163, 208].

Система координат в некоторой области U дифференцируемого многообразия класса k размерности D (или, что то же самое, *карта* области U) есть взаимно-однозначное отображение области U на некоторую область арифметического пространства \mathbb{R}^D . Сама область U при этом называется областью карты или системы координат.

Таким образом, каждой точке x области U ставится в соответствие упорядоченный набор из D вещественных чисел, которые и называются координатами данной точки. При этом если

U_1 и U_2 — области некоторых двух карт и $x \in U_1 \cap U_2$, то координаты точки x в одной из этих карт являются функциями класса C^k координат точки x в другой карте с ненулевым якобианом преобразования. Чаще всего рассматриваются многообразия класса C^∞ с бесконечно дифференцируемыми функциями преобразований.

Свойства многообразия в целом описываются наборами карт, области которых покрывают всё многообразие (атласами). Другими словами, объединение областей карт некоторого атласа тождественно всему многообразию.

Итак, система координат в многообразии и, в частности, в четырехмерном римановом пространстве есть просто способ приписать каждой его точке (событию) определенный “адрес” или метку в виде набора чисел, и различные способы адресации должны быть связаны друг с другом гладкими функциями.

Отсюда ясно, что координаты в пространстве-времени, вообще говоря, не несут никакого физического смысла, если, конечно, их значения не заданы специально как функции каких-либо физических величин (например, длин, промежутков времени, значений напряженности электрического поля и т.д.). В то же время уравнения физической теории, рассматривающей объекты или процессы в пространстве-времени, содержат координаты точек в качестве переменных или параметров. Но никакие свойства физических тел и явлений и никакие измеримые величины, очевидно, не должны зависеть от того, какие метки привязаны к точкам-событиям. (Так свойства дорожного покрытия и мощности моторов проезжающих автомобилей не могут зависеть от того, размечена дорога в километрах, милях, верстах или локтях и от какого пункта ведется отсчет.) Отсюда следует, что *физические законы должны допускать универсальную формулировку, не зависящую от выбора системы координат* (так называемую общековариантную запись). Это утверждение составляет содержание *принципа общей ковариантности*.

Принцип общей ковариантности есть, пожалуй, математиче-

ское выражение здравого смысла: любая претендующая на разумность физическая теория должна допускать общековариантную формулировку.

2.2.2. Системы отсчета и относительность

Перейдем к понятию *систем отсчета* (СО). По определению, СО есть воображаемое, лишенное массы, вообще говоря, произвольным (но плавным, гладким) образом деформирующееся тело — тело отсчета, существующее в некоторой области пространства-времени и снабженное во всех своих точках идеальными линейками, позволяющими измерять длины, и идеальными часами, позволяющими измерять промежутки времени.

Таким образом, в отличие от систем координат, представляющих собой чисто математическое понятие, СО — физическое понятие, необходимое для описания связи теории с измерениями. Как известно, реальными прототипами идеальных часов являются атомные часы, а прототипами идеальных линеек — действующие эталоны длины, привязанные к длинам волн определенных спектральных линий.

Общий принцип относительности утверждает эквивалентность всех СО при формулировке законов природы. Это содержательный физический принцип, из которого следует отсутствие в природе привилегированных СО. Заметим, что даже теории, в которых такие привилегированные СО имеют место, допускают запись своих уравнений в общековариантном виде. Очевидный пример — СТО с привилегированной ролью инерциальных СО (ИСО).

Отсутствие привилегированных СО при формулировке законов физики, конечно, не означает, что результаты физических измерений одинаковы во всех СО: в самом деле, измерительный прибор всегда находится в определенной СО, и это неизбежно влияет на его показания. Например, в СТО среди ИСО нет привилегированных, однако измерения длины одного и того же стержня, проводимые приборами, расположенными в разных

ИСО, дают, как мы видели, разные результаты.

2.2.3. Системы отсчета и хронометрические инварианты

Координаты в принципе можно использовать независимо от СО. Например, при описании города в его естественной СО, связанной с домами и улицами, ничто не запрещает использовать координаты, привязанные к какой-либо системе регулярно движущихся автомобилей, или к теням от плывущих по небу облаков.

Однако обычно проще и удобнее привязывать координаты к определенной СО, предполагая, что тело отсчета покрыто неподвижной относительно него трехмерной координатной сеткой x^i , а временная координата x^0 меняется вдоль мировых линий фиксированных точек тела отсчета (линий времени). В таком случае говорят, что *система координат принадлежит данной СО*. Мировые линии частиц, неподвижных в данной СО, описываются уравнением $x^i = \text{const}$ (не зависят от x^0). Если производится преобразование координат (2.24), то условие, что координаты y^μ принадлежат той же самой СО, что и x^μ , состоит в том, что новые пространственные координаты y^i не зависят от x^0 .

Таким образом, преобразования

$$y^i = y^i(x^1, x^2, x^3), \quad (2.26)$$

$$y^0 = y^0(x^0, x^1, x^2, x^3) \quad (2.27)$$

— это наиболее общие преобразования между системами координат, принадлежащими одной и той же СО. Равенство (2.26) описывает трехмерные пространственные преобразования, меняющие пространственную координатную сетку. Равенство (2.27) описывает произвольные хронометрические преобразования, меняющие ход произвольных (координатных) часов и их синхронизацию от одной пространственной точки к другой.

В любой системе координат x^μ часто рассматриваются *поверхности одновременности, или пространственные сечения*

— совокупности мировых точек, определяемых условием $x^0 = \text{const}$.

Из формулы (2.27) ясно, что условие $x^0 = \text{const}$ отнюдь не гарантирует $y^0 = \text{const}$. Иначе говоря, даже в координатах, принадлежащих одной и той же СО, пространственные сечения могут быть разными, так как на одном и том же теле отсчета могут быть разные наборы часов, по-разному синхронизованные и идущие с разной (и по-разному переменной!) скоростью.

Пример — разные временные шкалы на поверхности Земли, то есть заведомо в одной и той же СО: время по Гринвичу, время в других часовых поясах, истинное солнечное время и т.д.. А также время, отсчитываемое часами с маятниками фиксированной длины (их скорость хода зависит от силы тяжести, которая меняется от точки к точке) и атомными часами.

Измеримые физические величины в данной СО должны быть ковариантными по отношению преобразований (2.26) и инвариантными относительно преобразований (2.27) (*хронометрически инвариантными*). Это — введенный Зельмановым [12] *принцип хронометрической инвариантности*. Он диктуется теми же соображениями, что и принцип общей ковариантности, но применительно к заданной СО.

Несложно получить выражения для хронометрических инвариантов на основе компонент любых четырехмерных тензоров. Так, у любого 4-вектора A_μ можно выделить следующие хронометрически инвариантные (ХИ) составляющие:

$$A^i, \quad i = 1, 2, 3; \quad A_t = A_0/\sqrt{g_{00}}. \quad (2.28)$$

Для тензоров более высокого ранга надо поступить аналогичным образом с каждым индексом. А именно, составленные из компонент произвольного тензора $A_{\mu_1\mu_2\dots\mu_r}$ ранга r величины

$$A_{0\dots 0}^{i_1 i_2 \dots i_m} (g_{00})^{-(r-m)/2}, \quad (2.29)$$

где m меняется от нуля до r , а число нижних индексов (нулей) равно $r - m$, образуют контравариантные ХИ тензоры от-

носителем пространственных преобразований (2.26). Этот результат проверяется непосредственно. Метрический коэффициент g_{00} всегда отличен от нуля в координатах, принадлежащих какой-либо СО.

Дальнейшее развитие теории хронометрических инвариантов включает запись всех уравнений через ХИ величины и ХИ операторы дифференцирования, а также описание СО с помощью ХИ характеристик — ускорения, вращения и деформации [12, 13, 24].

Следует отметить, что иногда применяются системы координат, не принадлежащие никаким СО — например, световые координаты, и в них вполне может оказаться $g_{00} = 0$ в каких-то регулярных точках или даже во всем пространстве.

Универсальное описание СО осуществляется с помощью монадного формализма [4, 13, 25], который рассматривает хронометрические инварианты как одну из “калибровок”.

Более известен тетрадный (реперный) формализм (см., например, [24, 30, 227]), в котором СО характеризуется выбором ортонормированного репера в касательном пространстве Минковского в каждой мировой точке. Заметим, что для задания СО согласно приведенному выше определению достаточно фиксировать лишь один (временноподобный) вектор в каждой точке (т.е. монаду), а остальные три вектора тетрады, характеризующие положения пространственных осей, в этом смысле избыточны. Однако тетрадный формализм чрезвычайно полезен для расчетов наблюдаемых величин и для описания фермионов в искривленном пространстве-времени. Более того, многие авторы рассматривают именно реперы (а не метрику) в качестве динамических переменных, описывающих гравитационное поле в ОТО и ее обобщениях.

2.2.4. Ковариантность и относительность

Итак, как мы видели, принцип общей ковариантности — это всего лишь констатация возможности пользоваться любыми координатами для описания пространства-времени и происходящих

в нем явлений.

Напротив, различные виды принципа относительности — важные постулаты физических теорий. Так, принцип относительности Галилея, действующий в ньютоновой механике, утверждает независимость законов физики от выбора ИСО в ньютоновском абсолютном пространстве, где связь между ИСО осуществляется преобразованиями Галилея. Частный принцип относительности Эйнштейна утверждает в точности то же самое, но в пространстве Минковского, где ИСО связаны между собой преобразованиями Лоренца. В обоих случаях в теории выделяется класс привилегированных СО, внутри которого и происходят преобразования.

Декартовы координаты в ньютоновском (евклидовом) пространстве и координаты Минковского в пространстве Минковского носят привилегированный характер с двух точек зрения. Математически они выделены тем, что их координатные линии совпадают с орбитами группы изометрий соответствующих пространств (впрочем, этим свойством обладают также сферические и цилиндрические координаты), а физически — удобством описания ИСО с их помощью. Известно, однако, что в задачах классической механики часто используются разнообразные криволинейные координаты, а также изучаются физические явления в различных неинерциальных СО. В СТО также возникает необходимость рассматривать ускоренные СО и вводить системы координат, им принадлежащие (пример — равноускоренная СО и координаты Риндлера [258]). При этом оказываются полезными методы, разработанные в ОТО и связанные с общековариантной записью всех уравнений.

Надо сказать, что слово “относительность”, входящее в названия как СТО, так и ОТО, как бы едино в двух лицах. С одной стороны, этим словом названы принципы, говорящие о противоположном свойстве физических законов — их абсолютном характере, независимости от СО; но применять их приходится в конкретных СО, или относительно выбранного класса наблю-

дателей. С другой стороны, и в СТО, и в ОТО многие вещи, бывшие абсолютными в классической теории, становятся относительными, т.е. зависящими от СО. В СТО это одновременность, длины, промежутки времени; в ОТО (как и в других метрических теориях гравитации) к этому добавляется конечность или бесконечность пространственного объема и даже топология 3-пространства. В дальнейшем мы в этом убедимся на конкретных примерах.

2.3. Риманово пространство-время

Приведем некоторые определения и формулы, важные для следующего изложения.

Прежде всего, контравариантные компоненты метрики, $g^{\mu\nu}$, образуют матрицу, обратную матрице $g_{\mu\nu}$:

$$g_{\mu\alpha}g^{\alpha\nu} = \delta_{\mu}^{\nu}, \quad (2.30)$$

где δ_{μ}^{ν} — символ Кронекера, равный единице для совпадающих индексов и нулю для несовпадающих. Тензоры $g_{\mu\nu}$ и $g^{\mu\nu}$ используются для поднятия и опускания индексов произвольных тензоров.

Частные производные любой скалярной функции $f(x^{\mu})$ по координатам, $\partial_{\mu}f$, образуют ковариантный вектор — градиент f . Частные производные компонент вектора A_{μ} или A^{μ} в общем случае не образуют тензора в силу, вообще говоря, нелинейного характера координатных преобразований. Для ковариантной записи физических уравнений и для многих других целей необходимо обобщение понятия производной, делающее ее тензором. Это достигается введением ковариантных производных

$$\begin{aligned} \nabla_{\mu}A_{\nu} &= \partial_{\mu}A_{\nu} - \Gamma_{\mu\nu}^{\alpha}A_{\alpha}; \\ \nabla_{\mu}A^{\nu} &= \partial_{\mu}A^{\nu} + \Gamma_{\mu\alpha}^{\nu}A^{\alpha}, \end{aligned} \quad (2.31)$$

где величины $\Gamma_{\mu\nu}^{\alpha}$ (не образующие тензора!) называются символами Кристоффеля, или коэффициентами аффинной связности,

согласованной с метрикой (метрической связности); они выражаются через метрический тензор и его первые частные производные:

$$\Gamma_{\mu\nu}^{\sigma} = \frac{1}{2}g^{\sigma\alpha}(\partial_{\nu}g_{\alpha\mu} + \partial_{\mu}g_{\alpha\nu} - \partial_{\alpha}g_{\mu\nu}), \quad (2.32)$$

причем для свертки $\Gamma_{\mu\alpha}^{\alpha}$ с помощью (2.32) получается выражение

$$\Gamma_{\mu\alpha}^{\alpha} = \partial_{\mu}(\ln \sqrt{-g}), \quad g := \det(g_{\mu\nu}). \quad (2.33)$$

Символы Кристоффеля симметричны по нижним индексам, следовательно, в общем случае в них 40 различных компонент (10 вариантов нижней пары индексов и 4 значения верхнего).

Тензорный характер преобразования компонент ковариантных производных при преобразованиях координат проверяется непосредственно.

Для тензоров любого ранга ковариантные производные вычисляются по методике (2.31), примененной отдельно к каждому верхнему и каждому нижнему индексу. Например, для смешанного тензора T_{μ}^{ν} имеем

$$\nabla_{\alpha}T_{\mu}^{\nu} = \partial_{\alpha}T_{\mu}^{\nu} + \Gamma_{\alpha\beta}^{\nu}T_{\mu}^{\beta} - \Gamma_{\alpha\mu}^{\beta}T_{\beta}^{\nu}. \quad (2.34)$$

В силу (2.32) (фактически — в силу согласованности связности с метрикой) метрический тензор ковариантно постоянен:

$$\nabla_{\alpha}g_{\mu\nu} = \nabla_{\alpha}g^{\mu\nu} = 0. \quad (2.35)$$

Из этого следует, что операции ковариантного дифференцирования и поднятия и опускания индексов коммутируют друг с другом, что весьма удобно при преобразованиях сложных тензорных выражений.

Следующий важный момент — повторное применение ковариантных производных. Как известно, в случае обычных частных производных ∂_{μ} порядок дифференцирования несуществен; то же справедливо для ковариантных производных скаляра f .

Первая, $\nabla_\mu f \equiv \partial_\mu f$, совпадает с частной, вторая применяется уже к вектору (градиенту), но, тем не менее,

$$(\nabla_\alpha \nabla_\beta - \nabla_\beta \nabla_\alpha)f = 0.$$

Применительно к вектору, коммутация ковариантных производных дает

$$\begin{aligned} (\nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu)A_\rho &= R^\alpha_{\rho\mu\nu}A_\alpha; \\ (\nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu)A^\rho &= -R^\rho_{\alpha\mu\nu}A^\alpha, \end{aligned} \quad (2.36)$$

где величины $R^\rho_{\alpha\mu\nu}$ являются компонентами тензора, называемого тензором кривизны, или тензором Римана (иногда его называют также тензором Римана–Кристоффеля):

$$R^\sigma_{\mu\rho\nu} = \partial_\nu \Gamma^\alpha_{\mu\rho} - \partial_\rho \Gamma^\alpha_{\mu\nu} + \Gamma^\alpha_{\beta\nu} \Gamma^\beta_{\rho\alpha} - \Gamma^\alpha_{\beta\rho} \Gamma^\beta_{\nu\alpha}. \quad (2.37)$$

Тензор Римана играет центральную роль в римановой геометрии, так как именно он характеризует отличие заданной метрики от метрики плоского пространства, для которой все его компоненты тождественно равны нулю (и, в силу его тензорного характера, все его компоненты равны нулю при вычислении в произвольной системе координат, а не только в координатах Минковского).

Симметрии тензора Римана удобнее описывать, представляя его целиком в ковариантных компонентах, $R_{\mu\nu\rho\sigma} = g_{\mu\alpha}R^\alpha_{\nu\rho\sigma}$. А именно: по построению он симметричен относительно перестановки первой и второй пар индексов и антисимметричен внутри каждой пары:

$$R_{\mu\nu\rho\sigma} = R_{\rho\sigma\mu\nu} = -R_{\nu\mu\rho\sigma} = -R_{\mu\nu\sigma\rho}. \quad (2.38)$$

Кроме того, компоненты тензора Римана удовлетворяют тождеству Риччи

$$R_{\mu\alpha\beta\gamma} + R_{\mu\beta\gamma\alpha} + R_{\mu\gamma\alpha\beta} = 0, \quad (2.39)$$

в котором первый индекс остается на месте, а остальные циклически переставляются. В силу (2.38) можно переписать это тождество с любым другим неподвижным индексом.

С учетом симметрии тензора Римана число его независимых компонент в четырехмерном пространстве равно 20. В общем случае D -мерного пространства это число равно $D^2(D^2 - 1)/12$. В частных примерах пространств, обладающих значительными симметриями, независимых компонент гораздо меньше: например в пространстве де Ситтера, имеющем максимально возможную симметрию, все компоненты тензора Римана выражаются через одну постоянную.

Свертки тензора Римана приводят к тензору Риччи $R_{\mu\nu}$ и скалярной кривизне R , называемой также скаляром Риччи:

$$R_{\mu\nu} = R_{\mu\alpha\nu}^{\alpha}; \quad R = g^{\mu\nu} R_{\mu\nu} = R_{\alpha}^{\alpha}. \quad (2.40)$$

Непосредственной проверкой можно установить, что тензор Римана, помимо алгебраических тождеств (2.38) и (2.39), удовлетворяет дифференциальным тождествам

$$\nabla_{[\sigma} R_{\alpha\beta]} \gamma_{\delta} = 0, \quad (2.41)$$

называемым тождествами Бьянки.

Их свертка по одной паре индексов дает

$$\nabla_{\sigma} R_{\alpha\beta\gamma}^{\sigma} + \nabla_{\gamma} R_{\alpha\beta} - \nabla_{\beta} R_{\alpha\gamma} = 0, \quad (2.42)$$

а дальнейшая свертка приводит к равенству, весьма важному для теории гравитации:

$$\nabla_{\alpha} G_{\mu}^{\alpha} = 0, \quad G_{\mu}^{\nu} := R_{\mu}^{\nu} - \frac{1}{2} \delta_{\mu}^{\nu} R. \quad (2.43)$$

Тензор G_{μ}^{ν} называется тензором Эйнштейна.

В заключение раздела приведем формулу для инвариантного элемента объема в римановом пространстве произвольной размерности D : если требуется найти объем малого параллелепипеда, заданного D векторами dx_a^{μ} (a — номер вектора, μ — как

обычно, номер его компоненты), то инвариантный элемент объема есть

$$dV = \sqrt{|g|} |\det(dx_a^\mu)|, \quad g := \det(g_{\mu\nu}). \quad (2.44)$$

Следовательно, инвариантный интеграл от скалярной функции $f(x)$ по некоторому объему V имеет вид

$$\int_V \sqrt{|g|} f(x) d^D x, \quad d^D x := dx^1 dx^2 \cdots dx^D. \quad (2.45)$$

Очевидно, аналогичные формулы справедливы для интегрирования по поверхностям любых размерностей $d < D$, если под x^μ иметь в виду координаты, заданные на поверхности, а под $g_{\mu\nu}$ — внутреннюю метрику поверхности, индуцированную метрикой объемлющего пространства.

2.4. Действие для гравитационного поля и динамические уравнения

2.4.1. Уравнения Эйнштейна

В ОТО динамическими переменными, характеризующими гравитационное поле, являются компоненты метрического тензора $g_{\mu\nu}$. Динамические уравнения ОТО выводятся из вариационного принципа Гильберта

$$\delta S = 0, \quad S = \int \frac{R}{2\kappa} \sqrt{-g} d^4 x + S_m, \quad (2.46)$$

где $S_m = \int L_m \sqrt{-g} d^4 x$ — действие материи, т.е. вещества и всех полей, кроме гравитационного. Условие $\delta S = 0$ приводит к уравнениям Гильберта — Эйнштейна (чаще говорят — уравнениям Эйнштейна)

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = -\kappa T_{\mu\nu}. \quad (2.47)$$

Здесь $\kappa = 8\pi G/c^4$ — эйнштейновская гравитационная постоянная (G — ньютоновская гравитационная постоянная), а $T_{\mu\nu}$ — (метрический) тензор энергии-импульса (ТЭИ) материи:

$$T_{\mu\nu} = \frac{2}{\sqrt{-g}} \frac{\delta S_m}{\delta g^{\mu\nu}}. \quad (2.48)$$

Система (2.47) в общем случае состоит из 10 нелинейных уравнений в частных производных. Но, во-первых, свобода выбора системы координат дает возможность наложить 4 произвольных координатных условия, которые можно сформулировать в виде равенств с участием коэффициентов $g_{\mu\nu}$, поэтому остается всего 6 независимых уравнений. Во-вторых, среди оставшихся уравнений имеются 4 дифференциальные зависимости, связанные с тождествами (2.43), и в результате остаются лишь 2 динамических уравнения. Остальные четыре — уравнения связи, не содержащие вторых производных по времени. Эти обстоятельства важны для всех динамических процессов в ОТО и прежде всего для гравитационных волн, у которых в результате имеются всего две независимые поляризации.

В силу свернутых тождеств Бьянки (2.43) и уравнений Эйнштейна (2.47) для ТЭИ любой материи справедливы дифференциальные равенства, имеющие смысл законов сохранения²

$$\nabla_\alpha T_\mu^\alpha = 0, \quad (2.49)$$

из которых можно получить уравнения движения материи. Таким образом, уравнения движения материи следуют из уравнений поля. Это обстоятельство принципиально отличает ОТО от большинства классических теорий поля, включая ньютоновскую теорию гравитации, в которых уравнения движения источников поля приходится вводить отдельно от уравнений поля.

²В общем случае эти дифференциальные равенства не приводят к интегральным законам сохранения (так как для этого они должны были бы иметь вид $\partial f_\mu^\nu / dx^\nu = 0$), и точнее было бы сказать, что они выражают изменение ТЭИ, связанное с изменением метрики. Тем не менее название “законы сохранения” утвердилось за равенствами (2.49), и, более того, часто любой тензор T_μ^ν , удовлетворяющий условиям (2.49), называют консервативным.

2.4.2. Уравнения геодезических

Уравнения движения свободных частиц в римановом пространстве-времени можно получить вариацией действия (2.21) (записанного теперь для риманова интервала). Вариационное уравнение, как и в пространстве Минковского, имеет смысл уравнения траекторий (геодезических), реализующих экстремум длины мировой линии между двумя заданными мировыми точками (и одновременно экстремум действия, аналогичного (2.21)) и записывается в виде

$$\frac{du^\alpha}{ds} + \Gamma_{\mu\nu}^\alpha u^\mu u^\nu = 0. \quad (2.50)$$

Здесь s — интервал, который совпадает с собственным временем наблюдателя, движущегося по геодезической, если он временноподобен, и с истинной длиной вдоль геодезической, если он пространственноподобен. Во всех случаях s — канонический параметр.³

Выведем уравнение (2.50) из закона сохранения (2.49). Это будет иллюстрацией возможности вывода уравнений движения материи из уравнений Эйнштейна.

Начнем с выражения для ТЭИ идеальной жидкости, которое получается как естественное обобщение соответствующего выражения из СТО [23] на римановы пространства:

$$T_{\mu\nu} = (\varepsilon + p)u_\mu u_\nu - p g_{\mu\nu}, \quad (2.51)$$

где u_μ — 4-скорость частиц жидкости, $\varepsilon = \rho c^2$ — ее плотность энергии, p — давление. В частности, для пылевидного вещества,

³Канонический параметр определяется как именно тот параметр кривой, при использовании которого уравнения геодезических записываются в виде (2.50) (см., напр., [288]). Легко видеть, что если s — канонический параметр, то и $\bar{s} = as + b$, где a и b — числа, тоже канонический параметр; в теории гравитации, однако, обычно используется именно параметр, равный интервалу, и свобода его выбора сводится к возможности сдвигов $\bar{s} = s + b$. Использование других, неканонических параметров усложнило бы вид уравнений (2.50).

состоящего из невзаимодействующих частиц,

$$p = 0, \quad T_{\mu\nu} = \varepsilon u_\mu u_\nu. \quad (2.52)$$

Пылинки движутся без влияния внешних сил, кроме гравитации, следовательно, их уравнения движения совпадают с уравнениями движения свободных частиц. Получим его, дифференцируя тензор (2.52):

$$\nabla_\alpha T_\nu^\alpha = u_\mu \nabla_\alpha (\rho u^\alpha) + \rho u^\alpha \nabla_\alpha u_\mu = 0. \quad (2.53)$$

Умножив это равенство на u^μ , учитывая, что $u^\mu u_\mu = 1 \Rightarrow \nabla_\alpha (u^\mu u_\mu) = 0$, получаем из (2.53) уравнение непрерывности $\nabla_\mu (\rho u^\mu) = 0$ (его смысл — сохранение массы пыли) и уравнение движения

$$u^\mu \nabla_\mu u_\alpha = 0, \quad (2.54)$$

которое можно переписать в виде

$$u^\mu \nabla_\mu u^\alpha = 0 \Rightarrow \frac{dx^\mu}{ds} \frac{\partial u^\alpha}{\partial x^\mu} + \Gamma_{\mu\nu}^\alpha u^\mu u^\nu = 0.$$

Отсюда окончательно получается уравнение (2.50):

$$\frac{du^\alpha}{ds} + \Gamma_{\mu\nu}^\alpha u^\mu u^\nu = 0.$$

Геодезические делятся на пространственноподобные, временноподобные и светоподобные, причем этот характер геодезической не меняется вдоль нее, так как $\partial_\alpha (u_\mu u^\mu) = 0$.

2.4.3. Принцип соответствия

1. Поскольку риманово пространство-время совпадает с пространством Минковского (своим касательным пространством) в малой окрестности любой мировой точки, в любой точке приближенно справедливы законы СТО. Возможен и переход к СТО во всем пространстве в пределе слабой гравитации, а условие

слабости гравитации формулируется в виде условия малости отклонения метрики от плоской, или малости тензора Римана.

Заметим, однако, что формальный переход $\varkappa \rightarrow 0$ в решениях уравнений Эйнштейна приводит, вообще говоря, не к плоской метрике, а к некоторому неплоскому вакуумному (с $T_{\mu\nu} \equiv 0$) решению уравнений Эйнштейна.

2. Закон тяготения Ньютона получается из ОТО при условиях малых скоростей ($v \ll c$) и слабого гравитационного поля (подходящим образом введенный ньютоновский потенциал $V \ll c^2$). В дальнейшем мы явным образом убедимся в существовании такого перехода на примере метрики Шварцшильда и в справедливости соотношения $\varkappa = 8\pi G/c^4$.

Переход к ньютоновской гравитации при некоторых дополнительных условиях на вид метрики можно осуществить и как формальный переход $c \rightarrow \infty$, а разложение метрики и характеристик материи по степеням c^{-1} (точнее, по степеням v/c и V/c^2) удобно использовать для описания наблюдаемых эффектов релятивистской гравитации как в ОТО (постньютоновское приближение), так и в других метрических теориях гравитации (параметризованное постньютоновское приближение) [308].

2.5. Макроскопическая материя и негравитационные поля в ОТО

Для всех видов материи, кроме гравитационного поля, существует описание в рамках СТО. Для их описания в рамках ОТО (и вообще в любой теории, формулируемой в римановом пространстве) чаще всего используется так называемый *принцип минимального взаимодействия*, согласно которому все уравнения, известные в СТО, переносятся в искривленное пространство-время путем замены всех частных производных на ковариантные. Заметим, что этот прием заодно позволяет свободнее действовать уже в рамках СТО, не ограничиваясь координатами Минковско-

го, а вводя произвольные криволинейные координаты и СО с произвольными ускорениями, как поступательными, так и вращательными.

Приведем некоторые соотношения, справедливые для негравитационной материи в искривленном пространстве-времени согласно принципу минимального взаимодействия.

2.5.1. Идеальная жидкость

Ранее уже приводилось выражение (2.51) для ТЭИ идеальной жидкости в римановом пространстве-времени. Из закона сохранения для него выведем уравнения движения жидкости — общерелятивистские аналоги уравнения непрерывности и уравнения Эйлера.

Перепишем тензор (2.51) в смешанных компонентах,

$$T_{\mu}^{\nu} = (\varepsilon + p)u_{\mu}u^{\nu} - p\delta_{\mu}^{\nu}, \quad (2.55)$$

подействуем на него оператором ∇_{ν} и приравняем нулю:

$$\nabla_{\nu}T_{\mu}^{\nu} = (\partial_{\nu}w)u_{\mu}u^{\nu} + w\nabla_{\nu}(u_{\mu}u^{\nu}) - \partial_{\mu}p = 0, \quad (2.56)$$

где $w := \varepsilon + p$ — тепловая функция жидкости. Свертывая равенство (2.56) с u^{μ} (т.е. проектируя его на направление u^{μ}) и учитывая, что $\partial_{\mu}(u_{\nu}u^{\nu}) = 0$, получим:

$$\nabla_{\nu}(wu^{\nu}) - u^{\mu}\partial_{\mu}p = 0. \quad (2.57)$$

Теперь спроектируем (2.56) на направление, перпендикулярное u^{μ} — такая проекция имеет вид $\nabla_{\nu}T_{\mu}^{\nu} - u^{\nu}u_{\mu}\nabla_{\lambda}T_{\nu}^{\lambda} = 0$. В результате приходим к уравнению движения идеальной жидкости — общерелятивистскому аналогу уравнения Эйлера:

$$wu^{\nu}\nabla_{\nu}u_{\mu} = \partial_{\mu}p - u_{\mu}u^{\nu}\partial_{\nu}p. \quad (2.58)$$

Уравнение непрерывности в нерелятивистской гидродинамике, как известно, представляет собой закон сохранения массы. В

СТО и тем более в ОТО масса не сохраняется, и аналоги уравнения непрерывности получаются лишь для сохраняющихся величин, например, числа частиц, если можно пренебречь их возможным рождением и поглощением. Тогда можно ввести ток числа частиц $n^\mu = nu^\mu$, где n — плотность числа частиц в СО, в которой жидкость, состоящая из них, неподвижна, а u^μ — 4-скорость этой жидкости. Закон сохранения числа частиц (справедливый при отсутствии реакций с их рождением, уничтожением и превращениями) выражается равенством

$$\nabla_\mu(nu^\mu) = 0, \quad (2.59)$$

вполне аналогичным закону сохранения электрического заряда (2.67).

2.5.2. Скалярные поля

Для скалярного поля ϕ с произвольным самодействием, описываемым потенциалом $V(\phi)$, лагранжиан в искривленном пространстве в случае минимальной связи с гравитацией записывается точно так же, как в пространстве Минковского:

$$L_s = \frac{1}{2}g^{\mu\nu}\phi_{,\mu}\phi_{,\nu} - V(\phi), \quad (2.60)$$

а вариация по ϕ с учетом зависимости метрики от координат приводит к уравнению, обобщающему релятивистское уравнение Клейна — Гордона,

$$\square\phi + dV/d\phi = 0, \quad (2.61)$$

с общерелятивистским оператором д'Аламбера

$$\square = \nabla^\alpha\nabla_\alpha = \frac{1}{\sqrt{-g}}\partial_\alpha\left(\sqrt{-g}g^{\alpha\beta}\partial_\beta\right). \quad (2.62)$$

ТЭИ скалярного поля получается из лагранжиана (2.60) вариацией согласно (2.48):

$$T^\nu{}_\mu = \phi_{,\mu}\phi^{,\nu} - \delta^\nu{}_\mu L_s. \quad (2.63)$$

2.5.3. Электромагнитное поле

Электромагнитное (безмассовое векторное) поле характеризуется вектор-потенциалом A_μ и тензором напряженности, называемым также тензором Максвелла

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (2.64)$$

Лагранжиан электромагнитного поля непосредственно обобщает соответствующее выражение для плоского пространства. Для поля с источниками лагранжиан имеет вид

$$L_{e-m} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - j^\mu A_\mu, \quad (2.65)$$

где j^μ — плотность тока электрических зарядов. Его вариация относительно A_μ дает динамическое уравнение, соответствующее в обычной электродинамике второй паре уравнений Максвелла,

$$\nabla_\nu F^{\mu\nu} = \frac{1}{\sqrt{-g}}\partial_\nu (\sqrt{-g}F^{\mu\nu}) = j^\mu, \quad (2.66)$$

причем в силу (2.66) автоматически выполняется закон сохранения электрического заряда

$$\nabla_\mu j^\mu \equiv \frac{1}{\sqrt{-g}}\partial_\mu (\sqrt{-g}j^\mu) = 0. \quad (2.67)$$

Первой паре уравнений Максвелла отвечает тождество, вытекающее из (2.64),

$$\nabla_\mu F_{\nu\sigma} + \nabla_\sigma F_{\mu\nu} + \nabla_\nu F_{\sigma\mu} = 0, \quad (2.68)$$

или, эквивалентно,

$$\nabla_\mu *F^{\mu\nu} = 0, \quad (2.69)$$

где $*F^{\mu\nu}$ — тензор, дуальный по отношению к $F_{\mu\nu}$:

$$*F^{\mu\nu} = \frac{1}{2} \frac{\varepsilon^{\mu\nu\rho\sigma}}{\sqrt{-g}} F_{\rho\sigma}. \quad (2.70)$$

ТЭИ электромагнитного поля получается вариацией действия поля $F_{\mu\nu}$ по метрике $g^{\mu\nu}$:

$$T_{\mu e-m}^{\nu} = -F_{\mu\alpha}F^{\nu\alpha} + \frac{1}{4}\delta_{\mu}^{\nu}F_{\alpha\beta}F^{\alpha\beta}. \quad (2.71)$$

В этой книге не рассматриваются другие виды полей, такие, как спинорные (спин 1/2), массивное векторное, поле спина 3/2. Подробные обсуждения свойств этих полей в искривленных пространствах и их применений в задачах гравитации, физики частиц и астрофизики можно найти в [24].

Нередко исследуются также неминимальные взаимодействия между материальными полями и гравитацией. Они вводятся добавлением в лагранжиан членов, включающих материальное поле и какие-либо инварианты кривизны, чаще всего — скалярной кривизны. Например, неминимальное взаимодействие скалярного поля с гравитацией можно ввести, добавляя в (2.60) член $\xi R\phi^2$, $\xi = \text{const}$, а неминимальное взаимодействие электромагнитного поля и гравитации может описываться, например, членом в лагранжиане вида $\xi R F_{\mu\nu}F^{\mu\nu}$. В дальнейшем мы столкнемся с неминимальными взаимодействиями скалярных полей и метрики в рамках скалярно-тензорных и многомерных теорий гравитации.

2.6. Наиболее симметричные пространства

2.6.1. Группы изометрий и векторы Киллинга

Плоское пространство-время обладает максимально возможной симметрией, которая выражается инвариантностью интервала относительно группы Пуанкаре (G_{10} , т.е. 10-параметрической группой). В общем случае риманово пространство-время \mathbb{V}_4 не обладает никакой симметрией, а в тех частных случаях, когда симметрия есть, она выражается группой изометрий с таким же или меньшим числом параметров. Это число равно числу линейно независимых решений ξ_{α} уравнения Киллинга [227]

$$\nabla_{\alpha}\xi_{\beta} + \nabla_{\beta}\xi_{\alpha} = 0 \quad (2.72)$$

Векторы ξ^μ , удовлетворяющие уравнению (2.72) (векторы Киллинга), обладают следующим свойством: при смещении всех точек на $\xi^\alpha d\lambda$ ($d\lambda$ — бесконечно малая величина) все метрические соотношения в \mathbb{V}_4 остаются неизменными. Последовательные сдвиги на $\xi^\alpha d\lambda$ приводят к движению мировой точки вдоль *орбиты вектора Киллинга*, все точки которой эквивалентны между собой. (Пример: на произвольной поверхности вращения в обычном трехмерном пространстве имеется вектор Киллинга, направленный вдоль “параллели” и описывающий сдвиги вдоль нее на малые азимутальные углы, а вся “параллель” есть орбита этого вектора Киллинга.)

Известно, что максимально симметричные (G_{10}) метрики \mathbb{V}_4 являются решениями уравнений Эйнштейна в вакууме

$$R_\mu^\nu - \delta_\mu^\nu \Lambda = 0 \quad (2.73)$$

с космологической постоянной Λ , которой соответствует эффективный ТЭИ: $T_\mu^\nu = \delta_\mu^\nu \Lambda / \kappa$. Это пространства, называемые пространствами постоянной кривизны: нулевой — пространство Минковского ($\Lambda = 0$), положительной — пространство де Ситтера (при $\Lambda > 0$) и отрицательной — пространство анти-де Ситтера (при $\Lambda < 0$). В пространствах постоянной кривизны тензор Римана определяется одной постоянной Λ и выражается через метрический тензор алгебраически по формуле

$$R_{\mu\nu\rho\sigma} = \frac{\Lambda}{3}(g_{\mu\rho}g_{\nu\sigma} - g_{\mu\sigma}g_{\nu\rho}), \quad (2.74)$$

откуда для тензора и скаляра Риччи получаем

$$R_{\mu\nu} = \Lambda g_{\mu\nu}, \quad R = 4\Lambda. \quad (2.75)$$

Явный вид метрик пространств де Ситтера [279] и анти-де Ситтера (название условное, автор по фамилии Анти-де Ситтер вымышлен) выпишем несколько позднее.

2.6.2. Изотропная космология. Пространства dS и AdS

Меньшей, но высокой симметрией (G_6) обладают однородные изотропные космологические модели, которые будут подробнее рассматриваться в следующей главе. Здесь мы ограничимся получением метрик пространств, обладающих дополнительной по отношению к G_6 симметрией — метрик де Ситтера и анти-де Ситтера.

Общий вид метрики однородного и изотропного пространства, которая в этом случае называется метрикой Фридмана — Робертсона — Уокера (ФРУ), следующий:

$$ds^2 = dt^2 - a^2(t)dl^2, \quad dl^2 = \frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2), \quad (2.76)$$

где $k = 1, 0, -1$ соответствует закрытым (сферическим), открытым пространственно плоским и открытым гиперболическим моделям; t — физическое время по часам сопутствующих наблюдателей. Это метрика с максимально симметричными пространственными сечениями: в случае $k = +1$ каждое такое сечение является трехмерной сферой, в случае $k = 0$ — трехмерным евклидовым пространством, а в случае $k = -1$ — трехмерным пространством Лобачевского; группа изометрий G_6 каждого из них, включающая трансляции по трем направлениям и повороты вокруг трех осей, одновременно является группой изометрий всего четырехмерного пространства-времени.

В силу симметрии пространства ТЭИ материи, независимо от ее природы, обладает структурой $T_\mu^\nu = \text{diag}(\rho, -p, -p, -p)$, где ρ — плотность энергии, p — изотропное давление.

Компоненты уравнений Эйнштейна $\binom{0}{0}$ и $\binom{1}{1}$ записываются в виде

$$\frac{3}{a^2}(\dot{a}^2 + k) = \varkappa\rho, \quad (2.77)$$

$$\frac{1}{a^2}(2a\ddot{a} + \dot{a}^2 + k) = -\varkappa p, \quad (2.78)$$

где точка обозначает d/dt . Уравнение (2.78) является следствием из (2.77) с учетом закона сохранения (2.49), который дает

$$\frac{d\rho}{\rho + p} = -\frac{3 da}{a}. \quad (2.79)$$

Таким образом, получилась система двух независимых уравнений с тремя неизвестными функциями времени: $a(t)$, $p(t)$, $\rho(t)$. Чтобы сделать систему определенной, надо добавить к ней уравнение состояния материи — соотношение между p и ρ , от которого и будет зависеть вид решения.

Космологическая постоянная соответствует уравнению состояния $p = -\rho = -\varkappa\Lambda$, и (2.79) дает $\rho = \text{const} = \Lambda/\varkappa$. Решения для разных k и Λ следующие: при $\Lambda > 0$ получаем

$$k = 0 : \quad a(t) = a_0 e^{\pm H_0 t}, \quad a_0 = \text{const}; \quad (2.80)$$

$$k = 1 : \quad a(t) = H_0^{-1} \text{ch}[H_0(t - t_0)], \quad t_0 = \text{const}; \quad (2.81)$$

$$k = -1 : \quad a(t) = H_0^{-1} \text{sh}[H_0(t - t_0)], \quad t_0 = \text{const}, \quad (2.82)$$

где $H_0 := \sqrt{\Lambda/3}$. При $\Lambda < 0$ уравнение (2.77) имеет решения только для гиперболических моделей, $k = -1$:

$$a(t) = H_0^{-1} \sin[H_0(t - t_0)] \quad t_0 = \text{const}, \quad l = \sqrt{-3/\Lambda}. \quad (2.83)$$

Метрики (2.80)–(2.82) описывают пространство де Ситтера в различных координатах, *принадлежащих разным СО*. Легко видеть, что даже топология пространственных сечений различна: это \mathbb{R}^3 для моделей (2.80) и (2.82) и \mathbb{S}^3 (трехмерная сфера) для модели (2.81). Анализ показывает (см., напр., книгу Хокинга и Эллиса [163]), что пространство-время де Ситтера описывается полностью метрикой (2.81), тогда как (2.80) и (2.82) описывают лишь некоторые его части.

Метрика (2.83) описывает пространство анти-де Ситтера (АдС).

Глава 3

Черные дыры

В этой главе мы рассмотрим некоторые вопросы физики черных дыр. Черной дырой (ЧД) называют область пространства-времени, в которой гравитационное поле настолько сильно, что не позволяет никаким материальным телам и даже свету покинуть эту область и уйти на бесконечность [26].

По мнению большинства астрофизиков, ЧД широко распространены во Вселенной, они в большом количестве возникали на ранних этапах ее развития, они образуются в результате эволюции массивных звезд, а сверхмассивные ЧД ($\gtrsim 10^6$ масс Солнца) располагаются в центральных областях многих типов галактик, включая нашу. ЧД обладают многими интересными и необычными свойствами, которые подробно описываются в книгах [26, 163, 227, 299] и др. Здесь мы даем лишь краткое элементарное введение в эту сложную область физики, в основном ограничиваясь случаем сферической симметрии. Впрочем, некоторые из приводимых здесь результатов (например, методика построения диаграмм Картера — Пенроуза в общем случае, точные решения со скалярными полями и др.), насколько нам известно, до настоящего времени публиковались лишь в журнальных статьях и появляются в книге впервые.

3.1. Сферически-симметричные гравитационные поля

Сферическая симметрия — естественное предположение для описания простейших изолированных тел и островных конфигураций. Сферически-симметричные пространства инвариантны относительно пространственных вращений, образующих группу изометрий G_3 .

В общем случае сферически-симметричную метрику можно записать в виде

$$ds^2 = e^{2\gamma} dt^2 - e^{2\alpha} du^2 - e^{2\beta} d\Omega^2, \quad d\Omega^2 = d\theta^2 + \sin^2 \theta d\phi^2, \quad (3.1)$$

где α , β , γ — в общем случае функции радиальной координаты u и времени t . Будем также использовать обозначение $r \equiv e^\beta$; таким образом, r — радиус координатной сферы $u = \text{const}$, $t = \text{const}$, или радиальная координата Шварцшильда. В выражении (3.1) имеется свобода выбора системы отсчета (СО): различные СО соответствуют различному распределению радиальных скоростей тел отсчета.

В случае статического пространства-времени (группа симметрии G_4 , к вращениям добавляются временные сдвиги) можно выбрать СО таким образом, что α , β , γ зависят только от u . Но остается возможность замены радиальной координаты u при помощи преобразований вида $u = u(u_{\text{new}})$; можно зафиксировать радиальную координату, постулируя какое-либо соотношение между функциями α , β , γ .

Для решения различных задач удобны разные варианты таких координатных условий. Перечислим некоторые из них.

- 1) $e^\alpha \equiv 1$, $du = dl$ — гауссовы полярные координаты; координата l представляет собой истинную длину вдоль радиального направления, отсчитываемую от некоторой фиксированной сферы $l = 0$.

- 2) $u = r$, $\gamma = \gamma(r)$, $\alpha = \alpha(r)$ — координаты кривизн: r есть радиус кривизны сферы $r = \text{const}$.
- 3) $e^{2\beta(u)} = e^{2\alpha(u)}u^2$ — изотропные координаты, в которых пространственная часть метрики записывается в конформно-плоском виде:

$$dl^2 = e^{2\alpha(u)}(du^2 + u^2d\Omega^2) = e^{2\alpha(u)}d\vec{x}^2, \quad (3.2)$$

где в трехмерном плоском линейном элементе введены декартовы координаты.

- 4) $\alpha(u) = 2\beta(u) + \gamma(u)$ — гармоническая координата u , удобная, в частности, для решения задач со скалярными полями.
- 5) $\alpha = -\gamma$ — квазиглобальная координата u . Название связано с тем, что (как мы увидим позднее) она более, чем другие, удобна для исследования черных дыр и других пространств с горизонтами.
- 6) $\alpha(u) = \gamma(u)$ — координата u , иногда называемая “черепашью”, так как во многих важных случаях в терминах этой координаты метрические функции меняются очень медленно. В этом случае метрика (t, u) -подпространства имеет конформно-плоский вид, что, например, упрощает запись волновых уравнений.

Применяются и другие координатные условия.

Приведем вид некоторых величин для метрики (3.1) с произвольной координатой u , не предвещая заранее ее выбор. Найти те же величины в конкретных координатах, например в перечисленных выше, несложно, подставив соответствующие координатные условия.

Тензор Риччи для статической метрики (3.1) диагонален; в случае же зависимости от t появляется компонента $R_{01} \neq 0$.

Приведем его ненулевые компоненты в общем случае:

$$\begin{aligned}
R_0^0 &= e^{-2\gamma}[2\dot{\beta} + \ddot{\alpha} + 2\dot{\beta}^2 + \dot{\alpha}^2 - \dot{\gamma}(2\dot{\beta} + \dot{\alpha})] - \\
&\quad - e^{-2\alpha}[\gamma'' + \gamma'(2\beta' + \gamma' - \alpha')]; \\
R_1^1 &= e^{-2\gamma}[\ddot{\alpha} + \dot{\alpha}(2\dot{\beta} - \dot{\gamma} + \dot{\alpha})] - \\
&\quad - e^{-2\alpha}[2\beta'' + \gamma'' + 2\beta'^2 + \gamma'^2 - \alpha'(2\beta' + \gamma')]; \\
R_2^2 &= e^{-2\beta} + e^{-2\gamma}[\ddot{\beta} + \dot{\beta}(2\dot{\beta}) - \dot{\gamma} + \dot{\alpha}] - \\
&\quad - e^{-2\alpha}[\beta'' + \beta'(2\beta' + \gamma' - \alpha')] = R_3^3; \\
R_{01} &= 2[\dot{\beta}' + \dot{\beta}\beta' - \dot{\alpha}\beta' - \dot{\beta}\gamma'].
\end{aligned} \tag{3.3}$$

Выражения R_μ^ν для статической метрики получаются отсюда, если приравнять нулю все производные по времени.

Из компонент тензора Эйнштейна нам в первую очередь будет нужна компонента G_1^1 , поскольку она не содержит вторых производных по u ; для статической метрики

$$G_1^1 = -e^{-2\beta} + e^{-2\alpha}\beta'(\beta' + 2\gamma'). \tag{3.4}$$

Выражения (3.3) и (3.4) необходимы для подстановки в основные уравнения, определяющие свойства самогравитирующих систем в ОТО — уравнениях Эйнштейна (2.47), которые можно записывать в двух эквивалентных формах:

$$G_\mu^\nu \equiv R_\mu^\nu - \frac{1}{2}\delta_\mu^\nu R = -\varkappa T_\mu^\nu \tag{3.5}$$

и

$$R_\mu^\nu \equiv R_\mu^\nu = -\varkappa(T_\mu^\nu - \frac{1}{2}\delta_\mu^\nu T), \tag{3.6}$$

где $T \equiv T_\alpha^\alpha$ — след тензора энергии-импульса и принято $\varkappa = 8\pi G$ в соответствии с ньютоновским пределом ОТО.

Для выяснения свойств регулярности или сингулярности пространства-времени в тех или иных точках часто используется скаляр Кречмана (иногда условно называемый квадратом тензора Римана) $\mathcal{K} = R_{\alpha\beta\gamma\delta}R^{\alpha\beta\gamma\delta}$. Для статической метрики

(3.1) он является суммой квадратов всех ненулевых компонент $R_{\alpha\beta}{}^{\gamma\delta}$ тензора Римана:

$$\begin{aligned}
\mathcal{K} &= 4K_1^2 + 8K_2^2 + 8K_3^2 + 4K_4^2, \\
K_1 &= e^{-\alpha-\gamma}(\gamma' e^{\gamma-\alpha})' = -R_{01}{}^{01}, \\
K_2 &= e^{-2\alpha}\beta'\gamma' = -R_{02}{}^{02} = -R_{03}{}^{03}, \\
K_3 &= e^{-\alpha-\beta}(\gamma' e^{\beta-\alpha})' = -R_{12}{}^{12} = -R_{13}{}^{13}, \\
K_4 &= -e^{-2\beta} + e^{-2\alpha}\beta'^2 = -R_{23}{}^{23}.
\end{aligned} \tag{3.7}$$

Существенно то, что все K_i инвариантны относительно репараметризации координаты u , т.е. преобразований $u \mapsto f(u)$, иными словами, ведут себя при таких преобразованиях как скаляры. Так же ведут себя и смешанные компоненты всех тензоров с двумя индексами, включая R_μ^ν и G_μ^ν — именно поэтому, записывая уравнения Эйнштейна, мы предпочитаем иметь дело с ними, а не с компонентами $R_{\mu\nu}$, $G_{\mu\nu}$ или $R^{\mu\nu}$, $G^{\mu\nu}$, чувствительными к репараметризациям u . Заметим, что репараметризация $u \mapsto f(u)$ — частный случай чисто пространственных преобразований координат, оставляющих без изменения систему отсчета (см. главу 2).

Так как скаляр \mathcal{K} есть сумма квадратов, для его конечности необходимо и достаточно, чтобы были конечны *все* его составляющие K_i , иными словами, все ненулевые компоненты тензора Римана $R^{\mu\nu}{}_{\alpha\beta}$. Следовательно, если конечен скаляр \mathcal{K} , то конечны *все* инварианты, которые можно составить алгебраически из тензора Римана и метрического тензора, например, скалярная кривизна R , “квадрат тензора Риччи” $R_{\alpha\beta}R^{\alpha\beta}$, инвариант $R_{\alpha\beta\gamma\delta}R^{\alpha\gamma}R^{\beta\delta}$ и т.д. Таким образом, конечность \mathcal{K} в какой-либо точке пространства-времени означает отсутствие в этой точке сингулярности кривизны.

Условия регулярного центра и плоской асимптотики

Центр в статическом сферически-симметричном пространстве-времени есть, по определению, точка, линия или поверхность,

где $r \equiv e^\beta = 0$, т.е. место, где координатные сферы стягиваются в точки. Центр может быть регулярным или сингулярным; регулярность, как и у любой точки пространства-времени, определяется конечностью всех K_i в выражении (3.7). Необходимо отметить, что центра в сферически-симметричном пространстве-времени может вовсе не быть — так происходит, если величина r нигде не обращается в нуль во всем пространстве-времени или, по крайней мере, в его статической области. С примером такого поведения мы вскоре столкнемся при обсуждении геометрии Шварцшильда.

При использовании произвольной координаты u необходимые и достаточные условия регулярности центра ($r = 0$) получаются в виде

$$\gamma = \gamma_0 + O(r^2), \quad |\beta'| e^{-\alpha+\beta} = 1 + O(r^2), \quad (3.8)$$

где γ_0 — некоторая постоянная. Второе условие получается из требования конечности величины K_4 в (3.7). Его смысл заключается в том, что отношение длины окружности к ее радиусу должно принимать правильное значение (2π) для малых окружностей, описанных вокруг центра. Это гарантирует локальную плоскостность пространства в центре и наличие там касательного пространства — свойство, присущее любой регулярной точке. Для центра приходится вводить специальные условия регулярности в силу того, что центр является особой точкой используемого нами класса сферических систем координат.

Плоская асимптотика. При $r \rightarrow \infty$, вдали от источника гравитационного поля, во многих важных случаях (хотя и далеко не всегда) геометрия пространства-времени совпадает с геометрией Минковского — такие пространства и называют асимптотически плоскими. Прежде всего это означает обращение в нуль всех компонент тензора Римана, т.е. все $K_i \rightarrow 0$. Последнего условия, однако, недостаточно: кроме того, для асимптотической плоскостности необходимо потребовать правильного значения (2π) отношения длины окружности к радиусу для больших координатных

окружностей, иначе даже при равном нулю тензоре Римана у трехмерного пространства получится асимптотика типа воронки с дефицитом или избытком полного телесного угла по сравнению с его обычным значением 4π . Такая геометрия характерна для конфигураций особого вида, называемых *глобальными монополями*. (Двумерный аналог подобных трехмерных поверхностей — коническая поверхность, метрика на которой плоская во всех точках, кроме вершины, однако отношение длины окружности, описанной вокруг вершины, к ее радиусу, отлично от 2π .)

Как и центра, плоской асимптотики в сферически-симметричном пространстве-времени может не быть; более того, может вообще не быть предела $r \rightarrow \infty$. С примером такой геометрии мы уже встречались: закрытый мир Фридмана будучи сферически-симметричным содержит в каждом пространственном сечении два центра, но не содержит пространственной бесконечности.

Выпишем необходимые и достаточные условия плоской асимптотики для метрики (3.1) с произвольной координатой u при $r \rightarrow \infty$:

$$\gamma = \gamma_\infty + O(1/r), \quad |\beta'| e^{-\alpha+\beta} = 1 + O(1/r), \quad (3.9)$$

где γ_∞ — некоторая постоянная.

Напомним, что в выражениях (3.8) и (3.9), как и всюду, равенство $y = O(x)$ означает, что либо x и y — величины одного порядка ($y \sim x$), либо y много меньше x (обозначение $y \ll x$ или $y = o(x)$).

3.2. Решение Райснера — Нордстрема — (анти)-де Ситтера

3.2.1. Решение уравнений Эйнштейна

Найдем класс точных статических сферически-симметричных решений уравнений Эйнштейна для вакуума, а также в присутствии электромагнитного поля и космологической постоянной.

Этот класс, во-первых, содержит метрики, имеющие наибольшее число астрофизических приложений среди всех сферически-симметричных метрик, и, во-вторых, послужит в дальнейшем источником явных примеров при обсуждении общих свойств сферически-симметричных пространств, включая пространства с черными дырами.

Удобно решать названную задачу в координатах кривизн, в которых два независимых уравнения Эйнштейна можно записать в виде (3.5):

$$G_0^0 + \Lambda = e^{-2\alpha} \left(\frac{1}{r^2} - \frac{2\alpha'}{r} \right) - \frac{1}{r^2} + \Lambda = -\varkappa T_0^0, \quad (3.10)$$

$$G_1^1 + \Lambda = e^{-2\alpha} \left(\frac{1}{r^2} + \frac{2\gamma'}{r} \right) - \frac{1}{r^2} + \Lambda = -\varkappa T_1^1, \quad (3.11)$$

где штрих означает d/dr , а ТЭИ в данном случае соответствует электромагнитному полю.

Для электромагнитного поля лагранжиан и ТЭИ имеют вид

$$L_{e-m} = -\frac{1}{8\pi} F_{\mu\nu} F^{\mu\nu}, \quad T_\mu^\nu = \frac{1}{16\pi} [-4F_{\mu\alpha} F^{\nu\alpha} + \delta_\mu^\nu F_{\alpha\beta} F^{\alpha\beta}]. \quad (3.12)$$

Уравнения Масквелла $\nabla_\alpha F^{\alpha\beta} = 0$ надо записать для случая сферической симметрии, когда отличными от нуля могут быть лишь радиальное электрическое поле ($F_{01} = -F_{10}$) и радиальное магнитное поле ($F_{23} = -F_{32}$). Последнее возможно, лишь если предположить существование магнитных монополей. Ограничимся электрическим полем. Единственное нетривиальное уравнение Максвелла дает

$$(\sqrt{-g} F^{01})' = 0 \quad \Rightarrow \quad F^{01} = Q e^{-\alpha-\gamma}/r^2, \\ F_{10} = Q e^{\alpha+\gamma}/r^2, \quad (3.13)$$

где константа Q интерпретируется как электрический заряд. ТЭИ имеет вид

$$T_\mu^\nu = \frac{Q^2}{r^4} \text{diag}(+1, +1, -1, -1). \quad (3.14)$$

Займемся теперь решением уравнений Эйнштейна. Уравнение (3.10) приводится к виду

$$\begin{aligned} [r(A-1)]' &= -(\Lambda + \varkappa T_0^0)r^2 \Rightarrow \\ \Rightarrow A &= 1 - \frac{\Lambda r^2}{3} - \frac{\varkappa}{r} \int T_0^0 r^2 dr, \end{aligned} \quad (3.15)$$

где $A(r) = e^{-2\alpha}$. С другой стороны, разность уравнений (3.10) и (3.11) при условии $T_0^0 = T_1^1$ дает

$$\alpha' + \gamma' = 0,$$

откуда

$$e^{\alpha+\gamma} = 1 \quad (3.16)$$

при подходящем выборе масштаба по оси времени. В результате интегрирования (3.15) метрика принимает вид

$$\begin{aligned} ds^2 &= A(r)dt^2 - \frac{dr^2}{A(r)} - r^2 d\Omega^2, \\ A(r) &= 1 - \frac{\Lambda}{3}r^2 - \frac{2GM}{r} + \frac{GQ^2}{r^2}, \end{aligned} \quad (3.17)$$

где константа интегрирования M интерпретируется как активная гравитационная масса источника.

3.2.2. Частные случаи

Метрика (анти-)де Ситтера

В частном случае нулевых заряда и массы функция $A(r)$ в метрике (3.17) принимает простой вид $A = 1 - (\Lambda/3)r^2$. При $\Lambda = 0$ приходим к метрике Минковского. Далее, при $\Lambda > 0$, как показывает анализ, получаются в статическом виде (т.е. в координатах, принадлежащих другой СО) уже знакомое нам по предыдущей главе пространство де Ситтера, а при $\Lambda < 0$ — пространство анти-де Ситтера.

Рассмотрим вариант $\Lambda = 3H^2 > 0$ несколько подробнее. Метрика имеет вид

$$ds^2 = (1 - H^2 r^2) dt^2 - (1 - H^2 r^2)^{-1} dr^2 - r^2 d\Omega^2. \quad (3.18)$$

При $r = 0$ имеется регулярный центр — как и должно быть, поскольку будучи пространством постоянной кривизны, пространство де Ситтера однородно, все его точки равноправны и регулярны, и любая из них может быть “назначена” центром.

При $r = 1/H$ оба метрических коэффициента g_{00} и g_{11} обращаются в ноль (эта сфера называется горизонтом), а при $r > 1/H$ они принимают отрицательные значения, в результате чего координаты r и t меняются ролями: величина r носит характер временной координаты, а t — пространственной. Такая область называется Т-областью; общее описание горизонтов, R- и Т- областей см. в следующем разделе. В Т-области метрику можно переписать в виде

$$\begin{aligned} ds^2 &= (H^2 r^2 - 1)^{-1} dr^2 - (H^2 r^2 - 1) dx^2 - r^2 d\Omega^2 = \\ &= d\tau^2 - \text{sh}^2(H\tau) dx^2 - H^{-2} \text{ch}^2(H\tau) d\Omega^2. \end{aligned} \quad (3.19)$$

где “бывшее время” t переименовано в x , а переменная τ , введенная вместо r , имеет смысл собственного времени в получившейся однородной анизотропной космологической модели. Однородность следует из того, что, помимо сферической симметрии, метрические коэффициенты не зависят от пространственной координаты x ; анизотропия же заключается в том, что имеется два различных масштабных фактора, $|g_{xx}| = \text{sh}^2(H\tau)$ и $|g_{\theta\theta}| = r^2 = \text{ch}^2(H\tau)$, характеризующих расширение Вселенной, соответственно, в направлении x и в двух угловых направлениях. Сферически-симметричные космологические модели с топологией пространственного сечения $\mathbb{R} \times \mathbb{S}^2$, частным случаем которых является (3.19), называются моделями Кантовского–Сакса. Забегая вперед, укажем, что именно такими моделями описываются Т-области всех сферически-симметричных черных дыр.

В модели (3.19) космологическое расширение начинается с горизонта $\tau = 0$ — крайне анизотропного состояния, $g_{xx} = 0$, и становится изотропным и экспоненциальным (с постоянной Хаббла H) в пределе больших τ . Это — еще один вид метрики де Ситтера в дополнение к (2.80)–(2.82).

Подчеркнем, что в данном описании метрики (3.18) и (3.19) следует рассматривать отдельно друг от друга, так как при $r = H^{-1}$ в силу $A = 0$ в обеих метриках имеют место координатные сингулярности. То же относится ко всем R- и T-областям до конца раздела 6.2. На всех горизонтах, соответствующих регулярным нулям функции $A(r)$, как нетрудно проверить с помощью формул (3.7), 4-метрика регулярна, и естественно предположить, что переход к новым координатам позволит избавиться от координатных сингулярностей и получить более полную картину пространства-времени, естественно включающую различные R- и T-области. Такие преобразования будут описаны в последующих разделах.

Метрика Шварцшильда и закон Ньютона

В случае $\Lambda = 0$, $Q = 0$ метрика (3.17) превращается в метрику Шварцшильда

$$ds^2 = \left(1 - \frac{2GM}{r}\right) dt^2 - \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 - r^2 d\Omega^2, \quad (3.20)$$

где постоянная интегрирования M имеет смысл активной гравитационной массы. Для подтверждения этой интерпретации используем уравнения геодезических (2.50) в случае пробной частицы, мгновенно покоящейся при большом (по сравнению с $2GM$) значении радиуса r :

$$\frac{d^2 x^i}{ds^2} + \Gamma_{00}^i (u^0)^2 = 0,$$

где $ds^2 = g_{00} dt^2 \approx dt^2$ и $u^0 \approx 1$ в силу принятых допущений, а среди символов Кристоффеля Γ_{00}^i отличен от нуля лишь

$\Gamma_{00}^1 = \gamma' e^{2\gamma-2\alpha}$ (в терминах метрики (3.1)), или, с учетом (3.20), $\Gamma_{00}^1 \approx GM/r^2$. Таким образом, частица испытывает ускорение $-GM/r^2$, направленное к центру, что и требовалось.

Метрикой Шварцшильда описывается поле тяготения в пустоте вокруг произвольного изолированного сферически-симметричного тела в случаях, когда вдали от него можно считать пространство-время асимптотически плоским — а это так для подавляющего большинства явлений, представляющих астрофизический интерес. Свойства метрики (3.20) подробно обсуждаются в разделе 3.4. Здесь упомянем лишь, что, как и метрика де Ситтера (3.18), метрика (3.20) содержит один горизонт при $r = r_g = 2GM$, внутри которого при $r < r_g$ находится Т-область, описываемая некоторой моделью Кантовского-Сакса.

Область $r \leq r_g$ представляет собой простейший пример черной дыры.

Метрика Райснера — Нордстрема

Метрика Райснера — Нордстрема получается из (3.17) при $\Lambda = 0$ и описывает внешнее гравитационное поле сферического тела с массой M и электрическим зарядом Q . Геометрические и, следовательно, физические свойства пространства-времени Райснера — Нордстрема зависят от соотношения между массой и зарядом. Введем “геометризованные” массу и заряд, имеющие размерность длины:

$$m = GM, \quad q = \sqrt{G}Q, \quad (3.21)$$

тогда при условии $m^2 > q^2$ квадратный трехчлен (см. (3.17))

$$r^2 A(r) = r^2 - 2mr + q^2 \quad (3.22)$$

имеет два положительных корня

$$r_{\pm} = m \pm \sqrt{m^2 - q^2}, \quad (3.23)$$

так что $A(r) > 0$ (метрика статическая) при $r_+ < r < \infty$ — это внешняя R-область, а при $0 < r < r_-$ — внутренняя R-область. Промежуточные значения $r_- < r < r_+$ соответствуют T-области, в которой, как и в решениях Шварцшильда и де Ситтера, метрика описывает некоторую однородную анизотропную модель Кантовского–Сакса. Сферы $r = r_+$ и $r = r_-$, разделяющие R- и T-области, называют внешним и внутренним горизонтами ЧД Райснера — Нордстрема.

При условии $m^2 = q^2$, когда $A = (r - m)^2$, два горизонта сливаются в один, $r = r_h = m = |q|$ (его называют двойным, или экстремальным горизонтом, а область $r \leq m$ (впрочем, иногда так называют всю конфигурацию) — экстремальной ЧД Райснера — Нордстрема). По обе стороны от $r = r_h$ функция $A(r)$ положительна, т.е. имеют место R-области.

Наконец, при “больших зарядах”, $m^2 < q^2$, всё пространство $r > 0$ занимает одна R-область.

При $r \rightarrow \infty$ вкладом заряда в функцию $A(r)$ можно пренебречь, и метрика приближенно совпадает с метрикой Шварцшильда.

Напротив, при $r \rightarrow 0$ свойства метрики полностью определяются зарядом; так как это во всех случаях R-область, значение $r = 0$ соответствует центру, который, как следует из сравнения с формулами (3.7), сингулярен. Поскольку $g_{00} = A \rightarrow \infty$ при $r \rightarrow 0$, нетрудно установить, что центр отталкивает от себя пробные частицы.

Метрики с ненулевой космологической постоянной

При $\Lambda \neq 0$ основные свойства метрик (3.17) по-прежнему определяются поведением функции $A(r)$ и, в первую очередь, количеством и расположением ее нулей, каждый из которых соответствует горизонту, разделяющему R- и T-области.

При $Q = 0$, $M > 0$ и $\Lambda > 0$ (метрика Шварцшильда-де Ситтера) число горизонтов может быть от 0 до 2, но при $r \rightarrow 0$ и $r \rightarrow \infty$ имеют место T-области. Если еще и заряд отличен от ну-

ля, то при $r \rightarrow 0$ имеет место R-область с сингулярным центром, как в решении Райснера — Нордстрема, при $r \rightarrow \infty$ метрика ведет себя как в решении де Ситтера (T-область), а полное число горизонтов — от 1 до 3.

3.3. Горизонты и геодезические в статической сферически-симметричной метрике

После первого знакомства с горизонтами в пространстве де Ситтера, в метриках Шварцшильда, Райснера — Нордстрема и их обобщениях на ненулевую космологическую постоянную имеет смысл уточнить понятие горизонта и его тесную связь с представлениями о черных дырах. Как упоминалось в начале главы, черной дырой (ЧД) называют область пространства-времени, которую не может покинуть никакое материальное тело или световой сигнал. Поскольку движение массивных и безмассовых частиц в метрических теориях гравитации происходит по геодезическим, вопрос о том, какая область пространства-времени является ЧД решается при помощи исследования свойств геодезических.

По этой причине данный раздел начинается с описания общего вида геодезических в произвольном статическом сферически-симметричном пространстве-времени, затем будет обсуждаться поведение геодезических вблизи горизонтов и его связь с понятием черной дыры, а в заключение раздела мы кратко остановимся на разных видах горизонтов, упоминаемых в литературе по теории гравитации.

3.3.1. Общий вид уравнений геодезических

Вернемся к общей статической сферически-симметричной метрике (3.1) с произвольной радиальной координатой u и рассмотрим уравнения геодезических (2.50) как уравнения относительно

неизвестных функций $x^0 = t(\lambda)$, $x^1 = u(\lambda)$, $x^2 = \theta(\lambda)$, $x^3 = \phi(\lambda)$ — координат точки на траектории как функций канонического параметра λ .

Для простоты и в связи с симметрией задачи будем полагать, что геодезическая находится в экваториальной плоскости $\theta = \pi/2$; точкой обозначаем $d/d\lambda$, штрихом d/du . Уравнения и интегралы двух из них имеют вид

$$\ddot{t} + 2\gamma'\dot{t}\dot{u} = 0 \quad \Rightarrow \quad \dot{t} = E e^{-2\gamma}, \quad (3.24)$$

$$\ddot{u} + \gamma' e^{2\gamma-2\alpha}\dot{t}^2 + \alpha'\dot{u}^2 - \beta' e^{2\beta-2\alpha}\dot{\phi}^2 = 0, \quad (3.25)$$

$$\ddot{\phi} + 2\beta'\dot{\phi}\dot{u} = 0 \quad \Rightarrow \quad |\dot{\phi}| = L e^{-2\beta}, \quad (3.26)$$

где E и L — константы интегрирования. Уравнение, содержащее $\dot{\theta}$, тривиально.

Уравнения (3.24)–(3.26) не независимы: существует связь $u_\alpha u^\alpha = k$, где $u^\mu = dx^\mu/d\lambda$; $k = +1$ для временноподобных геодезических (в этом случае u^μ — 4-скорость, а параметр λ совпадает с собственным временем), $k = 0$ для светоподобных и $k = -1$ для пространственноподобных. Выпишем явный вид этой связи, которая представляет собой интеграл уравнения (3.25):

$$e^{2\gamma}\dot{t}^2 - e^{2\alpha}\dot{u}^2 - e^{2\beta}\dot{\phi}^2 = k, \quad (3.27)$$

и подставим в него интегралы (3.24) и (3.26). После домножения на $e^{2\gamma}$ получим:

$$e^{2\alpha+2\gamma}\dot{u}^2 + k e^{2\gamma} + L^2 e^{2\gamma-2\beta} = E^2. \quad (3.28)$$

Последнее соотношение носит характер закона сохранения энергии частицы, движущейся в потенциальном поле вдоль оси u : роль полной энергии играет величина $E^2 \geq 0$, первый член представляет собой аналог кинетической энергии частицы, сумма второго и третьего — аналог потенциальной энергии, причем эффективный потенциал

$$V(u) = e^{2\gamma}(k + L^2 e^{-2\beta}) \quad (3.29)$$

играет для геодезического движения ту же роль, что и потенциал в классической механике одномерного движения материальной точки: движение возможно только в области, в которой $E^2 \geq V(u)$, а значения координаты u , в которых $E^2 = V$, соответствуют точкам поворота.

Постоянная L , связанная с изменением азимутального угла ϕ , может быть интерпретирована как сохраняющийся момент импульса частицы относительно тяготеющего центра.

3.3.2. Горизонты, геодезические и квазиглобальная координата

Для дальнейшего описания мы будем часто использовать квазиглобальную координату ρ (см. раздел 6.1), которая обладает важными преимуществами по сравнению с другими вариантами выбора радиальной координаты.

Так, забегая вперед, укажем, что координата ρ ведет себя на горизонтах так же, как координаты, в которых осуществляется аналитическое продолжение, и потому может использоваться по обе стороны от горизонта (с чем и связано ее название). Кроме того, как мы сейчас убедимся, в некоторых важных случаях поведение координаты ρ может служить критерием геодезической полноты пространства.

С помощью квазиглобальной координаты $u = \rho$ статическая метрика (3.1) записывается в виде

$$ds^2 = A(\rho) dt^2 - \frac{d\rho^2}{A(\rho)} - r^2(\rho) d\Omega^2, \quad (3.30)$$

где $A(\rho) \equiv e^{2\gamma}$.

Понятие горизонта мы ввели выше на примере метрик де Ситтера, Шварцшильда и других, в которых координата $r = \rho$ оказывается одновременно координатой кривизны и квазиглобальной. В соответствии со сказанным, в произвольном пространстве

с метрикой (3.30) горизонт есть регулярная сфера $\rho = \rho_h$, вблизи которой

$$A(\rho) \sim (\rho - \rho_h)^n, \quad (3.31)$$

где $n \in \mathbb{N}$ — порядок горизонта. Легко заметить, что, в терминах метрики (3.30) R-область — это область, где $A > 0$, T-область — область, где $A < 0$, а горизонт — поверхность, где $A = 0$. Горизонт нечетного порядка n отделяет R-область от T-области, а горизонт четного порядка разделяет две R-области или две T-области.⁴

Рассмотрим поведение геодезических метрики (3.30) вблизи горизонтов. Согласно (3.28), в обозначениях метрики (3.30) ($u = \rho$, $e^{2\gamma} = e^{-2\alpha} = A(\rho)$), учитывая, что $\dot{u} = du/d\lambda$, получаем

$$\pm \frac{d\lambda}{d\rho} = \left[E^2 - A \left(L^2/r^2 + k \right) \right]^{-1/2}. \quad (3.32)$$

В частности, при $\rho \rightarrow \rho_h$ (где величина ρ_h может быть конечной или бесконечной) имеем $A \rightarrow 0$ (т.е. это возможный горизонт) и

$$d\lambda \approx E d\rho. \quad (3.33)$$

Если E — конечная константа, то, очевидно, вблизи возможного горизонта координата ρ ведет себя как канонический параметр на любых подходящих к нему геодезических, пространственно-, свето- и временноподобных. Для последних можно утверждать, что *горизонт достигается по геодезической за конечное собственное время $\lambda \equiv \tau$ тогда и только тогда, когда этот горизонт соответствует конечному значению квазиглобальной координаты ρ .*

⁴Нецелые n нарушили бы аналитичность метрики даже при конечных инвариантах кривизны, что обесмыслило бы понятие аналитического продолжения за горизонт. Так, если $\rho > \rho_h$ — статическая область, то при дробных n в области $\rho < \rho_h$ метрика теряет смысл. Хотя при рассуждениях, ограниченных рамками статической СО, можно было бы рассматривать и дробные n , мы для простоты сразу полагаем $n \in \mathbb{N}$.

Если $A \rightarrow 0$ при $\rho \rightarrow \infty$, такая поверхность может быть названа удаленным горизонтом: по геодезической она достигается за бесконечное время, и, по-видимому, никакие негеодезические траектории не могут сделать это время конечным.

Постоянная E равна нулю в единственном случае (см. (3.24)) — для чисто пространственных геодезических, полностью находящихся в пространственном сечении $t = \text{const}$. В этом случае канонический параметр λ , совпадающий с длиной l , определяется интегралом

$$l = \int \frac{du}{\sqrt{A(1 - L^2/r^2)}}. \quad (3.34)$$

Подынтегральное выражение бессмысленно при $r < L$. Следовательно, $r = L$ — минимальное значение r , достигаемое чисто пространственными геодезическими. Таков простой геометрический смысл константы L .

Согласно (3.32), если $E = 0$, то в пределе $A \rightarrow 0$, если $\rho \rightarrow \infty$, то и канонический параметр вдоль геодезических λ стремится к бесконечности, как и при ненулевых значениях E . Таким образом, получаем следующий общий результат для статических сферически-симметричных пространств:

Если $\rho \rightarrow \infty$ на поверхности, где $A \rightarrow 0$ (т.е. на предполагаемом горизонте), то эта поверхность (будем называть ее удаленным горизонтом) есть граница рассматриваемого пространства-времени, недостижимая ни для каких геодезических при конечном значении канонического параметра.

Пространство-время геодезически полно на удаленном горизонте, и продолжения за такой горизонт не требуется.

Вернемся к обычным (не удаленным) горизонтам и посмотрим, как на них ведет себя координатное время t (оно же — время по часам удаленного неподвижного наблюдателя в случае асимптотически плоского пространства) для приближающихся к горизонту временноподобных и световых геодезических. Из

(3.24) и (3.28) для метрики (3.30) получаем

$$\frac{d\rho}{dt} = \pm A \sqrt{1 - V(\rho)/E^2}, \quad (3.35)$$

где $E > 0$, а потенциал имеет вид $V = A(k + L^2/r^2)$. На горизонте потенциал обращается в нуль, и в силу (3.31) из (3.35) получаем

$$t \sim \pm \int \frac{d\rho}{(\rho - \rho_h)^n} \rightarrow \pm\infty. \quad (3.36)$$

Следовательно, *координатное время t бесконечно на горизонте для всех пересекающих его геодезических, кроме чисто пространственных*: оно равно $+\infty$ при движении внутрь и $-\infty$ при движении наружу из-под горизонта. Таким образом, горизонт находится в абсолютном будущем или абсолютном прошлом для любого наблюдателя, находящегося в статической области, и для перехода через горизонт и совместного описания областей по обе стороны горизонта необходим переход к координатам, не принадлежащим статической СО. Более того, уже из данного рассуждения, проведенного целиком в рамках статической СО, следует, что горизонт как предел поверхностей $r = \text{const}$ делится на две части — *горизонт прошлого* и *горизонт будущего*.

Проще всего уравнение (3.35) выглядит для радиально движущихся фотонов: $d\rho/dt = \pm A$. Оно показывает, что *координатная скорость фотонов (как и массивных частиц) с точки зрения любого статического наблюдателя стремится к нулю при приближении к горизонту* и стремится столь быстро, что координатное время t пролета фотона от горизонта или к горизонту бесконечно.

Таким образом, если у данной статической области пространства-времени имеется горизонт будущего, то этот горизонт и область, расположенная за ним, реализуют понятие черной дыры с точки зрения этой статической области, так как, попросту говоря, туда улететь можно, а обратно — из абсолютного будущего —

нельзя в принципе. Аналогично, горизонт прошлого и расположенную за ним область можно назвать *белой дырой*, из которой могут лететь фотоны и любые другие частицы, но в которую из данной статической области попасть нельзя, так как они расположены в абсолютном прошлом.

3.3.3. Горизонты, R- и T-области

Описанные выше горизонты статических (или однородных, если речь идет о T-области) сферически-симметричных пространств являются частными случаями более общих видов горизонтов, обсуждаемых в литературе по теории гравитации. Упомянем некоторые из них, кратко поясняя их смысл, но не углубляясь в тонкости, которые потребовали бы введения большого числа понятий из таких областей математики как дифференциальная и алгебраическая топология и дифференциальная геометрия в целом. Заинтересованный читатель найдет подробный анализ этих понятий и их применение в теории гравитации в книгах [26, 163, 227, 299] и обзорных статьях.

Горизонтом событий (event horizon) называют границу черной дыры, причем оба этих понятия строго определены в асимптотически плоском пространстве-времени, а в других случаях, например в пространствах с ненулевой космологической постоянной, используются естественные обобщения этих понятий. Известно (теорема Пенроуза), что горизонт событий образован световыми геодезическими.

Горизонт видимости (apparent horizon) определяется как двумерная поверхность, выходящие с которой ортогональные световые геодезические имеют нулевую расходимость. В стационарных черных дырах горизонт видимости совпадает с горизонтом событий.

Горизонт Коши есть граница области пространства-времени, в которой эволюция физических полей и распределений материи может быть найдена из решения уравнений движения с начальными данными, заданными на некоторой гиперповерхности

Коши (пространственно- или светоподобной). Возникновение горизонтов Коши чаще всего связано с наличием сингулярностей, которые могут быть источниками воздействий, непредсказуемых в рамках классической теории гравитации.

Горизонт Киллинга — по определению, поверхность, нормаль к которой светоподобна и совпадает с вектором Киллинга. Иными словами, это поверхность, на которой некоторый вектор Киллинга светоподобен. Очевидно, горизонты Киллинга могут иметь место только в пространствах с изометриями, которые и характеризуются векторами Киллинга.

Для статических пространств особую важность представляют горизонты Киллинга, на которых становится световым временноподобный вектор Киллинга $\xi^\mu = (1, 0, 0, 0)$. Это происходит там, где $g_{00} = e^{2\gamma} = 0$. С точки зрения статической СО, на горизонте имеет место сингулярность, хотя 4-кривизна должна оставаться регулярной⁵.

Наконец, иногда пользуются понятием горизонта, связанного с поведением некоторого скаляра. Так, например, в сферически-симметричных пространствах с метрикой (3.1) величина $r(u, t)$ представляет собой скаляр с точки зрения двумерного подпространства, параметризованного координатами u и t . Исходя из поведения функции $r(u, t)$, можно ввести понятия **R-области** и **T-области**, обобщающие ранее введенные на произвольные нестатические метрики. По определению, **R-область** — область пространства-времени, в которой градиент функции $e^\beta = r(u, t)$, рассматриваемой как скаляр в (u, t) -подпространстве, пространственноподобен:

$$r^{;\alpha} r_{;\alpha} = e^{-2\gamma} \dot{r}^2 - e^{-2\alpha} r'^2 < 0. \quad (3.37)$$

В этом случае можно выбрать величину r в качестве пространственной координаты. **T-область** — по определению, область,

⁵В определении горизонта Киллинга неявно предполагается регулярность данной поверхности, иначе не имело бы смысла давать такое определение — мы имели бы дело с сингулярностью, а не с горизонтом.

где $r^{\cdot\alpha}r_{,\alpha} > 0$, следовательно, там можно выбрать r в качестве временной координаты. Далее, **r -горизонт** — поверхность, на которой $r^{\cdot\alpha}r_{,\alpha} = 0$.

Все горизонты, рассматривавшиеся в предыдущих разделах, являются, очевидно, r -горизонтами и горизонтами Киллинга. Горизонтами событий и одновременно горизонтами видимости являются сфера $r = 2m = 2GM$ в метрике Шварцшильда и сфера $r = r_+$ в метрике Райснера — Нордстрема. Сфера $r = r_-$ в метрике Райснера — Нордстрема есть горизонт Коши, как и многие горизонты в метриках с $\Lambda \neq 0$.

3.4. Черная дыра Шварцшильда. Геодезические и глобальное описание

3.4.1. R- и T-области

Как мы уже отмечали, в решении Райснера — Нордстрема — (анти-)де Ситтера, (3.17) радиальная координата является одновременно координатой кривизн и квазиглобальной координатой ($r \equiv \rho$). Это позволяет непосредственно воспользоваться результатами предыдущего раздела при обсуждении свойств решения.

Обратимся к простейшему частному случаю — метрике Шварцшильда (3.20). При $r = 2GM = 2m = r_g$ (это значение r называют гравитационным радиусом, соответствующим массе M , или радиусом Шварцшильда) имеет место простой горизонт, а в области $r < 2GM$ (T-области) координаты r и t меняются ролями: r становится временной координатой, а t — пространственной. T-область описывается метрикой Кантовского — Сакса

$$ds^2 = \frac{dr^2}{|A(r)|} - |A(r)|dy^2 - r^2d\Omega^2, \quad (3.38)$$

где $|A(r)| = 2GM/r - 1$, а пространственная координата y введена вместо t . Если полагать, что r уменьшается от значения $2GM$ до нуля, то однородная модель (3.38) сжимается (коллапсирует)

до нуля в двух угловых направлениях вдоль координатных сфер, а в третьем направлении (y) неограниченно растягивается; конечное состояние $r = 0$ — пространственноподобная сингулярность⁶. Обратный процесс анизотропной эволюции (расширение координатных сфер $\theta = \text{const}$, $\phi = \text{const}$ и сжатие от бесконечности до нуля вдоль оси y) наблюдается при изменении от $r = 0$ до $r = 2GM$.

Метрика (3.38) описывает как расширяющуюся (в смысле роста r), так и сжимающуюся космологическую модель, так как направление хода времени не связано однозначно с ростом или убыванием r : физическое время τ связано с r соотношением

$$\tau = \pm \int dr / \sqrt{A(r)}, \quad (3.39)$$

сходимость которого как при $r \rightarrow 2GM$, так и при $r \rightarrow 0$ говорит о том, что эволюция от горизонта к сингулярности (или наоборот) проходит за конечное собственное время по часам, покоящимся в метрике (3.38) в фиксированной точке пространства (y, θ, ϕ) .

При этом, однако, расширяющаяся или сжимающаяся T -область не связана с исходной R -областью, хотя на разделяющем их горизонте $r = \tau = 2GM$ 4-кривизна не имеет особенности: скаляр Кречмана есть

$$\mathcal{K} = 48(GM)^2 / r^6. \quad (3.40)$$

⁶Иногда встречается утверждение, что пространство-время Шварцшильда имеет сингулярность в центре. Как видно из сказанного, такое утверждение в корне неверно. Сингулярность $r = 0$ находится в T -области, представляющей собой однородную космологическую модель, в которой все пространственные точки равноправны, в то время как термином “центр” обозначают некоторую выделенную точку пространства. Более того, в координатах, в которых записана метрика (3.38), сингулярность достигается всеми точками пространства одновременно.

3.4.2. Геодезические в R-области

Поведение геодезических вблизи горизонта соответствует данному выше общему описанию. Согласно (3.28) и (3.24), в метрике Шварцшильда

$$\begin{aligned} \left(\frac{dr}{d\lambda}\right)^2 + \left(1 - \frac{2m}{r}\right)\left(\frac{L^2}{r^2} + k\right) &= E^2, \\ \frac{dt}{d\lambda} &= \frac{E}{1 - 2m/r}, \quad m := GM. \end{aligned} \quad (3.41)$$

Первое из этих уравнений вблизи горизонта ($r \rightarrow 2m$) дает, независимо от характера геодезической и от углового момента L , $dr/d\lambda \approx \pm E$, откуда (в предположении $E > 0$) $d\lambda \approx \pm dr/E$. Таким образом, *все геодезические пересекают горизонт при конечном значении канонического параметра λ* . В частности, для временноподобных геодезических это означает, что *горизонт пересекается при конечном значении собственного времени s* .

В исключительном случае чисто пространственных геодезических ($E = 0$) из (3.34) получаем:

$L > 2m$: $r = L$ — минимальное значение r , такие геодезические не достигают горизонта.

$L = 2m$: горизонт достигается, но длина $l \sim \int dr/(r - 2m)$ логарифмически расходится. Иными словами, расстояние от любой фиксированной точки до горизонта вдоль такой (касательной) геодезической бесконечно.

$L < 2m$: горизонт достигается, и интеграл (3.34) сходится при $r = 2m$.

Несложно также непосредственно проверить, что 3-кривизна гиперповерхностей $t = \text{const}$ сингулярна на горизонте.

Остановимся на движении массивных частиц в поле Шварцшильда, то есть на временноподобных геодезических ($k = 1$,

$d\lambda = ds$), где ds имеет смысл приращения собственного времени частицы. Из (3.41) получаем

$$\left(\frac{dr}{ds}\right)^2 + V(r) = E^2 \quad (3.42)$$

с потенциалом

$$V(r) \equiv \left(1 + \frac{L^2}{r^2}\right) \left(1 - \frac{2m}{r}\right). \quad (3.43)$$

Взяв производную d/ds , получаем уравнение движения в виде

$$\frac{d^2r}{ds^2} = -\frac{1}{2}V'(r). \quad (3.44)$$

Устойчивые круговые орбиты можно найти из условий $V' = 0$, $V'' > 0$. Подставляя выражение для потенциала и учитывая, что $m = GM$, имеем

$$3L^2GM/r^2 - L/r + GM = 0. \quad (3.45)$$

Решение этого квадратного уравнения дает радиусы стационарных орбит:

$$r_{\text{orb}}(L) = 2 \left(\frac{1}{3GM} \pm \sqrt{\frac{1}{(3GM)^2} - \frac{4}{3L^2}} \right)^{-1}. \quad (3.46)$$

Видно, что стационарные орбиты отсутствуют при малых моментах:

$$L \leq L_{\text{min}} = 2\sqrt{3}GM.$$

Минимальный радиус орбиты —

$$r_{\text{min}} = r_{\text{orb}}(L_{\text{min}}) = 6GM = 3r_g. \quad (3.47)$$

Максимум потенциала находится на радиусе

$$r = r_{\text{max}}(L) = \frac{6GM}{1 + \sqrt{1 - 12(GM/L)^2}}, \quad (3.48)$$

зависящем от момента L . Напоминаем, что в искривленном пространстве радиус координатной сферы r , вообще говоря, не равен расстоянию от центра, если таковой имеется. А в поле Шварцшильда, как мы уже убедились, центра вообще нет, так как значение $r = 0$ относится к Т-области.

3.4.3. Захват частиц черной дырой

Вернемся к уравнению (3.42) и рассмотрим задачу о захвате частицы черной дырой Шварцшильда. Пусть частица движется из бесконечности с “энергией” $E \geq 1$ (где $E = 1$ соответствует нулевой скорости на бесконечности). Если она не захватывается черной дырой, то уходит снова на бесконечность. Значит, должен быть момент наибольшего сближения, когда

$$dr = 0 \Rightarrow V(r) = E^2.$$

Напротив, если захват происходит, то уравнение $V(r) = E^2$, где $V(r)$ задано выражением (3.43), не должно иметь решения.

Для частицы единичной массы, движущейся из бесконечности с начальной скоростью $v_\infty \ll 1$, имеем следующие выражения для энергии и момента импульса на бесконечности: $E_\infty = 1 + v_\infty^2/2 \simeq 1$ и $L = bv_\infty$, где b — прицельный параметр. Мы приходим к алгебраическому уравнению

$$\left(1 + \frac{J^2}{r^2}\right) \left(1 - \frac{r_g}{r}\right) \simeq 1,$$

решение которого отсутствует при

$$L \lesssim 2r_g \Rightarrow b \lesssim 2r_g/v_\infty.$$

Следовательно, сечение захвата в нерелятивистском случае имеет вид

$$\sigma = \pi b^2 = 4\pi r_g^2/v_\infty^2 = 16\pi m^2/v_\infty^2. \quad (3.49)$$

В общем, релятивистском случае условие захвата то же самое — отсутствие решений уравнения (3.42) с производной dr/ds , равной нулю. Имеем по определению момента импульса

$$J \equiv r^2 \frac{d\phi}{ds}.$$

На больших расстояниях от черной дыры $\phi = b/r$, откуда $d\phi = -dr(b/r^2)$ (b — прицельный параметр). Подставляя в уравнение, получаем $dr/ds = L/b$. Значит,

$$E^2 = 1 + \frac{J^2}{b^2}.$$

Подставляя в (3.42), имеем

$$\left(1 + \frac{L^2}{r^2}\right) \left(1 - \frac{r_g}{r}\right) = 1 + \frac{L^2}{b^2}.$$

В ультрарелятивистском пределе, $L/r \gg 1$, $L/b \gg 1$, уравнение упрощается:

$$\frac{b^2}{r^2} \left(1 - \frac{r_g}{r}\right) \simeq 1.$$

Решения этого уравнения не существует при

$$b \leq \frac{3\sqrt{3}}{2} r_g.$$

Следовательно, сечение захвата ультрарелятивистской частицы черной дырой есть

$$\sigma = \pi b^2 = \frac{27}{4} \pi r_g^2. \quad (3.50)$$

В предыдущих выводах не учитывалось излучение гравитационных волн при взаимодействии частицы с черной дырой. В случае взаимодействия двух черных дыр оказывается, что этим явлением пренебрегать нельзя, и оно существенно увеличивает

сечение взаимозахвата черных дыр. Точнее, сечение гравитационного захвата двух черных дыр с массами M_0 и M в тесную пару имеет вид [229]

$$\sigma = 2\pi \left(\frac{85\pi}{6\sqrt{2}} \right)^{2/7} \frac{G^2 (M_0 + M)^{10/7} M_0^{2/7} M^{2/7}}{c^{10/7} v_{\text{rel}}^{18/7}}, \quad (3.51)$$

где v_{rel} — относительная скорость двух ЧД. После захвата черные дыры быстро сливаются за счет излучения гравитационных волн, и в момент слияния производят мощный гравитационный всплеск. Сечение этого механизма слияния двух ЧД превышает сечение прямых столкновений ЧД с учетом гравитационной фокусировки в большом интервале масс и скоростей ЧД.

3.4.4. Глобальное описание — метрика Крускала

Переход от статической метрики Шварцшильда к координатам, обеспечивающим ее полное описание одной картой, был получен в 1960 г. независимо Крускалом (Kruskal) и Секерешем (Szekeres) [204, 289].

Опишем этот переход. В преобразовании нуждается лишь (r, t) -подпространство, так что угловые координаты θ , ϕ остаются прежними.

Прежде всего преобразуем r , переходя к “черепашьей” координате r_* , такой, что интересующая нас двумерная метрика в (r, t) -подпространстве станет конформно-плоской:

$$\begin{aligned} r_* &= \int \left(1 - \frac{2m}{r} \right)^{-1} dr = r + 2m \ln \left| \frac{r}{2m} - 1 \right|, \\ ds_2^2 &= \left(1 - \frac{2m}{r} \right) dt^2 - \left(1 - \frac{2m}{r} \right)^{-1} dr^2 = \\ &= \left(1 - \frac{2m}{r} \right) (dt^2 - dr_*^2). \end{aligned} \quad (3.52)$$

Горизонту соответствует предел $r_* \rightarrow -\infty$.

В метрике (3.52) удобно перейти к световым координатам V, W :

$$2t = V + W, \quad 2r_* = V - W. \quad (3.53)$$

При фиксированном V горизонту соответствует $W \rightarrow \infty$, следовательно, $t \rightarrow \infty$. Наоборот, при фиксированном W горизонту соответствует $V \rightarrow -\infty$, следовательно, $t \rightarrow -\infty$. Таким образом (в согласии с предыдущим рассмотрением геодезических), горизонт находится по отношению к неподвижному наблюдателю либо в абсолютном прошлом, либо в абсолютном будущем и распадается на горизонт прошлого и горизонт будущего. В терминах V и W метрика Шварцшильда имеет вид

$$ds^2 = \left(1 - \frac{2m}{r}\right) dV dW - r^2 d\Omega^2. \quad (3.54)$$

Сингулярность при $r = 2m$ пока не устранена: для этого нужен еще один шаг — преобразование вида $V = V(v), W = W(w)$. Рассмотрим, например, переход через горизонт будущего: фиксируем V и рассматриваем предел $W \rightarrow \infty$. Надо взять такую функцию $W(w)$, чтобы коэффициент при $dV dw$ в метрике стал конечным. Как нетрудно убедиться, этим свойством обладает функция $W = -2m \ln |w|$, так как $1 - 2m/r = (r - 2m)/r \sim r_*/(2m) \sim -w/(2m)$ при $r \rightarrow 2m$. Аналогично происходит регуляризация метрики при переходе через горизонт прошлого. Легко проверить, что при преобразовании

$$V = 2m \ln |v|, \quad W = -2m \ln |w| \quad (3.55)$$

справедливо соотношение

$$1 - \frac{2m}{r} = -vw e^{-r/(2m)}, \quad (3.56)$$

и метрика приобретает регулярный вид

$$ds^2 = \frac{32m^3}{r} e^{-r/(2m)} dv dw - r^2 d\Omega^2. \quad (3.57)$$

Можно перейти от световых координат к пространственной (R) и временной (T) координатам, полагая

$$v = \frac{1}{2}(T + R), \quad w = \frac{1}{2}(T - R), \quad dv dw = dT^2 - dR^2, \quad (3.58)$$

откуда

$$ds^2 = \frac{32m^3}{r} e^{-r/(2m)} (dT^2 - dR^2) - r^2 d\Omega^2. \quad (3.59)$$

Это и есть метрика Крускала, дающая полную картину пространства-времени Шварцшильда. Исходной области $r > 2m$ в ней соответствует квадрант $v > 0$, $w < 0$, или $R > 0$, $-R < T < R$ (область I). Кроме нее, имеются две T-области II и IV и еще одна R-область III (см. рис. 3.1) Сингулярности $r = 0$ соответствуют две гиперболы: $T = -\sqrt{1 + R^2}$ (сингулярность прошлого) и $T = +\sqrt{1 + R^2}$ (сингулярность будущего). Именно в сингулярность будущего падает всч, что падает под горизонт событий $r = 2m$, который на диаграмме Крускала изображается двумя светоподобными прямыми $T = \pm R$.

В силу светового характера координат v и w , радиальные световые геодезические на диаграмме Крускала параллельны осям v или w . Остальные причинные кривые (временноподобные, а также нерадиальные светоподобные) имеют в каждой точке углы наклона меньше 45° от вертикали.

3.4.5. От диаграммы Крускала к диаграмме Картера-Пенроуза для поля Шварцшильда

При исследовании структуры пространств, содержащих горизонты Киллинга, часто используются координатные преобразования крускаловского типа, которые обеспечивают гладкие переходы между областями, разделенными горизонтами. Результаты чаще всего представляются в виде двумерных диаграмм (диаграмм Картера-Пенроуза), вид которых несет основную информацию о глобальной структуре пространства-времени. Такие диаграммы удобнее, чем диаграмма Крускала, так как на них каждая R- или

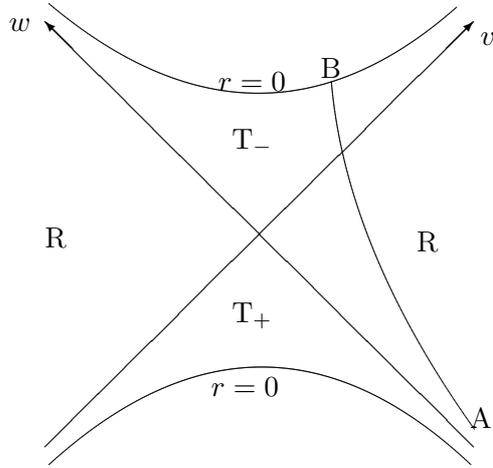


Рис. 3.1. Диаграмма Крускала

T-область изображается квадратом или треугольником конечного фиксированного размера.

Построим такую диаграмму для пространства-времени Шварцшильда, используя метрику Крускала (3.57):

$$ds^2 = \frac{32m^3}{r} e^{-r/(2m)} dv dw - r^2 d\Omega^2, \quad (3.60)$$

где радиус r связан с координатами v и w соотношением (3.56), или

$$-vw = \frac{r - 2m}{2m} e^{r/(2m)}. \quad (3.61)$$

В координатах v , w , т.е. на диаграмме Крускала, горизонт $r = 2m$ изображается парой пересекающихся прямых $v = 0$, $w = 0$, сингулярность $r = 0$ — двумя ветвями гиперболы $vw = 1$. Исходная R-область соответствует квадранту $v > 0$, $w < 0$;

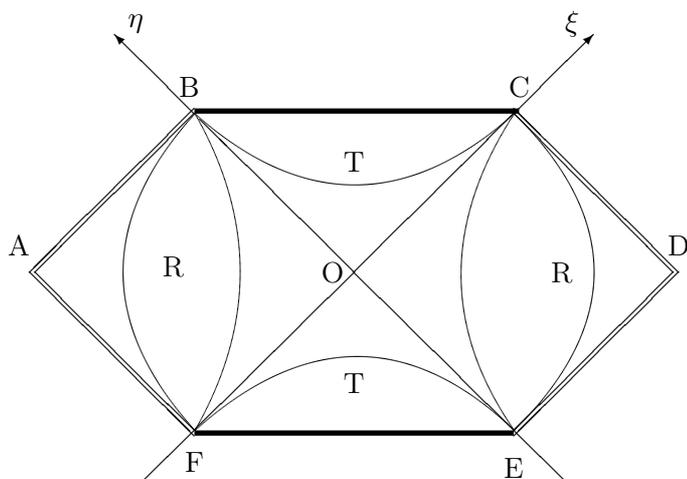


Рис. 3.2. Диаграмма Карттера–Пенроуза для метрики Шварцшильда

квадрант $v < 0, w > 0$ — это еще одна R-область, причинно не связанная с исходной.

Диаграмма становится компактной, если вместо координат $v \in \mathbb{R}, w \in \mathbb{R}$ ввести координаты ξ, η , заданные на отрезке $(\pi/2, \pi/2)$, полагая

$$v = \operatorname{tg} \xi, \quad w = \operatorname{tg} \eta. \quad (3.62)$$

Легко видеть (рис. 3.2), что горизонт теперь описывается двумя отрезками $\xi = 0, \eta = 0$ (BE и CF), сингулярности $r = 0$ отвечают отрезки, на которых $\xi + \eta = \pm\pi/2$ (BC и EF). Значения $\xi = \pm\pi/2$ и $\eta = \pm\pi/2$ (ломаные FAB и CDE) соответствуют бесконечному радиусу r . Всю двумерное (r, t) многообразие отображилось во внутренность шестиугольника $ABCDEF$. Световой характер координат ξ, η позволяет легко различать простран-

ственноподобные, временноподобные и светоподобные направления в любой точке точно так же, как на диаграмме Крускала.

3.5. Глобальная причинная структура пространств с горизонтами

3.5.1. Переход через горизонт в общем случае

Преобразования, аналогичные примененным к метрике Шварцшильда, носят достаточно общий характер, в результате чего диаграммы Картера-Пенроуза можно строить в общем случае двумерных сечений пространства-времени, пользуясь простыми правилами соответствия между поведением метрических функций и графическими образами.

Рассмотрим произвольное статическое или стационарное пространство-время с зависимостью метрики от пространственной координаты u и ограничимся его **двумерным сечением** с метрикой

$$ds_2^2 = e^{2\gamma(u)} dt^2 - e^{2\alpha(u)} du^2 \quad (3.63)$$

(если метрика зависит от других пространственных координат, предполагаем их фиксированными). Пусть поверхность $u = u^*$ — горизонт, следовательно, $e^{\gamma(u^*)} = 0$. Введем две новые "радикальные" координаты x и ρ соотношениями

$$\begin{aligned} dx &= \pm e^{\alpha-\gamma} du \Rightarrow \\ &\Rightarrow ds_2^2 = e^{2\gamma}(dt^2 - dx^2), \quad x \text{ — "черепашня"} \end{aligned} \quad (3.64)$$

$$\begin{aligned} d\rho &= \pm e^{\alpha+\gamma} du, \\ ds_2^2 &= e^{2\gamma} dt^2 - e^{-2\gamma} d\rho^2, \quad \rho \text{ — квазиглобальная.} \end{aligned} \quad (3.65)$$

При $u \rightarrow u^*$ на горизонте $x \rightarrow \pm\infty$ (легко убедиться, что конечность $x(u^*)$ означала бы сингулярность [см. (3.7)], а координата ρ может стремиться как к конечному, так и к бесконечному пределу. Как мы уже убедились в разделе 3.3, если $\rho \rightarrow \pm\infty$ при

$u \rightarrow u^*$, то рассматриваемое пространство-время геодезически полно и без продолжения за горизонт, такой “отдаленный” горизонт является естественной границей пространства-времени, и продолжения не требуется.

Рассмотрим более сложный случай $|\rho(u^*)| < \infty$. Без потери общности предположим, что $\rho(u^*) = 0$, $x \rightarrow -\infty$ и $\rho \rightarrow +0$ при $u \rightarrow u^*$. Кроме того, предположим, что в некоторой конечной окрестности $u = u^*$

$$e^{2\gamma} = \rho^q F(\rho), \quad (3.66)$$

где $q = \text{const} \geq 1$, а $F(\rho)$ — аналитическая функция с конечным значением $F(0)$. Именно такое поведение метрики встречается в большинстве физически интересных примеров.

Для перехода через горизонт введем сначала световые координаты

$$V = t + x, \quad W = t - x. \quad (3.67)$$

Предел $V \rightarrow -\infty$ при конечном фиксированном W соответствует горизонту прошлого (так как $t \rightarrow -\infty$), а предел $W \rightarrow \infty$ при конечном фиксированном V — горизонту будущего. Далее, если ввести новые световые координаты v и w , связанные с V и W зависимостями

$$V = V(v), \quad W = W(w), \quad (3.68)$$

то при некоторых простых требованиях на эти функции смешанная координатная карта (v, W) покрывает горизонт прошлого $V = -\infty$, а карта (V, w) покрывает горизонт будущего $W = +\infty$.

В самом деле, рассмотрим горизонт будущего. Легко проверить, что конечное значение метрического коэффициента g_{Vw} при $w = 0$ достигается, если при $W \rightarrow \infty$

$$w \sim \begin{cases} e^{-WF(0)/2}, & (q = 1), \\ W^{-q/(q-1)}, & (q > 1). \end{cases} \quad (3.69)$$

Продолжение через горизонт будущего соответствует гладкому переходу координаты w через ноль.

С другой стороны, нетрудно убедиться, что при сделанных предположениях при $x \rightarrow -\infty$ справедливо следующее соотношение между ρ и x :

$$\rho \sim \begin{cases} e^{xF(0)/2} & (q = 1), \\ |x|^{-1/(q-1)} & (q > 1). \end{cases} \quad (3.70)$$

Следовательно, согласно (3.66), при переходе через горизонт координата ρ ведет себя точно так же, как w — гладко переходит через нулевое значение. Таким образом, именно этот выбор статической пространственной координаты позволяет описывать область за горизонтом, несмотря на невозможность описания всего пространства-времени в рамках статической СО. Именно по этой причине координата ρ названа квазиглобальной.

В зависимости от четности q , область за горизонтом будет R- или T-областью.

Продолжение метрики через горизонт прошлого происходит совершенно аналогично при помощи карты (v, W) и приводит к тем же результатам.

Замечание. В конкретных задачах часто бывает, что преобразование к квазиглобальной координате во всем пространстве в явном виде провести трудно, например для этого требуется решить трансцендентное уравнение. Заметим в этой связи, что для проведенного рассуждения существенно лишь, чтобы координата ρ удовлетворяла соотношению (3.65) в окрестности горизонта, а в остальном пространстве она остается произвольной.

Мы убедились, что продолжение метрики через горизонт может быть описано в терминах некоторой выделенной координаты (w или ρ), которая при переходе меняет знак. Однако если в представлении (3.66) число q дробное, то в области $\rho < 0$ метрические коэффициенты теряют смысл, следовательно, предполагаемый горизонт в действительности представляет собой *осо-*

бый вид пространственно-временной сингулярности — нарушение аналитичности метрики, не связанный с особенностями алгебраических инвариантов тензора кривизны. Особенностями в этих случаях обладают, очевидно, дифференциальные инварианты, содержащие производные вида $(\nabla_\alpha \nabla_\beta \dots) R_{\mu\nu\rho\sigma}$.

Мы будем иногда называть такие поверхности *сингулярными горизонтами*.

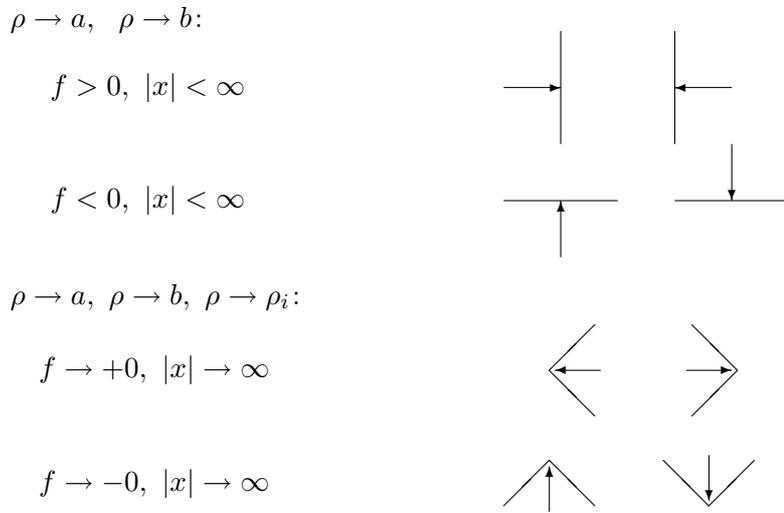
Полученное условие “квантования” $q \in \mathbb{N}$ может оказаться не единственным, если дробные степени ρ появляются в других метрических коэффициентах помимо $g_{\rho\rho}$ и g_{tt} . Такие примеры действительно иногда рассматриваются в литературе.

3.5.2. Построение диаграмм Пенроуза

Как следует из изложенного, для анализа глобальной структуры двумерных лоренцевых поверхностей с метрикой (3.63) удобно воспользоваться координатой ρ , заданной в (3.65) и связанной со световыми координатами, обеспечивающими гладкое продолжение метрики. Итак, пусть при некотором выборе координат метрика двумерного сечения пространства-времени имеет вид

$$ds^2 = f(\rho)dt^2 - \frac{d\rho^2}{f(\rho)}, \quad (3.71)$$

где координата ρ задана в некоторой конечной или бесконечной области $a < \rho < b$ на вещественной прямой. Согласно (3.64), $x = \pm \int d\rho/f(\rho)$. Тогда с помощью рассуждений, подобных приведенным выше при обсуждении глобальных свойств метрики Шварцшильда, нетрудно убедиться, что в некоторой системе координат концевым точкам a , b и нулям функции $f(\rho)$ можно сопоставить следующие графические образы ($\rho = \rho_i$; предполагаем, что нули, не совпадающие с a или b , существуют):



(стрелками отмечено направление, по которому мы приближаемся к a , b или ρ_i). Построение начинается с любой из конечных точек. Каждому интервалу знакопостоянства $f(\rho)$ соответствует внутренняя область некоторого треугольника или квадрата на диаграмме. Для перехода через $\rho = \rho_i$ (горизонт) к каждому из двух имеющихся наклонных отрезков, отвечающих $\rho = \rho_i$, надо построить перпендикулярный ему отрезок той же длины с одного из концов. Этот конец выбирается в зависимости от знака $f(\rho)$ при $\rho < \rho_i$ и $\rho > \rho_i$ (см. рис. 3.3). Новый отрезок отвечает тому же значению $\rho = \rho_i$.

Можно сформулировать удобное правило для продолжений через горизонты: **при RT- и TR-переходах** (т.е. от R- к T-области или наоборот) **старые и новые отрезки образуют X-образную фигуру**, а **при RR- и TT-переходах** — **зигзагообразную фигуру**. Правило объясняется тем, что горизонт с данным значением ρ_i расположен по отношению к R-области на диаграмме либо справа, либо слева (и делится на горизонт прошлого и горизонт будущего), а по отношению к T-области —

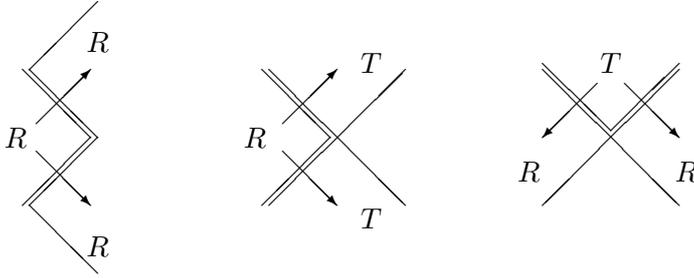


Рис. 3.3. Примеры переходов через горизонты. Двойные наклонные отрезки изображают горизонты исходной области, одинарные отрезки — достраиваемые горизонты, стрелки — направление перехода. Символ R обозначает область $f(\rho) > 0$ (R -область), символ T — область $f(\rho) < 0$ (T -область). Так же обозначены области на последующих рисунках.

либо вверх, либо вниз.

Построение может продолжаться неограниченно или заканчивается, когда диаграмма покрывает конечную область, точки границы которой отвечают значениям $\rho = a$ и $\rho = b$, которые могут быть как конечными, так и бесконечными.

Для каждой пары областей, разделенных наклонным отрезком, существует общая система светоподобных координат типа (3.67).

Сформулированных правил достаточно для построения любых двумерных диаграмм Пенроуза, если метрика приводится к виду (3.71). Для иллюстрации метода приведем примеры диаграмм для метрики Райснера — Нордстрема — де Ситтера (3.17), для которой в обозначениях (3.71)

$$f(r) = 1 - 2m/r + q^2/r^2 - \Lambda r^2/3, \quad (3.72)$$

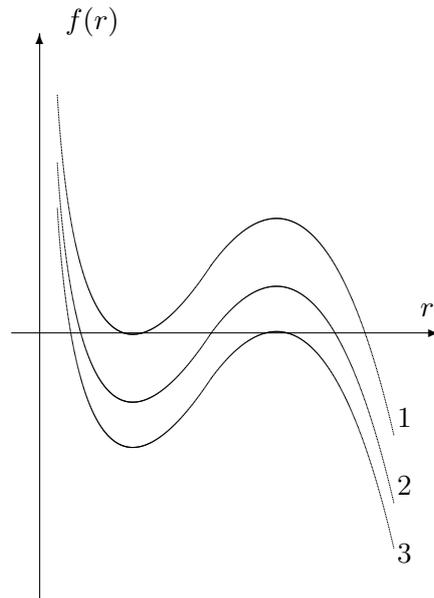


Рис. 3.4. Качественный вид функции $f(r)$ в метрике (3.72) при $0 < \Lambda q^2 < 1/4$ с двумя и тремя нулями.

где m и q — соответственно масса и заряд в геометризованных единицах. Введенная в (3.65) координата ρ здесь совпадает с обычной (шварцшильдовской) радиальной координатой r .

Поведение функции $f(r)$ при $r \in (0, \infty)$ довольно разнообразно и зависит от значений параметров m , q , Λ . На рис. 3.4 показан качественный вид $f(r)$ в случаях, когда она имеет три однократных корня или один однократный и один двукратный, а на рис. 3.5–3.7 — соответствующие диаграммы Пенроуза. В других, более простых случаях, когда $f(r)$ имеет один корень или два однократных корня, структура пространства-времени хорошо известна — см., напр., книги [163, 227]. Пользуясь сформу-

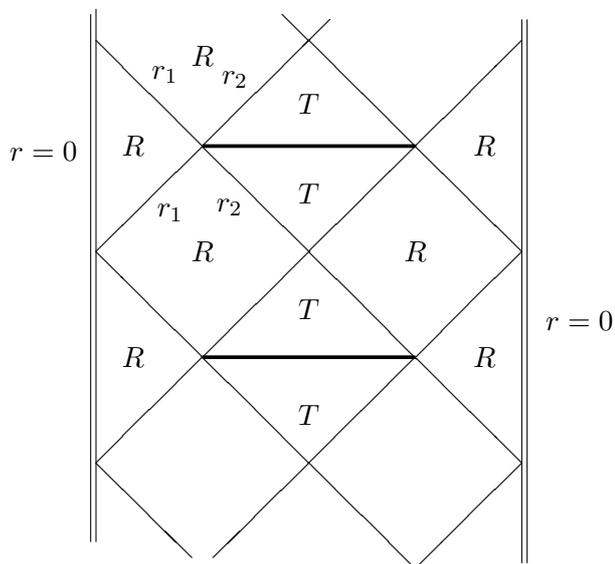


Рис. 3.5. Диаграмма Пенроуза, соответствующая положению 1 кривой $f(r)$ на рис. 3.4. Диаграмма неограниченно продолжаема вверх и вниз. Толстые линии соответствуют $r = \infty$.

лированными правилами, читатель может легко построить соответствующие диаграммы самостоятельно.

На рис. 3.5–3.7 наклонные отрезки изображают горизонты Киллинга, соответствующие которым корни r_i функции $f(r)$ нумеруются в порядке возрастания. Не включенные в диаграмму квадраты на рис. 3.6 и отрезки, изображенные толстыми линиями на рис. 3.5 и 3.7 могут играть роль точек ветвления (их обход не обязан замыкаться в исходной области), так что возможны любые многолистные усложнения этих диаграмм. С другой стороны, возможны не нарушающие причинности отождествления

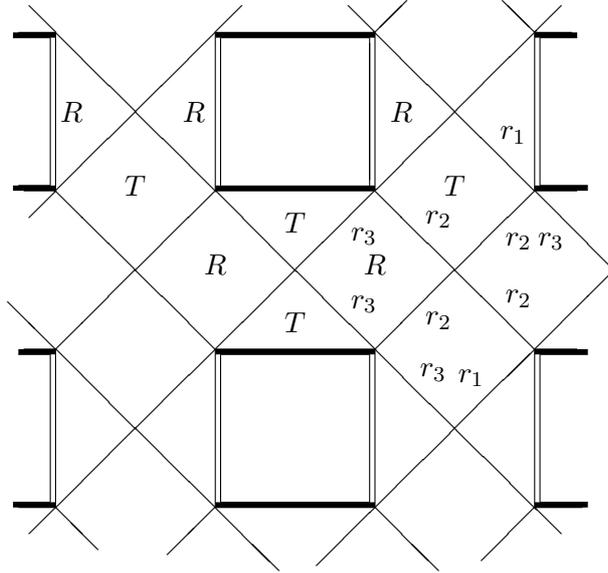


Рис. 3.6. Диаграмма Пенроуза, соответствующая положению 2 кривой $f(r)$ на рис. 3.4. Диаграмма неограниченно продолжаема во все стороны и занимает всю плоскость, кроме квадратов, ограниченных двойными и толстыми линиями. Двойные линии соответствуют $r = 0$, толстые — $r = \infty$.

изометричных гиперповерхностей.

3.6. Черная дыра как результат гравитационного коллапса

Как известно из астрофизики, после выгорания ядерного топлива звезды испытывают катастрофическое сжатие (коллапс) за счет сил гравитации, которые более не могут быть уравновешены

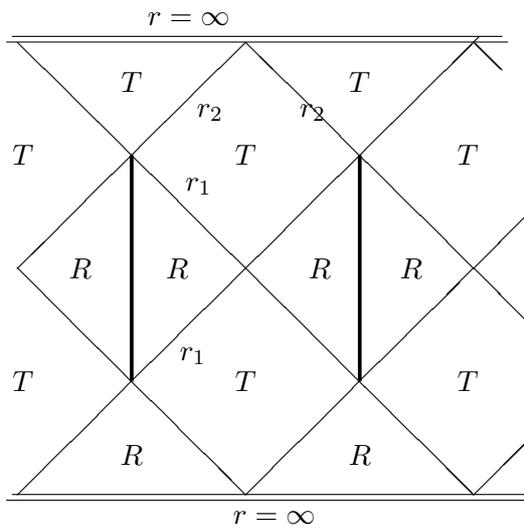


Рис. 3.7. Диаграмма Пенроуза, соответствующая положению 3 кривой $f(r)$ на рис. 3.4. Диаграмма неограниченно продолжаема вправо и влево. Толстые линии соответствуют $r = 0$.

гидродинамическим давлением и давлением излучения. Результат гравитационного коллапса зависит от того, какая масса в нем участвует (это, вообще говоря, не вся звезда, а только внутренние ее слои). При массах меньше предела Чандрасекара (около $1,4 M_{\odot}$, где M_{\odot} — масса Солнца) внутреннее ядро звезды в конце ее эволюции образует белый карлик с плотностью порядка $10^6 \text{ г}\cdot\text{см}^{-3}$. При $m > m_{\odot}$ гравитационный коллапс приводит к возникновению нейтронной звезды с плотностью порядка ядерной ($10^{13} - 10^{15} \text{ г}\cdot\text{см}^{-3}$), а при дальнейшем увеличении масса коллапсирующего ядра превосходит предел устойчивости нейтронных звезд (порядка $3-4 M_{\odot}$, точное значение зависит от уравнения состояния нейтронного вещества и деталей строения

нейтронной звезды, а также от выбора теории гравитации), и сжатие приводит к образованию ЧД. По результатам расчетов, ЧД должна образовываться, если первоначальная масса звезды $\gtrsim 10M_{\odot}$; при этом более половины массы звезды сбрасывается в окружающее пространство, что для удаленного наблюдателя выглядит как взрыв сверхновой.

ЧД меньших масс могли образоваться за счет коллапса флуктуаций плотности в ранней Вселенной.

Расчеты гравитационного коллапса весьма сложны, существует лишь небольшое число точно решаемых моделей (например, коллапс сферически-симметричного облака пыли [291]; см., напр., [23]), в более сложных случаях задачи решаются лишь численно. Есть, однако, некоторые общие, не зависящие от модели особенности коллапса с образованием ЧД.

Главная такая особенность — то, что вне коллапсирующего тела пространство-время остается вакуумным, и к нему применимы все результаты, касающиеся свойств вакуумных решений уравнений гравитации. Более того, если коллапсирует сферически-симметричное тело, весь процесс происходит без нарушения сферической симметрии, и тогда справедлива знаменитая теорема Биркгофа [67],⁷ упрощенная формулировка которой гласит: сферически-симметричное гравитационное поле в вакууме в ОТО с необходимостью статическое. Оно, следовательно, сводится к решению Шварцшильда (или Шварцшильда — де Сит-

⁷После Биркгофа теорема была обобщена на сферические системы с космологической постоянной Λ , системы с электромагнитными и скалярными полями, некоторые скалярно-тензорные теории гравитации и т.д.. Наиболее общий подход к проблеме [3] состоял в исследовании общих условий, при которых можно доказать статичность или однородность системы. Это позволило включить в рассмотрение не только все ранее исследованные частные случаи теоремы в ОТО и СГТ, но и много новых. Теорема была обобщена в двух направлениях: включение новых типов пространственно-временной симметрии (плоская, цилиндрическая, псевдоплоская) и новых видов материальных источников гравитации (скалярные поля, калибровочные поля, идеальная жидкость и т.д.). Многомерное обобщение этого подхода дано в статье [96].

тера, если учтена космологическая постоянная). С более общей точки зрения, теорема указывает случай, когда уравнения поля при определенных условиях индуцируют дополнительную симметрию пространства-времени, не заданную исходно. Например, для Т-области того же пространства-времени Шварцшильда из теоремы вытекает не статичность (которой нет), а однородность — независимость метрики от пространственной координаты, в которую перешла временная координата t статической области при переходе через горизонт $r = 2m$.

Приведем схему доказательства теоремы Биркгофа для самого простого случая сферической симметрии [метрика (3.1)] и вакуума, $T_\mu^\nu \equiv 0$, следовательно, согласно уравнениям Эйнштейна, $R_\mu^\nu \equiv 0$. Предположим, что в рассматриваемой области пространства-времени справедливо неравенство (3.37), т.е. мы находимся в R-области, и можно выбрать $r = e^\beta$ в качестве радиальной координаты. Таким образом, в выражениях (3.3) для компонент тензора Риччи можно положить $\dot{\beta} = 0$. Тогда уравнение $R_{01} = -2\dot{\alpha}\beta' = 0$ немедленно приводит к $\dot{\alpha} = 0$ [$\beta' \neq 0$ в силу (3.37)], и из трех метрических функций может зависеть от времени только γ . В уравнение Эйнштейна $G_1^1 = R_1^1 - \frac{1}{2}R = 0$ функция γ входит только в комбинации $e^{-2\alpha}\beta'\gamma'$, откуда следует, что $\gamma' \equiv d\gamma/dr = \gamma'(r)$, следовательно, $\gamma = \gamma_1(r) + \gamma_2(t)$. Но функцию $\gamma_2(t)$ можно исключить подходящим выбором временной координаты, что приводит к полностью статической метрике. Теорема доказана.⁸

Справедливость теоремы Биркгофа связана с тем, что в ОТО гравитационное поле имеет спин 2 и, следовательно, у него отсутствуют монополярная и дипольная динамические степени сво-

⁸Если в условиях теоремы заменить неравенство (3.37) на противоположное, т.е. предполагать, что мы находимся в Т-области, то повторяя те же рассуждения, мы снова выясним, что метрика не зависит от t , но теперь t — пространственная координата, и независимость от нее означает однородность пространства. Если же полагать градиент r светоподобным, то теорема не доказывается, и при этом условии возможны волновые решения уравнений Эйнштейна.

боды. Иными словами, не существует монопольных и дипольных гравитационных волн.

Для коллапса сферических тел из теоремы Биркгофа следует, что пространство-время вне тела однозначно и на всех временах описывается метрикой Шварцшильда. Из свойств метрики следует, что, с точки зрения удаленного наблюдателя, сжатие поверхности звезды до горизонта происходит бесконечно долго, она как бы застывает на радиусе Шварцшильда $r = 2m$. Однако расчет показывает, что отклонения от $r = 2m$ затухают экспоненциально по времени того же удаленного наблюдателя, с характерным временем порядка $2m$, что составляет около 10 мкс, — время, за которое свет проходит $2m = 3$ км в случае $m = M_{\odot}$ [255]. Сигналы с поверхности коллапсирующей звезды быстро перестают поступать, и она почти мгновенно исчезает из поля зрения удаленного наблюдателя.

На самой поверхности проходит конечный отрезок собственного времени при сжатии от любого конечного радиуса не только до горизонта, но и до сингулярности $r = 0$ (напомним, что это не центральная, как ошибочно пишет ряд авторов, а космологическая сингулярность).

Глобальная структура пространства-времени коллапсирующей звезды существенно отличается от структуры “вечной” ЧД Шварцшильда, изображенной на диаграммах Крускала (рис. 3.1) и Картера-Пенроуза (рис. 3.2). Так, например, на рисунке 3.1 от пустого пространства-времени остается лишь область правее кривой АВ, изображающей мировую линию точки поверхности звезды, а всч, что левее этой кривой, остается неопределенным до тех пор, пока неизвестно динамическое решение уравнений Эйнштейна для внутренности звезды. Остается лишь одна R-область из двух (правая) и одна T-область из двух (верхняя, сжимающаяся).

3.7. Черные дыры и безмассовые скалярные поля

В предыдущих разделах мы убедились, что присутствие безмассового векторного (электрического или магнитного) поля не противоречит существованию ЧД. Иными словами, ЧД может обладать электрическим или/и магнитным зарядом и быть источником соответствующих полей. В данном разделе мы рассмотрим вопрос о том, может ли ЧД обладать также и безмассовым скалярным полем. Обсуждение будем вести при помощи точных решений уравнений поля, которые в этом случае получаются довольно просто и свойства которых представляют самостоятельный интерес.

3.7.1. Общая СТТ и преобразование Вагонера

Рассмотрим свойства сферически-симметричных полей тяготения для двух важнейших вариантов уравнения безмассового скалярного поля — с минимальной связью⁹ ($\psi = \phi^{\min}$) — уравнение $\square \psi = 0$ и с конформной связью ($\phi = \phi^{\text{conf}}$) — уравнение $\square \phi + \frac{1}{6}R\phi = 0$, где R — скалярная кривизна.

Приведем известное конформное преобразование [298], позволяющее формально сводить довольно широкий класс систем со скалярными полями в рамках обобщенной скалярно-тензорной теории (СТТ) Бергмана–Вагонера–Нордтведта к системам со скалярным полем ϕ^{\min} . Рассмотрим лагранжиан обобщенной СТТ

$$L_{\text{СТТ}} = f(\phi)\tilde{R} + h(\phi)\tilde{g}^{\mu\nu}\phi_{,\mu}\phi_{,\nu} + L_m, \quad (3.73)$$

заданной в пространстве \tilde{V}_4 с метрикой $\tilde{g}_{\mu\nu}$ и скалярной кривизной \tilde{R} ; L_m — лагранжиан остальной материи, возможно, также

⁹Выражение “с минимальной связью” означает (см. главу 2), что полный лагранжиан системы есть просто сумма лагранжианов составляющих ее полей.

содержащий зависимость от ϕ . Эта формулировка, соответствующая *конформной калибровке (картине) Йордана*, носит фундаментальный характер в СТТ, так как именно в ней ТЭИ материи T_ν^μ удовлетворяет обычным законам сохранения $\nabla_\nu T_\mu^\nu = 0$, что обеспечивает стандартные уравнения движения в атомной системе измерений. В частности, свободные частицы движутся по геодезическим.

Уравнения поля легче исследовать и решать в так называемой *конформной калибровке (картине) Эйнштейна*, в которой скалярное поле минимальным образом взаимодействует с гравитацией. А именно, конформное преобразование [298]

$$\tilde{g}_{\mu\nu} = g_{\mu\nu}/f(\phi), \quad \frac{d\phi}{d\psi} = f \left| fh + \frac{3}{2} \left(\frac{df}{d\phi} \right)^2 \right|^{-1/2} \quad (3.74)$$

приводит лагранжиан (3.73) (с точностью до полной дивергенции) к лагранжиану типа ОТО с ϕ^{\min} :

$$L = R + \varepsilon g^{\mu\nu} \psi_{;\mu} \psi_{;\nu} + \frac{L_m}{f^2(\phi(\psi))}, \quad \varepsilon := \text{sign} \left[fh + \frac{3}{2} \left(\frac{df}{d\phi} \right)^2 \right] \quad (3.75)$$

для системы со скалярным полем ψ в пространстве с метрикой $g_{\mu\nu}$. Предполагается, что функция $f(\phi)$, используемая в преобразовании (3.74), конечна и положительна во всей интересующей нас области.

В зависимости от $\varepsilon = \pm 1$, СТТ делятся на два класса — нормальные, в которых $\varepsilon = +1$, т.е. кинетическая энергия скалярного поля в картине Эйнштейна положительна, и аномальные, $\varepsilon = -1$, в которых она отрицательна. При всей необычности таких полей, они давно рассматриваются в литературе как чисто теоретический объект, в качестве мысленного эксперимента. В последнее время, однако, к ним проявляется значительный интерес в связи с поисками моделей для так называемой темной

энергии, приводящей к наблюдаемому возрастающему ускорению расширения Вселенной (подробнее см. гл. 5). Скалярные поля с $\varepsilon = -1$ иногда также называют “духовыми” (ghost fields) или “фантомными” (phantom fields).

Частными примерами теории (3.73) является ОТО с минимально связанным скалярным полем ($f \equiv 1$, $h = \varepsilon = \pm 1$) и с конформным скалярным полем $\phi = \phi^{\text{conf}}$:

$$f = 1 - \phi^2/6, \quad h = 1, \quad \varepsilon = +1, \quad \sqrt{6}\phi = \text{th} \frac{\psi + \psi_0}{\sqrt{6}}, \quad (3.76)$$

где ψ_0 — произвольная постоянная.

В инвариантности уравнения

$$\square \phi + \frac{1}{6}R\phi = 0 \quad (3.77)$$

относительно конформного преобразования $\tilde{g}_{\mu\nu} = \Omega^{-2}(x)g_{\mu\nu}$, сопровождаемого преобразованием скалярного поля $\tilde{\phi} = \Omega\phi$, можно убедиться непосредственно, используя формулу соответствующих преобразований для скалярной кривизны и для оператора д’Аламбера

$$\begin{aligned} \tilde{R} &= \Omega^2 R - 6\Omega \square \Omega + 12\Omega^{\cdot\alpha}\Omega_{,\alpha}, \\ \tilde{\square}\phi &= \Omega^2 \square \phi - 2\Omega\Omega^{\cdot\alpha}\phi_{,\alpha}. \end{aligned} \quad (3.78)$$

Легко видеть, что преобразование (3.74) — не то, относительно которого инвариантно уравнение (3.77). Это и позволяет перейти от (3.77) к уравнению $\square \psi = 0$.

3.7.2. Скалярное поле с минимальной связью

Рассмотрим конфигурации с метрикой (3.1) в присутствии безмассового скалярного поля $\psi = \phi^{\text{min}}$ с лагранжианом

$$\frac{1}{2}\varepsilon\psi^{\cdot\alpha}\psi_{,\alpha}, \quad \varepsilon = \pm 1, \quad (3.79)$$

и электромагнитного полем $F_{\mu\nu}$ с лагранжианом

$$L_{\text{em}} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \quad (3.80)$$

Наиболее общий вид электромагнитного потенциала, совместимый со сферической симметрией, есть

$$A_\mu = \delta_\mu^0 A(u) + \delta_\mu^3 q_m \cos \theta + \partial\Phi/\partial x^\mu, \quad (3.81)$$

где $\Phi(x^\mu)$ — произвольная функция, $q_m = \text{const}$ — магнитный заряд. Соответственно этому ТЭИ полей ψ и $F_{\mu\nu}$ имеют вид

$$T_\mu^\nu[\psi] = \frac{1}{2}\varepsilon e^{-2\alpha}\psi'^2 \text{diag}(1, -1, 1, 1); \quad (3.82)$$

$$T_\mu^\nu[F] = 2(\mathbf{E}^2 + \mathbf{B}^2) \text{diag}(1, 1, -1, -1), \quad (3.83)$$

где

$$\mathbf{E}^2 = F^{01}F_{10} = q_e^2/r^4; \quad \mathbf{B}^2 = F^{23}F_{23} = q_m^2/r^4, \quad (3.84)$$

\mathbf{E} и \mathbf{B} — значения напряженностей электрического и магнитного полей (напомним, что последнее в сферической конфигурации существует только при условии существования магнитного монополюсного заряда), причем

$$F^{01} = q_e e^{-\alpha-\gamma}/r^2, \quad F_{10} = q_e e^{\alpha+\gamma}/r^2, \quad (3.85)$$

где q_e — электрический заряд. Обозначим $Q^2 = \frac{1}{2}\varkappa(q_e^2 + q_m^2)$.

Уравнение поля ψ и его первый интеграл имеют вид

$$(e^{\gamma+2\beta-\alpha}\psi')' = 0, \quad \psi' = C e^{\alpha-2\beta-\gamma}, \quad (3.86)$$

где константа C интерпретируется как скалярный заряд.

Тензоры (3.82) и (3.83) обладают свойством $T_1^1 + T_2^2 = 0$. В этом случае удобно использовать гармоническую ($\square u = 0$) радиальную координату u , такую, что метрические коэффициенты в (3.1) удовлетворяет условию (см. раздел 3.1)

$$\alpha = 2\beta + \gamma. \quad (3.87)$$

В этом случае комбинация уравнений Эйнштейна $(1) + (2)$ сводится к уравнению Лиувилля

$$(\beta + \gamma)'' = e^{2\beta+2\gamma} \quad (3.88)$$

и легко интегрируется:

$$e^{-\beta-\gamma} = s(k, u) := \begin{cases} k^{-1} \operatorname{sh} ku, & k > 0; \\ u, & k = 0; \\ k^{-1} \sin ku, & k < 0; \end{cases} \quad (3.89)$$

где k — константа интегрирования; еще одна константа интегрирования устраняется выбором начала отсчета u . В результате приходим к метрике

$$ds^2 = e^{2\gamma} dt^2 - \frac{e^{-2\gamma}}{s^2(k, u)} \left[\frac{du^2}{s^2(k, u)} + d\Omega^2 \right]. \quad (3.90)$$

Без потери общности будем полагать $u \geq 0$, причем $u = 0$ соответствует пространственной бесконечности, на которой метрика асимптотически плоская, и можно полагать $\gamma(0) = 0$ при подходящем выборе масштаба временной координаты t .

На асимптотике ($u \rightarrow 0$) координата u ведет себя как $1/r$, так что

$$r = e^\beta \approx 1/u; \quad e^\alpha \approx 1/u^2; \quad \gamma' \rightarrow \gamma'(0) = -Gm. \quad (3.91)$$

Следует отметить, что очень многие полученные до сих пор точные статические сферически-симметричные решения относятся к системам с $T_1^1 + T_2^2 = 0$ (или к ним сводятся) и, следовательно, могут быть найдены описываемым здесь способом.

Уравнения поля ψ в метрике (3.1) в координатах (3.87) дает

$$\psi = Cu + \psi_0, \quad \psi_0 = \text{const}. \quad (3.92)$$

Из оставшихся уравнений Эйнштейна независимо только одно (что соответствует единственной оставшейся неизвестной $\gamma(u)$), однако имеет смысл выписать два уравнения: компоненту (0) из (3.6), имеющую вид уравнения Лиувилля, и компоненту

(¹₁) из (3.5), не содержащую вторых производных и являющуюся первым интегралом остальных уравнений:

$$e^{2\alpha} R_0^0 = -\gamma'' = -Q^2 e^{-2\gamma}, \quad (3.93)$$

$$e^{2\alpha} G_1^1 = \beta'(\beta' + 2\gamma') = \frac{1}{2}\varepsilon\psi'^2 - Q^2 e^{2\gamma}. \quad (3.94)$$

Интегрируя (3.93), приходим к следующим метрикам:

(а) при $Q \neq 0$:

$$ds^2 = \frac{dt^2}{Q^2 s^2(h, u + u_1)} - \frac{Q^2 s^2(h, u + u_1)}{s^2(k, u)} \left[\frac{du^2}{s^2(k, u)} + d\Omega^2 \right]; \quad (3.95)$$

константы k , h , u_1 , C , Q связаны между собой и с массой m соотношениями

$$\begin{aligned} s^2(h, u_1) &= 1/Gq^2; \\ G(Gm^2 - Q^2) &= h^2 \operatorname{sign} h = k^2 \operatorname{sign} k - \varepsilon\kappa C^2/2, \end{aligned} \quad (3.96)$$

где первое равенство следует из граничного условия $\gamma(0) = 0$, а второе — из подстановки выражения (3.89) для $\beta - \gamma$ и решения уравнения (3.93) в уравнение (3.94);

(б) при $Q = 0$:

$$\begin{aligned} ds^2 &= e^{-2hu} dt^2 - \frac{e^{2hu}}{s^2(k, u)} \left[\frac{du^2}{s^2(k, u)} + d\Omega^2 \right]; \\ Gm &= h; \quad k^2 \operatorname{sign} k = \varepsilon\kappa C^2/2 + h^2. \end{aligned} \quad (3.97)$$

Функция s всюду определяется формулой (3.89).

Решение (3.92)–(3.96) при $C \rightarrow 0$ переходит в электровакуумное решение Райснера-Нордстрема, а при $Q \rightarrow 0$ — в (3.97); в свою очередь, (3.97) при $C \rightarrow 0$ переходит в вакуумное решение Шварцшильда. В других координатах решение (3.97) (естественно, только для нормального знака энергии скалярного поля, $\varepsilon = 1$) было впервые найдено Фишером в 1948 г. [33], решение (3.92)–(3.96) — Пенни в 1969 г. [244].

Кратко рассмотрим свойства решений с $\varepsilon = +1$.

В электронейтральном случае (3.97) координата u задана в области $u > 0$, причем $u \rightarrow \infty$ соответствует голой сингулярности, которая является притягивающей ($g_{00} \rightarrow 0$) при $m > 0$. Аналогично ведет себя метрика (3.95) при $h \geq 0$, т.е. при малых зарядах по сравнению с массой в сопоставимых единицах, $Q^2 \leq Gm^2$. Сингулярность можно характеризовать как скалярно-доминированную, так как на ней $\phi \rightarrow \infty$ и полная энергия скалярного поля E_s бесконечна, а энергия электромагнитного поля E_e конечна. Соответствующая электровакуумная метрика Райснера — Нордстрема имеет горизонты (два при $Q^2 < Gm^2$ и один при $Q^2 = Gm^2$), которые, как мы убеждаемся, исчезают с появлением скалярного поля при сколь угодно малом скалярном заряде.

При $h < 0$ (большой заряд, $Q^2 > Gm^2$) координата u задана в области $0 < u < u_{\max}$, где u_{\max} — ближайший нуль функции $\sin[h|(u + u_1)]$. Значение $u = u_{\max}$ соответствует центральной голой отталкивающей ($e^\gamma \rightarrow \infty$) сингулярности райснер-нордстремовского типа с доминирующим электрическим полем ($\phi < \infty$, $E_s < \infty$, $E_e = \infty$).

Мы видим, что во всех вариантах решение со скалярным полем содержит голую сингулярность. Таким образом, по меньшей мере при условии сферической симметрии безмассовое поле ϕ^{\min} не может быть внешним полем черной дыры. Это утверждение, как мы увидим в дальнейшем, является частным случаем одной из так называемых “теорем об отсутствии волос” у черных дыр, где “волосами”, вслед за Уилером, называют любые внешние характеристики ЧД, кроме массы, электрического и магнитного зарядов и момента импульса.

3.7.3. Скалярное поле с конформной связью

Скалярно-электровакуумное поле тяготения с ϕ^{conf} получается из (3.92)–(3.97) при помощи преобразования (3.74), (3.76); в

результате метрика приобретает конформный множитель, существенно меняющий характер геометрии. Решение имеет вид:

(а) при $q \neq 0$

$$ds^2 = \frac{\text{ch}^2[C(u + u_0)]}{\text{ch}^2 C u_0} \left\{ \frac{dt^2}{Q^2 s^2(h, u + u_1)} - \frac{Q^2 s^2(h, u + u_1)}{s^2(k, u)} \left[\frac{du^2}{s^2(k, u)} + d\Omega^2 \right] \right\}, \quad (3.98)$$

$$\begin{aligned} s^2(h, u_1) &= 1/GQ^2, \quad Q = q \text{ch} C u_0, \\ k^2 \text{sign } k &= h^2 \text{sign } h + 3C^2; \end{aligned} \quad (3.99)$$

электрическое поле и шварцшильдовская масса m определяются согласно (3.84) и сравнением с метрикой Шварцшильда на асимптотике:

$$\begin{aligned} \mathbf{E}^2 &= F^{10} F_{01} = Q^2 e^{-4\beta} = Q^2 (g_{22})^{-2}; \\ [Gm + C \text{th}(C u_0)]^2 &= Q^2 + h^2 \text{sign } h; \end{aligned} \quad (3.100)$$

(б) при $Q = 0$

$$\begin{aligned} ds^2 &= \frac{\text{ch}^2[C(u + u_0)]}{\text{ch}^2 C u_0} \left[e^{-2hu} dt^2 - \frac{k^2 e^{2ku}}{\text{sh}^2 k u} \left(\frac{k^2 du^2}{\text{sh}^2 k u} + d\Omega^2 \right) \right]; \\ Gm &= h - C \text{th} C u_0; \quad k = \sqrt{3C^2 + h^2}, \end{aligned} \quad (3.101)$$

причем скалярное поле $\phi = \phi^{\text{conf}}$ в обоих случаях определяется формулой

$$\phi = \sqrt{6} \text{th}[C(u + u_0)]. \quad (3.102)$$

Отметим, что, в отличие от решения с ϕ^{min} , из-за примененного конформного преобразования здесь не выполняется координатное условие (3.87).

В случае $h < 0$ решение с ненулевым зарядом Q (3.98)–(3.100) задано на отрезке $0 < u < u_{\text{max}} < \infty$, где u_{max} — ближайший нуль функции $\sin[|h|(u + u_1)]$. Значение $u = u_{\text{max}}$ соответствует голой центральной ($r = 0$) отталкивающей ($g_{00} \rightarrow 0$)

сингулярности типа сингулярности Райснера — Нордстрема при $Q^2 > Gm^2$; однако, в отличие от подобного случая решения с полем ϕ^{\min} , здесь расходятся обе энергии E_e и E_s^{conf} . Аналогично ведет себя решение (3.98)–(3.100) при $h \geq 0, u_1 < 0$.

Случай $0 < h < |C|$, $u_1 > 0$ отличается от описанного тем, что решение задано на всей полупрямой $u > 0$, и сингулярность при $u = \infty$ скалярно-доминированная: $r \rightarrow 0$, $e^{2\gamma} = g_{00} \rightarrow \infty$, $E_s^{\text{conf}} = \infty$, $E_e < \infty$. Так же ведет себя решение (3.101), (3.102) при $h < |C|$.

В случае $|h| > |C|$, $u_1 > 0$ оба решения заданы на полупрямой $u > 0$, однако теперь значение $u = \infty$ соответствует притягивающей ($e^{2\gamma} \rightarrow 0$) голой сингулярности с бесконечным сферическим радиусом r типа “пространственного кармана” (термин предложен П. Йорданом). Бесконечность ($u = 0$) и сингулярность ($u = \infty$) разделены горловиной — сферой, где функция $e^\beta = r(u)$ имеет минимум. Энергия $E_e < \infty$ (равна нулю при $q = 0$), тогда как $E_s^{\text{conf}} = -\infty$. Аномальный знак энергии связан с отсутствием положительной определенности у нулевой составляющей ТЭИ.

При $h = |C|$, $u_1 > 0$ сфера $u = \infty$ регулярна, и решение должно быть продолжено за нее. Заметим, что эта регулярная сфера в картине Йордана соответствует сингулярности в картине Эйнштейна, т.е. всц пространство-время картины Эйнштейна отображается лишь на часть пространства-времени картины Йордана (это — пример так называемого конформного продолжения [77]).

Для продолжения решения за сферу $u = \infty$ введем новую радиальную координату

$$y = \text{cth } |Cu|. \quad (3.103)$$

Тогда оба решения с ϕ^{conf} принимают вид

$$E^2 = \sqrt{F^{10}F_{01}} = \frac{|Q|}{r^2} \frac{|Q|y^2}{h^2(y + y_0)^2(y + y_1)^2},$$

$$\begin{aligned}
\phi &= \sqrt{6} \frac{1 + yy_0}{y + y_0}; \\
ds^2 &= (y + y_0)^2 \left[\frac{dt^2}{(y + y_1)^2} - h^2 \frac{(y + y_1)^2}{y^4} (dy^2 + y^2 d\Omega^2) \right]; \\
y_0 &= \text{th } hu_0; \quad y_1 = \text{cth } hu_1.
\end{aligned} \tag{3.104}$$

При $Q = 0$ имеем $y_1 = 1$; пространственной бесконечности соответствует $y = \infty$, где $\phi = \phi_\infty = \sqrt{6}y_0$.

Если $y_0 < 0$, решение определено при $-y_0 < y < \infty$ и имеет голую притягивающую сингулярность в центре ($y \rightarrow -y_0 > 0$; $e^\gamma = e^\beta = 0$), где оба поля ϕ и $F_{\mu\nu}$ сингулярны.

В случае $y_0 > 0$ координата y задана при всех $y > 0$, причем при $y \rightarrow 0$, как и при $y \rightarrow \infty$, метрика становится плоской — имеется вторая пространственная бесконечность, где $e^{2\gamma} \rightarrow (y_0/y_1)^2$ и $\phi \rightarrow (1/y_0)\sqrt{6}$. Таким образом, получена регулярная кротовая нора с горловиной (минимумом r) при $y = \sqrt{y_0 y_1}$. Это решение — **единственное несингулярное среди решений с линейными полями**. Кротовые норы будут подробнее обсуждаться в главе 7.

Наконец, в случае $h = |C|$, $y_0 = 0$ (т.е. $\phi \rightarrow 0$ при $y \rightarrow \infty$) сфера $y = 0$ оказывается горизонтом ($r \rightarrow r_0 = hy_1$, $e^\gamma \rightarrow 0$). Удобно перейти к координатам кривизн: $r = h(y + y_1)$, откуда

$$\begin{aligned}
ds^2 &= (1 - r_0/r)^2 dt^2 - (1 - r_0/r)^{-2} dr^2 - r^2 d\Omega^2; \\
F^{10} &= F_{01} = q/r^2; \quad \phi = C/(r - r_0); \quad r_0 = Gm.
\end{aligned} \tag{3.105}$$

Это решение известно как черная дыра со скалярным зарядом [1, 74]. Метрика имеет тот же вид, что и в случае экстремальной ($Gm^2 = Q^2$) черной дыры Райснера-Нордстрема, к которой решение и сводится в случае $C = 0$. Отметим, что на горизонте ТЭИ конечен, несмотря на сингулярность поля ϕ .

Последующий анализ показал, что ЧД (3.105) неустойчивы относительно малых радиальных возмущений [90].

3.7.4. Аномальные поля. Решение анти-Фишера

В случае аномальных полей — как с минимальной, так и с конформной связью — имеется еще больше вариантов поведения решений, и, в частности, кротовые норы получаются как с минимальными, так и с конформными скалярными полями не в исключительном случае (как метрика (3.104), а в общем случае при $k < 0$. С формальной точки зрения это связано с тем, что в решениях (3.95)–(3.97) можно получить либо $k < 0$ при $h \geq 0$, либо $|k| > |h|$, когда одновременно $h < 0$ и $k < 0$. В этих случаях, при подходящих значениях константы u_1 , решение задано на интервале $0 < u < \pi/|k|$, и на обоих его концах имеется плоская пространственная асимптотика. С физической точки зрения, возникновение кротовых нор связано с нарушением светового энергетического условия $T_{\mu\nu}\xi^\mu\xi^\nu \geq 0$, $\forall\xi^\mu : \xi^\nu\xi_\mu = 0$ — см. главу 4.

В качестве примера рассмотрим свойства скалярно-вакуумного решения (3.101) при $\varepsilon = -1$ — “фантомного” аналога решения Фишера, которое, по аналогии с “анти-де Ситтером” будем называть “решением анти-Фишера”.

а) $k > 0$. В этом случае удобно перейти к квазиглобальной координате ρ преобразованием

$$e^{-2ku} = 1 - 2k/\rho, \quad (3.106)$$

и решение принимает вид

$$\begin{aligned} ds^2 &= \left(1 - \frac{2k}{\rho}\right)^a dt^2 - \left(1 - \frac{2k}{\rho}\right)^{-a} d\rho^2 - \left(1 - \frac{2k}{\rho}\right)^{1-a} \rho^2 d\Omega^2, \\ \phi &= -\frac{C}{2k} \ln\left(1 - \frac{2k}{\rho}\right), \end{aligned} \quad (3.107)$$

где константы связаны соотношением

$$a = h/k, \quad a^2 = 1 + C^2/(2k^2). \quad (3.108)$$

В случае $h < 0$, т.е. $m < 0$, имеем $a < -1$, что приводит, как и в решении Фишера, к отталкивающей центральной сингулярности при $u = \infty$, или $\rho = 2k$.

Ситуация существенно меняется при $m > 0$, или $a > 1$. Действительно, сферический радиус r в этом случае имеет минимум при $\rho = \rho_{\text{th}} = (a + 1)k$, соответствующий горловине размером

$$r(\rho_{\text{th}}) = r_{\text{th}} = k(a + 1)^{(a+1)/2}(a - 1)^{(1-a)/2}, \quad (3.109)$$

и стремится к бесконечности при $\rho \rightarrow 2k$. Более того, при $a = 2, 3, \dots$ метрика имеет горизонт порядка a при $\rho = 2k$ и допускает продолжение к меньшим значениям ρ . Особенностью этих горизонтов является их бесконечная площадь. Такие конфигурации получили название "холодных черных дыр", т.к. все они обладают нулевой температурой Хокинга.

Далее, все составляющие скаляра Кречмана K_i [см.(3.7)] ведут себя как P^{a-2} при $\rho \rightarrow 2k$. Исключением является значение $a = 1$, при котором $C = 0$, $\phi \equiv 0$, и воспроизводится решение Шварцшильда. Таким образом, при $\rho = 2k$ метрика имеет сингулярность кривизны, если $a < 2$ (за исключением $a = 1$), приводит к конечной кривизне при $a = 1$, $a = 2$ и к нулевой кривизне при $a > 2$.

При нецелых значениях $a > 2$ качественное поведение метрики при $\rho \rightarrow 2k$ то же, что и на горизонтах бесконечной площади, но продолжение за них невозможно в силу потери аналитичности функции $(1 - 2k/\rho)^a$ при $\rho = 2k$. Это вид сингулярности, упоминавшийся в разделе 3.5, который может быть назван "сингулярным горизонтом".

б) $k = 0$: решение определено в области $u \in \mathbb{R}_+$ и представляется через квазиглобальную координату $\rho = 1/u$ в виде

$$ds^2 = e^{-2h/\rho} dt^2 - e^{2h/\rho} [d\rho^2 + \rho^2 d\Omega^2], \quad \phi = C/\rho. \quad (3.110)$$

Как и ранее, $\rho = \infty$ — плоская бесконечность, а с другой стороны, при $\rho \rightarrow 0$ поведение различно при положительных и отрицательных массах $m = h/G$. Так, при $m < 0$, $\rho = 0$ — сингулярный центр ($r = 0$ и все K_i бесконечны). Наоборот, при

$m > 0$, $r \rightarrow \infty$ и все $K_i \rightarrow 0$ при $\rho \rightarrow 0$. Это снова сингулярный горизонт: несмотря на исчезающую кривизну, неаналитичность метрики в терминах ρ делает продолжение невозможным.

в) $k < 0$: решение описывает кротовую нору с двумя пространственными асимптотиками при $u = 0$ и $u = \pi/|k|$. Метрика имеет вид

$$\begin{aligned} ds^2 &= e^{-2hu} dt^2 - \frac{k^2 e^{2hu}}{\sin^2(ku)} \left[\frac{k^2 du^2}{\sin^2(ku)} + d\Omega^2 \right] = \\ &= e^{-2hu} dt^2 - e^{2hu} [d\rho^2 + (k^2 + \rho^2) d\Omega^2], \end{aligned} \quad (3.111)$$

где u выражается через квазиглобальную координату ρ , заданную на всей числовой оси \mathbb{R} , по формуле $|k|u = \text{arctg}(\rho/|k|)$. Дальнейшее обсуждение этой метрики см. в главе 7.

Из рассмотрения точных решений с безмассовыми скалярными полями можно сделать следующие основные выводы.

- 1) В ОТО с нормальным скалярным полем ϕ^{\min} ($\varepsilon = +1$) решений с горизонтами нет.
- 2) В ОТО с фантомным скалярным полем ϕ^{\min} ($\varepsilon = -1$) есть решения, содержащие горизонты бесконечной площади при специальных дискретных значениях параметров. Горизонтов конечной площади нет.
- 3) В ОТО с конформным скалярным полем ϕ^{conf} , с нормальным знаком кинетической энергии ($\varepsilon = +1$) есть специального вида неустойчивое решение в виде экстремальной (с горизонтом второго порядка) ЧД.
- 4) Другие виды скалярных полей с неминимальной связью: при $\varepsilon = +1$ в принципе возможны, но неизвестны случаи, подобные п. 3; при $\varepsilon = -1$ ситуация подобна п. 2.

3.8. Скалярные поля с произвольными потенциалами. Теоремы запрета

3.8.1. Для чего нужны теоремы запрета

В большинстве приложений необходимо рассматривать более общие скалярные поля по сравнению с предыдущим разделом — поля с ненулевыми потенциалами. В то же время уравнения для самогравитирующих скалярных полей с нетривиальными потенциалами могут быть явно проинтегрированы лишь в весьма редких случаях, даже для систем с высокой симметрией, таких, как изотропные модели в космологии и статические сферически-симметричные системы.

В связи с этим большую ценность представляют общие выводы, или теоремы о свойствах подобных систем, которые могут быть получены без полного решения полевых уравнений.

Для статических сферически-симметричных скалярно-вакуумных конфигураций в ОТО, описываемых действием

$$\begin{aligned} S &= \int d^4x \sqrt{|g|} [\mathcal{R}/\kappa^2 + L_{\text{sc}}], \\ L_{\text{sc}} &= g^{\mu\nu} \phi_{,\mu} \phi_{,\nu} - 2V(\phi), \end{aligned} \quad (3.112)$$

где \mathcal{R} — скалярная кривизна, а κ^2 — гравитационная постоянная, среди таких теорем известны:

- A.** Теоремы “no-hair” ([40, 63]), утверждающие, что асимптотически плоские черные дыры с нетривиальным скалярным полем не могут существовать при некоторых условиях на потенциал $V(\phi)$.
- B.** Обобщенная теорема Розена, утверждающая, что частицеподобное решение (ЧПР) (т.е. асимптотически плоское решение с регулярным центром) невозможно при $V \geq 0$ [100].
- C.** Теорема несуществования регулярных решений без центра (например, кротовых нор) [76].

D. Теорема о глобальной структуре [76], утверждающая, что список возможных типов глобальной причинной структуры пространства-времени (и соответствующие диаграммы Картера — Пенроуза) для конфигураций с произвольными потенциалами $V(\phi)$ и с любой пространственной асимптотикой не отличается от аналогичного списка в случае $\phi = \text{const}$, а именно, имеются структуры: Минковского (или АдС), Шварцшильда, де Ситтера и Шварцшильда — де Ситтера.

Ряд известных точных решений представляют собой примеры конфигураций, которые допускаются этими теоремами: чернотырные решения со скалярными полями, $V(\phi) \geq 0$, но без плоской асимптотики, асимптотически плоские ЧД и ЧПР с отрицательными потенциалами $V(\phi)$. Всч это вместе взятое дает достаточно ясную картину того, чего можно и чего нельзя ожидать от статических полей с минимальной связью в ОТО. Подробное описание этих результатов выходит за рамки нашего изложения — см., напр., [86] и приведенные там ссылки.

Имеется много обобщений действия (3.112), для которых представляет интерес знать, выполняются ли, и если да, то при каких дополнительных условиях, утверждения А–D. В частности:

- 1) Многомерные конфигурации в ОТО, с лагранжианом скалярного поля, приведенном в (3.112), и метрикой

$$ds^2 = e^{2\gamma} dt^2 - e^{2\alpha} du^2 - e^{2\beta} d\Omega_{d_0}^2, \quad (3.113)$$

где α , β , γ — функции радиальной координаты u и $d\Omega_{d_0}^2$ — линейный элемент на d_0 -мерной сфере единичного радиуса.

- 2) Более общие лагранжианы скалярного поля в ОТО, например, $L_{\text{sc}} = F(\phi, I)$, где $I = g^{\mu\nu} \phi_{,\mu} \phi_{,\nu}$ и F — произвольная функция двух переменных.

- 3) Наборы скалярных полей $\vec{\phi} = \{\phi^K\}$, $K = \overline{1, N}$ σ -модельного типа, рассматриваемые как координаты в N -мерном пространстве целей \mathbb{T}_ϕ , так что

$$L_{\text{sc}} = H_{KL}(\vec{\phi})(\partial\phi^K, \partial\phi^L) - 2V(\vec{\phi}) \quad (3.114)$$

где метрика пространства целей H_{KL} (обычно она предполагается положительно-определенной) и потенциал V — функции N переменных ϕ^K , и мы пользуемся обозначением

$$(\partial y, \partial z) = g^{\mu\nu} \partial_\mu y \partial_\nu z. \quad (3.115)$$

- 4) Скалярно-тензорные теории (СТТ) гравитации, с D -мерным действием

$$S_{\text{СТТ}} = \int d^D x \sqrt{g} [f(\phi)\mathcal{R} + h(\phi)(\partial\phi)^2 - 2U(\phi)], \quad (3.116)$$

где $(d\phi)^2 = (\partial\phi, \partial\phi)$ и f, h, U — произвольные функции скалярного поля.

- 5) Теории гравитации с высшими производными (нелинейные по кривизне), так что скалярная кривизна \mathcal{R} в (3.112) заменяется функцией $f(\mathcal{R})$, и более общие, включающие, например, инвариант $R_{\mu\nu}R^{\mu\nu}$.
- 6) Более общие многомерные конфигурации на многообразиях в виде произведений типа

$$\mathbb{M}^D = \mathbb{R}_u \times \mathbb{M}_0 \times \mathbb{M}_1 \times \mathbb{M}_2 \times \cdots \times \mathbb{M}_n, \quad (3.117)$$

где $\mathbb{M}_{\text{ext}} = \mathbb{R}_u \times \mathbb{M}_0 \times \mathbb{M}_1$ — “внешнее” многообразие с метрикой (3.113), $\mathbb{R}_u \subseteq \mathbb{R}$ — область определения радиальной координаты u , \mathbb{M}_1 — ось времени, $\mathbb{M}_0 = \mathbb{S}^{d_0}$. Далее, $\mathbb{M}_2, \dots, \mathbb{M}_n$ — “внутренние” фактор-пространства произвольных размерностей d_i , $i = 2, \dots, n$; согласно этому обозначению, $\dim \mathbb{M}_0 = d_0$ и $\dim \mathbb{M}_1 = d_1 = 1$.

Можно продолжать этот список и (или) объединять его пункты, получая новые обобщения.

Мы здесь рассмотрим сферически-симметричные системы в ОТО в пространстве-времени произвольной размерности D с достаточно общим σ -модельным источником.

3.8.2. Основные уравнения

Рассмотрим D -мерную ОТО с набором скалярных полей (3.114). Как и в четырехмерном случае, уравнения Эйнштейна можно записать в двух эквивалентных видах:

$$G_\mu^\nu \stackrel{\text{def}}{=} R_\mu^\nu - \frac{1}{2}\delta_\mu^\nu R = -\varkappa^2 T_\mu^\nu, \quad (3.118)$$

или

$$R_\mu^\nu = -\varkappa^2 \tilde{T}_\mu^\nu \stackrel{\text{def}}{=} -\varkappa^2 \left(T_\mu^\nu - \frac{\delta_\mu^\nu}{D-2} T^\alpha_\alpha \right), \quad (3.119)$$

где T_μ^ν — тензор энергии-импульса (ТЭИ), который для полей (3.114) имеет вид

$$T_\mu^\nu = \partial^\nu \vec{\phi} \partial_\mu \vec{\phi} - \frac{1}{2} \delta_\mu^\nu L_{\text{sc}}, \quad (3.120)$$

или, в варианте “с тильдами”,

$$\tilde{T}_\mu^\nu = \partial^\nu \vec{\phi} \partial_\mu \vec{\phi} - \delta_\mu^\nu V(\vec{\phi}) \quad (3.121)$$

где стрелки обозначают скалярное произведение в пространстве целей \mathbb{T}_ϕ : $\vec{a}\vec{b} = H_{KL}a^Kb^L$.

Статическая сферически-симметричная метрика (3.1) записана с произвольной радиальной координатой u . Для наших целей удобно использовать квазиглобальную координату $u = \rho$, соответствующую калибровке $\alpha + \gamma = 0$, так что метрика принимает вид

$$ds^2 = A(\rho)dt^2 - \frac{du^2}{A(\rho)} - r^2(\rho)d\Omega_{d_0}^2 \quad (3.122)$$

где обозначено $r(\rho) = e^\beta$ и $A(\rho) = e^{2\gamma} \equiv e^{-2\alpha}$. Как упоминалось выше (см. раздел 3.3), этот выбор предпочтителен для рассмотрения горизонтов Киллинга, которые описываются нулями функции $A(\rho)$. Причина заключается в том, что в малой окрестности горизонта координата ρ изменяется (с точностью до постоянного множителя) как заведомо допустимая координата типа Крускала, используемая для аналитического продолжения метрики.

При этом выборе координат уравнение скалярного поля и четыре различные компоненты уравнений (3.119) можно записать следующим образом:

$$\left[Ar^{d_0} H_{KL}(\phi^L)' \right]' = r^{d_0} \frac{\partial V}{\partial \phi^K}; \quad (3.123)$$

$$(A'r^{d_0})' = -(4/d_0)r^{d_0} \varkappa^2 V; \quad (3.124)$$

$$d_0 r''/r = -\varkappa^2 (\vec{\phi}')^2; \quad (3.125)$$

$$A(r^2)'' - r^2 A'' + (d_0 - 2)r'(2Ar' - A'r) = 2(d_0 - 1); \quad (3.126)$$

$$d_0(d_0 - 1)(1 - Ar'^2) - d_0 A' r r' = -Ar^2 (\vec{\phi}')^2 + 2r^2 V. \quad (3.127)$$

Уравнения (3.124), (3.125) и (3.126) — компоненты $\binom{t}{t}$, $\binom{t}{t} - \binom{\rho}{\rho}$ и $\binom{t}{t} - \binom{\theta}{\theta}$, соответственно, уравнений (3.119), а (3.127) — компонента $\binom{\rho}{\rho}$ уравнений (3.118). Здесь выписано $(N + 4)$ уравнений для $(N + 2)$ неизвестных ϕ^K , A и r ; однако среди уравнений (3.124)–(3.127) есть лишь два независимых, в частности, (3.127) — первый интеграл остальных уравнений. Система уравнений, таким образом, является определенной.

3.8.3. Теоремы о глобальной структуре

Первая теорема относится к вопросу о существовании кротовых нор, рогов (horns) и трубок с потоком (flux tubes). *Кротовая нора* — по определению, конфигурация, содержащая по меньшей мере

один регулярный минимум функции $r(\rho)$ (обсуждение см. в главе 4). *Трубка с потоком* — конфигурация с $r = \text{const} > 0$, представляющая собой статический $(d_0 + 1)$ -мерный цилиндр. *Рог* — конфигурация, переходящая в трубку с потоком на одной из своих асимптотик, это происходит, если $r(\rho) \rightarrow \text{const} > 0$ на одном из концов области определения ρ . Подобные “рогатые частицы” (“horned particles”) с плоской второй асимптотикой рассматривались как возможные остатки от испарения черных дыр [51].

Теорема о конфигурациях без центра. Уравнения (3.123)–(3.127) для $D \geq 4$ и положительно-определенной метрики H_{KL} не допускают (i) решений, где функция $r(\rho)$ имеет регулярный минимум, (ii) решений, содержащих рога, и (iii) решений в виде трубок с потоком и $\phi^K \neq \text{const}$.

Доказательство (не будем излагать его подробно) опирается на уравнение (3.125), из которого следует $r'' \leq 0$. Это неравенство фактически выражает световое энергетическое условие, выполняющееся для ТЭИ T_μ^ν , если матрица H_{KL} положительно-определенная. В результате невозможны не только кротовые норы как глобальные конфигурации, но даже горловины.

Другая теорема относится к возможному числу и порядку горизонтов Киллинга, совпадающих с числом и порядком нулей функции $A(\rho)$. Напомним, что простой (первого порядка) горизонт, так же, как горизонт любого нечетного порядка, отделяет статическую область, где $A > 0$ (называемую также R-областью), от нестатической области, где $A < 0$ и где (3.113) есть однородная космологическая метрика типа Кантовского — Сакса (T-область). Горизонт четного порядка разделяет области с одинаковым знаком $A(\rho)$.

Расположение горизонтов однозначно определяет глобальную структуру пространства-времени (с точностью до отждествления изометричных поверхностей, если таковые есть). Следующая теорема существенно ограничивает возможные расположения горизонтов.

Теорема о глобальной структуре. Рассмотрим решения уравнений (3.123)–(3.126) для $D \geq 4$. Пусть имеется статическая область $a < \rho < b \leq \infty$. Тогда:

- (i) все горизонты простые;
- (ii) нет горизонтов при $\rho < a$ и при $\rho > b$.

Для доказательства используются свойства уравнения (3.126), которое можно переписать в виде

$$r^4 B'' + (d_0 + 2)r^3 r' B' = -2(d_0 - 1) \quad (3.128)$$

где $B(\rho) = A/r^2$. В точках, где $B' = 0$, получается $B'' < 0$, следовательно, $B(\rho)$ не может иметь регулярного минимума. Таким образом, при движении по оси ρ направо или налево, функция $B(\rho)$ (и, следовательно, $A(\rho)$), однажды став отрицательной, не может вернуться к нулю и к положительным значениям.

Теорема о глобальной структуре говорит о том, что вокруг статической области не может существовать более двух простых горизонтов. Может возникнуть горизонт второго порядка между двумя Т-областями, но тогда этот горизонт единственный, и данная конфигурация вообще не содержит статической области.

Таким образом, возможное расположение нулей функции $A(\rho)$, а следовательно, и набор возможных глобальных причинных структур — те же, что и в случае вакуума с космологической постоянной, который соответствует решению уравнений (3.123)–(3.126) при $\phi^K = \text{const}$, $\varkappa^2 V = \Lambda = \text{const}$, с метрикой

$$ds^2 = A(r)dt^2 - \frac{dr^2}{A(r)} - r^2 d\Omega_{d_0}^2, \quad (3.129)$$

$$A(r) = 1 - \frac{2m}{(d_0 - 1)r^{d_0 - 1}} - \frac{2\Lambda r^2}{d_0(d_0 + 1)}. \quad (3.130)$$

Это — многомерное решение Шварцшильда — де Ситтера (или Тангерлини — де Ситтера). Его частные случаи соответствуют

решениям Шварцшильда ($d_0 = 2$) и Тангерлини ($d_0 \geq 2$)¹⁰ при $\Lambda = 0$ и решению де Ситтера в произвольном числе измерений при $m = 0$ (анти-де Ситтера, если $\Lambda < 0$. При $\Lambda > 0$, если величина m положительна, но меньше критического значения

$$m_{\text{cr}} = \frac{d_0 - 1}{d_0 + 1} \left[\frac{d_0(d_0 - 1)}{2\Lambda} \right]^{(d_0-1)/2}, \quad (3.131)$$

имеется два горизонта, один из которых, при меньшем значении r , интерпретируется как горизонт черной дыры, а другой — как космологический горизонт. Если $m = m_{\text{cr}}$, два горизонта сливаются, и пространство-время содержит две однородные нестатические области, разделенные двойным горизонтом. Решение с $m > m_{\text{cr}}$ — чисто космологическое и горизонтов не содержит. В случаях $m < 0$ и/или $\Lambda < 0$ имеется (как максимум) один простой горизонт. Все диаграммы Картера — Пенроуза для этих случаев хорошо известны, и мы здесь их приводить не будем.

В $(2+1)$ -мерной гравитации ($d_0 = 1$) список глобальных структур еще короче: возможно не более одного простого горизонта.

Обе теоремы из этого раздела весьма универсальны: они не зависят от вида потенциала скалярного поля и от каких бы то ни было предположений о пространственной асимптотике. Первая теорема справедлива в предположении положительной определенности метрики H_{KL} пространства целей (иными словами, все скалярные поля ϕ^K должны быть нормальными), а вторая не требует и этого и справедлива, даже если все или часть полей ϕ^K фантомные.

¹⁰Масса M в обычных единицах, например в граммах, получается, если написать $m = GM$, где G — $(d_0 + 2)$ -мерный аналог постоянной Ньютона. Коэффициент при m выбран в (3.132) и соответственно в (3.130) таким образом, что при больших r в случае $\Lambda = 0$, когда пространство-время асимптотически плоское, покоящаяся пробная частица испытывает ньютоновское ускорение, равное $-GM/r^{d_0}$.

3.8.4. Теоремы “no-hair”

Рассмотрим асимптотически плоское пространство-время, т.е. в метрике (3.129) при больших ρ , без потери общности, $r \approx \rho$ и функция $A(\rho) \approx A(r)$ принимает вид, характерный для метрики Тангерлини — (3.130) при $\Lambda = 0$:

$$A(r) = 1 - \frac{2m}{(d_0 - 1)r^{d_0-1}} \quad (3.132)$$

при $\rho \rightarrow \infty$. Тогда из уравнений поля следует, что компоненты ТЭИ и величины V и $(\vec{\phi})^2$ уменьшаются при больших $\rho \approx r$ быстрее чем $r^{-(d_0+1)}$.

Докажем следующую теорему “no-hair”, распространяющую на нашу систему теорему, известную в ОТО в четырех измерениях [40].

Теорема о внешнем скалярном поле черной дыры. *Единственное асимптотически плоское решение уравнений (3.123)–(3.127) при $D \geq 4$, с положительно-определенной матрицей $H_{KL}(\vec{\phi})$ и $V(\vec{\phi}) \geq 0$, описывающее черную дыру, характеризуется $\vec{\phi} = \text{const}$ и метрикой Тангерлини (3.129), (3.132) во всей области $h < \rho < \infty$, где $\rho = h$ — горизонт событий.*

На горизонте событий $\rho = h$, по определению, $A = A(h) = 0$, и $A > 0$ при $\rho > h$. По теореме о глобальной структуре, горизонт должен быть простым, следовательно, $A \sim \rho - h$ при $\rho \rightarrow h$. Рассмотрим функцию

$$\mathcal{F}_1(\rho) = \frac{r^{d_0}}{r'} [2V - A\vec{\phi}'^2]. \quad (3.133)$$

Легко убедиться, что

$$\mathcal{F}'(\rho) = \mathcal{F}_2(\rho) \stackrel{\text{def}}{=} r^{d_0-1} \left[2d_0V + (d_0 - 1) \frac{\vec{\phi}'^2}{r'^2} + A\vec{\phi}'^2 \right]. \quad (3.134)$$

Для этого при вычислении \mathcal{F}'_1 следует подставить $\vec{\phi}''$ из (3.123), r'' из (3.125) и A' из (3.127). Проинтегрируем (3.134) от h до

бесконечности:

$$\mathcal{F}_1(\infty) - \mathcal{F}_1(h) = \int_h^\infty \mathcal{F}_2(\rho) d\rho. \quad (3.135)$$

Поскольку $r'(\infty) = 1$ и $r'' \leq 0$, имеем $r' > 1$ во всей области определения ρ , но $r'(h) < \infty$. В самом деле, регулярность на горизонте подразумевает конечное значение скаляра Кречмана (3.7) и, следовательно, конечные значения всех его составляющих. Для нас здесь представляет интерес выражение $A'r'$. Поскольку $A'(h) > 0$ (горизонт простой в силу теоремы о глобальной структуре), его конечность означает $r'(h) < \infty$.

Величина $\mathcal{F}_1(h)$ должна быть конечной, т.к. в противном случае либо V , либо $A\vec{\phi}^2$ стремились бы к бесконечности, приводя к бесконечным значениям компонент ТЭИ (см. (3.120) и через уравнения Эйнштейна, к сингулярности кривизны.

Если, однако, допустить ненулевое значение $A\vec{\phi}^2$ при $\rho = h$, то интеграл в (3.135) разойдется при $\rho = h$ за счет второго члена в скобках в (3.134), и это, в свою очередь, приведет к бесконечному значению $\mathcal{F}_1(h)$. Поэтому $A\vec{\phi}^2 \rightarrow 0$ при $\rho \rightarrow h$ и

$$\mathcal{F}_1(h) = \frac{2r^{d_0}(h)}{r'(h)} V(h) \geq 0.$$

С другой стороны, $\mathcal{F}_1(\infty) = 0$ в силу асимптотической плоскостности. Таким образом, левая часть уравнения (3.135) неположительна, а правая неотрицательна. Единственный способ удовлетворить (3.135) — это положить $V \equiv 0$ и $\vec{\phi} \equiv 0$ во всей области $\rho > h$, и единственное решение для метрики имеет вид метрики Тангерлини. Теорема доказана.

Существует ряд других теорем аналогичного характера, общий смысл которых заключается в том, что любая стационарная (а не только статическая сферически-симметричная) ЧД полностью и однозначно характеризуется решением уравнений Эйнштейна, определяемым параметрами M (масса), Q (электрический заряд) и J (угловой момент). Именно это подразумевается

в высказывании Уилера [307] “Black holes have no hair” (“черные дыры не имеют волос”). Имеются аналоги подобных теорем для некоторых обобщений ОТО, а также для не асимптотически плоских конфигураций.

До настоящего момента мы обсуждали лишь сферически-симметричные ЧД, угловой момент которых равен нулю. В следующем разделе мы очень кратко опишем основные свойства вращающихся ЧД. В последующих разделах этой главы будем полагать ЧД электрически незаряженными как из соображений простоты, так и в связи с малой ролью возможного заряда ЧД в астрофизических условиях.

3.9. Черные дыры с вращением

ЧД с вращением образуются при коллапсе вращающихся тел — а таковы почти все небесные тела — а также при слиянии ЧД друг с другом. Следовательно, ЧД с вращением гораздо ближе к реальности, чем сферически-симметричные.

ЧД с вращением описывается метрикой Керра [185]

$$ds^2 = \left(1 - \frac{2mr}{\rho^2}\right) dt^2 + \frac{4mr a^2 \sin^2 \theta}{\rho^2} d\phi dt - \frac{\rho^2}{\Delta} dr^2 - \rho^2 d\theta^2 - \left(r^2 + a^2 + \frac{2mr a^2}{\rho^2}\right) \sin^2 \theta d\phi^2, \quad (3.136)$$

где

$$a = J/m, \quad \rho^2 = r^2 + a^2 \cos^2 \theta, \quad \Delta = r^2 - 2mr + a^2. \quad (3.137)$$

Здесь (t, r, θ, ϕ) — так называемые координаты Бойера — Линдквиста [72] (первоначально Керр получил свою метрику в менее удобных обозначениях). Метрика (3.136) представляет собой асимптотически плоское (при $r \rightarrow \infty$) стационарное аксиально-симметричное решение уравнений Эйнштейна в вакууме ($T_\mu^\nu \equiv 0$). Стационарность и аксиальная симметрия определяются существованием векторов Киллинга $\xi^\mu = (1, 0, 0, 0)$

и $\eta^\mu = (0, 0, 0, 1)$, соответственно (порядок перечисления координат обычный: t, r, θ, ϕ). Стационарность, а не статичность метрики проявляется в ее неинвариантности относительно замены $t \mapsto -t$.

Решение параметризовано двумя константами — массой m и угловым моментом J ; при $J = 0$ оно превращается в решение Шварцшильда.

Истинная сингулярность кривизны имеет место при $\rho = 0$, $r = 0$, $\theta = \pi/2$, и исследование показывает, что эта сингулярность имеет форму кольца (факт, не очевидный в координатах Бойера — Линдквиста и устанавливаемый при другой параметризации метрики — см., напр., [26, 227]). Поверхности же, на которых $\Delta = 0$, т.е.

$$r_{\pm} = m \pm \sqrt{m^2 - a^2} \quad (3.138)$$

являются горизонтами. Как видно, горизонты существуют при $m^2 \geq a^2$, причем в случае неравенства имеется два простых горизонта, а в случае равенства — один экстремальный (двойной).

Остановимся на метрике с горизонтами. В этом случае норма вектора Киллинга ξ^μ , равная

$$\xi_\mu \xi^\mu = g_{tt} = \frac{\Delta - a^2 \sin^2 \theta}{\rho^2}, \quad (3.139)$$

обращается в нуль на поверхности, где

$$\Delta = a^2 \sin^2 \theta, \quad (3.140)$$

т.е. на поверхности вне горизонта (с поверхностью $\Delta = 0$ — горизонтом $r = r_+$ — она имеет лишь две общие точки, а именно полюса $\theta = 0$ и $\theta = \pi$ — см. рис. 3.9). Поверхность (3.140) носит название предела стационарности, так как на самой этой поверхности и внутри нее не могут существовать стационарные наблюдатели. Кроме того, поверхность (3.140) является поверхностью бесконечного красного смещения для излучаемых с нее на пространственную бесконечность фотонов.

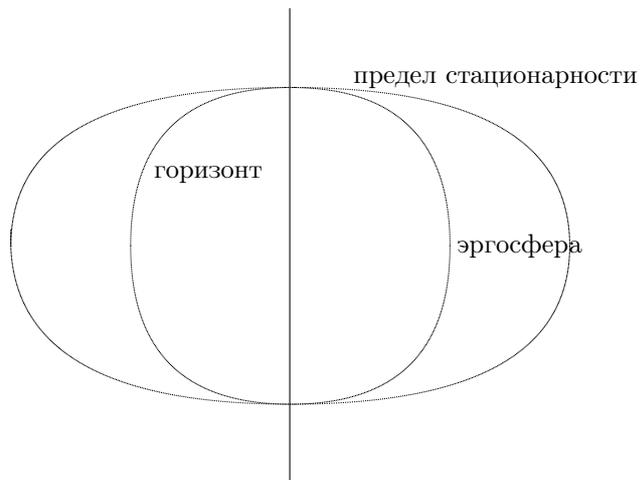


Рис. 3.8. Горизонт и эргосфера ЧД Керра — меридиональный разрез

Область пространства между горизонтом $r = r_+$ и поверхностью (3.140) получила название эргосферы. Название связано с принципиальной возможностью извлекать из керровской ЧД ее вращательную энергию с помощью процесса Пенроуза [245]. Дело в том, что в эргосфере имеются орбиты массивных частиц с отрицательной полной энергией (это означает, что энергия связи по абсолютной величине превышает массу). Тогда можно запустить в эргосферу, скажем, космический корабль, там он излучит частицы с отрицательной энергией и вернется в обычную область пространства с большей энергией, чем он имел первоначально. Частицы же с отрицательной энергией, движение которых направлено против вращения ЧД, при их поглощении уменьшают скорость вращения и полную вращательную энергию ЧД.

3.10. Термодинамика черных дыр

3.10.1. Четыре закона термодинамики ЧД

По определению, ЧД ничего не могут излучать (в рамках классической теории), так как их не могут покидать ни фотоны, ни, тем более, массивные частицы. Тем не менее, как впервые установил Бекенштейн [62], ЧД удовлетворяют некоторым соотношениям, идентичным известным формулам термодинамики, и им можно приписать такие термодинамические характеристики, как температура (вообще говоря, ненулевая) и энтропия.

Хокинг [162], рассматривая квантовые процессы вблизи горизонтов, обнаружил, что ЧД испускают излучение как черные тела с определенной температурой и вывел точное соотношение между энтропией и площадью горизонта ЧД. Несколько позднее были сформулированы четыре закона термодинамики ЧД [150]. В результате появилось последовательное описание термодинамических систем, включающих как обычную материю, так и ЧД.

Так называемый *нулевой закон термодинамики ЧД* формулируется следующим образом.

Нулевой закон. *Поверхностная гравитация ЧД не меняется от точки к точке на горизонте событий.*

Поверхностной гравитацией ЧД называется значение ускорения свободного падения, вычисленное на горизонте событий через 4-ускорение частицы. Для ЧД Шварцшильда поверхностная гравитация κ равна

$$\kappa = 1/(4m), \quad (3.141)$$

а для ЧД Керра довольно громоздкое вычисление дает (см., напр., [154])

$$\kappa = \frac{r_+ - m}{2mr_+} = \frac{\sqrt{m^2 - a^2}}{2m(m + \sqrt{m^2 - a^2})}. \quad (3.142)$$

В формулы (3.141) и (3.142) входят только параметры ЧД, а угловой зависимости нет, что и подтверждает Нулевой закон.

Первый закон термодинамики ЧД аналогичен первому закону обычной термодинамики, а именно:

$$dm = \frac{\varkappa}{8\pi} dA + \Omega dJ, \quad (3.143)$$

где A — площадь горизонта,

$$A = 4\pi(r_+^2 + a^2) \quad (3.144)$$

(при $a = 0$ получаем выражение для ЧД Шварцшильда, $A = 4\pi r_+^2 = 16\pi m^2$), а Ω — угловая скорость на горизонте

$$\Omega = \frac{a}{r_+^2 + a^2}. \quad (3.145)$$

Равенство (3.143) аналогично первому закону термодинамики $dU = TdS + dW$ (U — внутренняя энергия, S — энтропия, W — работа). Следовательно, (3.143) можно прочесть следующим образом: ΩdJ есть работа, произведенная над ЧД при добавлении углового момента dJ . Дальнейшее исследование приводит к следующему отождествлению для температуры и энтропии ЧД [162]:

$$T_{\text{ВН}} = \varkappa/(2\pi), \quad S_{\text{ВН}} = A/4. \quad (3.146)$$

В частности, из выражения (3.142) следует, что температура равна нулю лишь у экстремальной ЧД с двойным горизонтом ($|a| = m$).

Второй закон термодинамики ЧД аналогичен утверждению о неубывании энтропии.

Второй закон. *Никакой классический процесс не приводит к уменьшению площади горизонта ЧД.*

Квантовый процесс испарения ЧД за счет излучения Хокинга нарушает этот классический закон.

Наконец, третий закон утверждает недостижимость абсолютного нуля температуры, или, как иногда формулируют, несуществование отрицательных температур. Для ЧД аналогичное утверждение состоит в следующем [56].

Третий закон. *Не существует процедуры, приводящей температуру ЧД к нулю конечной последовательностью операций.*

Иначе говоря, из неэкстремальной ЧД нельзя получить экстремальную и, тем более, превысить экстремальный предел (получить $|a| > m$) и получить из ЧД голую сингулярность.

3.10.2. Испарение черных дыр

Как уже упоминалось, согласно [162], ЧД испаряются, излучая как черные тела с температурой $T_{\text{ВН}} = \varkappa/(2\pi)$, где поверхностная гравитация на горизонте \varkappa определяется формулами (3.141) и (3.142). Качественная физическая картина испарения ЧД заключается в том, что виртуальные пары частица-античастица, постоянно возникающие и исчезающие согласно квантовой теории поля, находятся в области горизонта в достаточно сильном гравитационном поле. Последнее с некоторой вероятностью (которая тем больше, чем больше кривизна пространства-времени) может растаскивать эти виртуальные пары, таким образом, что одна из частиц падает на горизонт, а другая уходит на бесконечность. В результате ухода частиц с ненулевой энергией ЧД теряет массу.

Подстановка соответствующих констант приводит к следующему приближенному выражению для температуры ЧД Шварцшильда в кельвинах:

$$T_{\text{ВН}} \approx 10^{-7} m_{\odot}/m \text{ К}, \quad (3.147)$$

где m_{\odot} — масса Солнца. Таким образом, у ЧД с массами порядка звездных и тем более галактических температура весьма мала, и их испарение происходит крайне медленно. Чем ЧД меньше, тем больше искривление пространства у ее горизонта и выше температура. В частности, при добавлении массы ЧД ее температура понижается, что говорит о ее отрицательной теплоемкости.

Потеря массы (энергии) ЧД может быть рассчитана по закону Стефана-Больцмана, $-\dot{E}/A = \sigma T_{\text{ВН}}^4$, где σ — констан-

та Стефана-Больцмана. Интегрирование этого выражения дает следующую величину массы ЧД как функции времени:

$$m(t) = (m_0^3 - 3Kt)^{1/3}, \quad K = \frac{1}{15360\pi} = \frac{\hbar c^4}{15360\pi G^2}, \quad (3.148)$$

где m_0 — масса ЧД в момент $t = 0$; после последнего знака равенства дано выражение в обычных единицах. В момент $t_f = m_0^3/(3K)$ испарение заканчивается, но пока неясно, исчезает ли ЧД полностью или остается какой-то объект с массой порядка планковской, $m \sim 10^{-5}$ г. Дело в том, что при приближении массы к планковскому значению (а следовательно, радиуса горизонта — к планковской длине $\sim 10^{-33}$ см) должны вступать в действие пока неизвестные квантово-гравитационные закономерности.

За секунду до окончания процесса, как нетрудно подсчитать, масса ЧД — порядка 10^9 г $\sim 10^{33}$ ГэВ. Именно такая энергия высвобождается в последнюю секунду испарения ЧД.

С другой стороны, если приравнять время жизни $t_f = m_0^3/3K$ ЧД с заданной массой m_0 и возраст Вселенной $t_U \sim 13 \cdot 10^9$ лет, получим массу $m_0 = m_U \sim 10^{15}$ г. Таким образом, все первичные ЧД массой меньше 10^{15} г, образовавшиеся в первые минуты или тысячелетия существования Вселенной, уже испарились и внесли свою лепту в химический состав вещества Вселенной, а оставшиеся ЧД с массой порядка m_U должны в современную эпоху время от времени взрываться. Более массивные ЧД проживут еще долго.

3.11. Регулярные черные дыры и черные вселенные

3.11.1. Различные виды регулярных черных дыр

В физике ЧД, как и в ОТО в целом, одной из известных и трудно решаемых проблем является проблема сингулярностей, которые

существуют внутри горизонта событий в ЧД-решениях, полученных при наиболее простых и естественных предположениях (решения Шварцшильда, Райснера — Нордстрема, Керра и Керра — Ньюмена в ОТО и их аналоги во многих альтернативных теориях гравитации). ОТО и другие классические теории гравитации, как известно, теряют смысл в сингулярностях. По этой причине для полного понимания теоретических предсказаний необходимо либо избавляться от сингулярностей, либо модифицировать сами теории, например, рассматривая квантовые эффекты. Известно много работ в этом направлении, в некоторых из них сингулярности внутри горизонтов ЧД предлагалось заменить регулярными “сердцевинами” с конечной плотностью.

Обсудим возможные геометрии регулярных ЧД, оставаясь в рамках статических асимптотически плоских сферически-симметричных конфигураций. Метрику предполагаем в виде (3.30), используя квазиглобальную радиальную координату:

$$ds^2 = A(\rho)dt^2 - \frac{d\rho^2}{A(\rho)} - r^2(\rho)d\Omega^2. \quad (3.149)$$

На плоской асимптотике ($\rho \rightarrow \infty$) имеем, без потери общности, $A(\rho) \rightarrow 1$ и $r(\rho) \approx \rho$. Центр $\rho = \rho_c$ (если он существует) соответствует $r = 0$; горизонты (если они есть) описываются регулярными нулями $A(\rho)$, число, порядок и расположение которых определяют глобальную причинную структуру пространства-времени.

Тогда известные в литературе виды регулярных геометрий классифицируются следующим образом (см. рис. 3.9).

- 1) ЧД с регулярным центром ($r \approx \text{const} \cdot (\rho - \rho_c)$, $A \rightarrow A_c > 0$, $A(dr/d\rho)^2 \approx 1 + O(r^2)$ при $\rho \rightarrow \rho_c$). Поскольку регулярный центр может располагаться только в R-области, у таких ЧД должно быть по меньшей мере два простых горизонта или один двойной, и их причинная структура описывается теми же диаграммами Картера — Пенроуза, что и, соответственно, у неэкстремальных и экстремальных ЧД Райсне-

ра — Нордстрема (диаграммы 1b и 1c на рис. 3.9). Большое число горизонтов также не исключено и приводит к более сложным структурам.

- 2) ЧД без центра, с горизонтами второго порядка, имеющими бесконечную площадь (так называемые холодные ЧД, т.к. подобные горизонты всегда характеризуются нулевой температурой Хокинга) [81, 82, 108]. Они могут иметь различную причинную структуру. В одном случае диаграмма Картера — Пенроуза та же, что для экстремальной ЧД Керра и состоит из бесконечной последовательности R -областей, но, конечно, не содержит керровской кольцевой сингулярности (график 2a и диаграмма 2b на рис. 3.9). В другом случае есть только четыре R -области (графики 2c и диаграмма 2d).
- 3) ЧД с причинной структурой типа неэкстремальной керровской ЧД, снова без сингулярного кольца [83, 113] (диаграмма 3b).
- 4) Регулярные ЧД с причинной структурой типа Шварцшильда [86] (диаграмма 4b), но с космологическим расширением вместо сингулярности. Такие ЧД получили название *черных вселенных* [84]

На рис. 3.9 показано качественное поведение метрических функций и диаграммы Картера–Пенроуза для указанных четырех типов регулярных статических сферически-симметричных ЧД. Диаграммы 1b и 1d (как у неэкстремальных и экстремальных ЧД Райснера — Нордстрема) соответствуют кривым A_1 и A_2 на графике 1a. Диаграмма 2b, такая же, как у экстремальной метрики Керра, соответствует графику 2a, диаграмма 2d — графику 2c, 3b — 3a и 4b — 4a. Буквы R и T обозначают R - и T -области пространства-времени. Диаграммы 1b, 1c, 2b и 3b бесконечно продолжаемы вверх и вниз. На всех диаграммах все

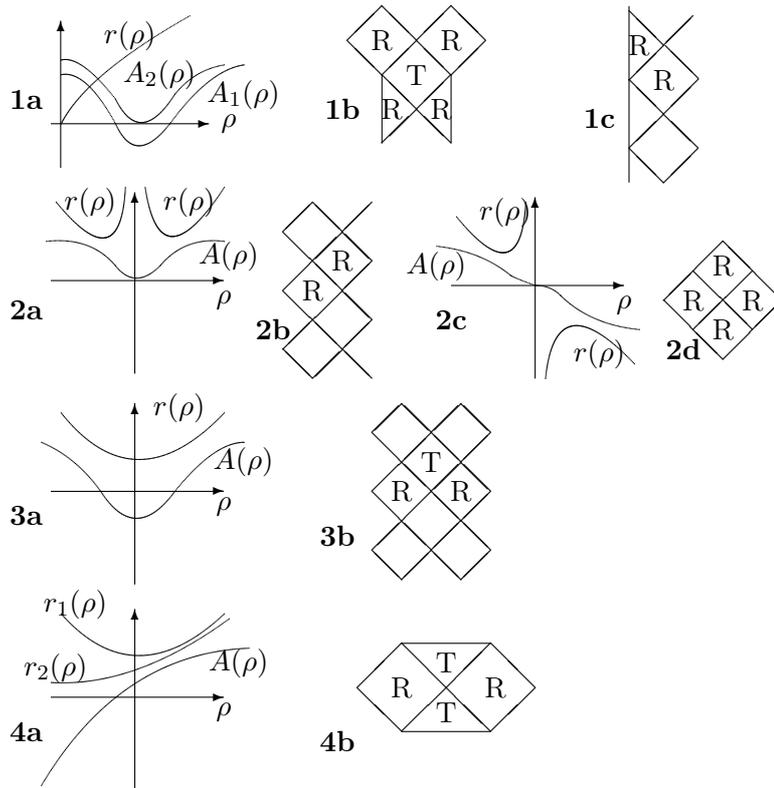


Рис. 3.9.

внутренние наклонные линии изображают горизонты, все границы диаграмм соответствуют $r = \infty$, за двумя исключениями: вертикальные линии на диаграммах 1b и 1d соответствуют регулярному центру $r = 0$; горизонтальные линии на диаграмме 4b соответствуют либо $r = \infty$, либо $r = r_0 > 0$, согласно виду кривых $r_1(\rho)$ или $r_2(\rho)$ при больших отрицательных ρ .

Тип 1 регулярных ЧД был, по-видимому, впервые упомянут в работе Бардина [55], в которой высказана идея регулярных ЧД

и в качестве примера предложена метрика вида

$$r \equiv \rho, \quad A(\rho) = 1 - \frac{M\rho^2}{(\rho^2 + q^2)^{3/2}}, \quad (3.150)$$

где $M, q = \text{const}$ и имеется два горизонта при условии $q^2 < (16/27)M^2$. Впоследствии появились многочисленные примеры ([50, 75, 137] и др.) решений в виде регулярных ЧД, где, как и в метрике (3.150), $r \equiv \rho$. Согласно уравнениям Эйнштейна, ТЭИ материи при этом должен удовлетворять условию

$$T_0^0 \equiv T_1^1, \quad (3.151)$$

или, другими словами, $\varepsilon = -p_r$ ($\varepsilon = T_0^0$ — плотность энергии, $p_r = -T_1^1$ — радиальное давление). Последнее условие инвариантно относительно радиальных бустов, так что такую материю имеет смысл назвать вакуумной [137]. В самом деле, именно для вакуума нельзя выделить сопутствующую СО: как бы мы ни двигались в радиальном направлении — вакуум с ТЭИ, удовлетворяющим условию (3.151), выглядит для нас одинаково. При этом в регулярном центре уравнение состояния получается с необходимостью такое же, как у космологической постоянной [85], $T_\mu^\nu \propto \delta_\mu^\nu$.

Было также показано [75], что регулярные ЧД с $r \equiv \rho$ и произвольной функцией $A(\rho)$, удовлетворяющей условиям регулярного центра, могут быть получены как решения в виде магнитных монополей в ОТО с источником в виде нелинейного электромагнитного поля с лагранжианом $L(F)$, $F := F_{\mu\nu}F^{\mu\nu}$ ($F_{\mu\nu}$ — тензор электромагнитного поля); произвол в задании $A(\rho)$ соответствует свободе выбора функции $L(F)$. Выяснилось, что регулярные модели с электрическим зарядом невозможны [75], каков бы ни был выбор $L(F)$, если функция $L(F)$ одна и та же во всем пространстве. Эту теорему, впрочем, можно обойти, предполагая различный вид $L(F)$ около центра и при больших r [105], т.е. предполагая нечто похожее на фазовые переходы при некоторых значениях радиальной координаты.

Были найдены и обсуждались и другие примеры регулярных ЧД этого типа, см., напр., [65, 165, 274] и приведенные там ссылки.

Тип 2 регулярных ЧД, электрически заряженных и нейтральных, получен [81, 82, 108] в рамках СТТ Бранса-Дикке с константой связи $\omega < -3/2$, когда скалярное поле носит фантомный характер. Для их существования требуется тонкая настройка в виде специальных соотношений между ω и константами интегрирования соответствующих точных решений. Интересно, что у этих регулярных холодных ЧД есть сингулярные аналоги в картине Эйнштейна, или, другими словами, в ОТО с минимально-связанным фантомным скалярным полем (см. решение анти-Фишера (3.111)) [80].

Тип 3 регулярных ЧД был получен [83, 113] среди статических сферически-симметричных решений эффективных уравнений гравитации в мире на бране типа RS2 [276]; было показано, что такие регулярные решения суть решения общего положения в некоторой области значений констант интегрирования [83] и что многие из них устойчивы, по крайней мере относительно некоторых видов возмущений [37]. Следует, однако, подчеркнуть, что эти четырехмерные уравнения не образуют замкнутую систему, а для получения полной пятимерной геометрии необходимо решать соответствующие пятимерные уравнения, и результаты на этом пути можно считать только предварительными [114].

Конфигурации типа 4 (черные вселенные) получены [86] как решения общего положения уравнений Эйнштейна и минимально-связанного фантомного скалярного поля с потенциалами определенного вида. Напомним, что именно такие поля позволяют получать в космологии уравнения состояния темной энергии с $p/\varepsilon = w < -1$, что, по некоторым данным, необходимо для описания современной фазы ускоренного расширения Вселенной (см., напр., обзоры [123, 271] и приведенные там ссылки). Такие конфигурации, по нашему мнению, представляют особый интерес, поэтому кратко опишем способ их получения и некоторые

свойства.

3.11.2. Черные вселенные с минимально-связанным скалярным полем

Рассмотрим действие для самогравитирующего фантомного скалярного поля в ОТО

$$S = \int \sqrt{g} d^4x [R - (\partial\phi)^2 - 2V(\phi)], \quad (3.152)$$

где $g = |\det(g_{\mu\nu})|$, $(d\phi)^2 = g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi$ и $V(\phi)$ — произвольный потенциал. С метрикой (3.1) и $\phi = \phi(\rho)$ уравнение скалярного поля и три независимых комбинации уравнений Эйнштейна имеют вид

$$(Ar^2\phi')' = -r^2 dV/d\phi, \quad (3.153)$$

$$(A'r^2)' = -2r^2 V; \quad (3.154)$$

$$2r''/r = \phi'^2; \quad (3.155)$$

$$A(r^2)'' - r^2 A'' = 2, \quad (3.156)$$

где штрих означает $d/d\rho$. Эти уравнения можно получить из (3.123)–(3.126) как частный случай. Уравнение скалярного поля (3.153) есть следствие трех уравнений (3.154)–(3.156), которые при заданном потенциале $V(\phi)$ образуют определенную систему уравнений для неизвестных $r(\rho)$, $A(\rho)$, $\phi(\rho)$. Уравнение (3.156) легко интегрируется:

$$B' \equiv (A/r^2)' = 2(\rho_0 - \rho)/r^4, \quad (3.157)$$

где $B(\rho) = A/r^2$ и ρ_0 — константа интегрирования.

Уравнение (3.157), как мы знаем, приводит к теореме о глобальной структуре [76] (см. раздел 5.8.3), согласно которой в присутствии скалярного поля набор возможных видов глобальной причинной структуры тот же, что и в решении Шварцшильда — де Ситтера с произвольными параметрами. Этот результат не

зависит от знака кинетической энергии скалярного поля, так как этот знак не влияет на вид уравнений (3.156) и (3.157).

Как показано в статье [86], система (3.153)–(3.156) имеет 16 различных типов регулярных решений с плоскими, деситтеровскими и АдС асимптотиками. Обсудим асимптотически плоские конфигурации, для которых $A(\rho) \rightarrow 1$ и $r(\rho) \approx \rho$ при $\rho \rightarrow \infty$. Тогда при уменьшении ρ от бесконечности производная r' также уменьшается в силу уравнения (3.155), а скорость уменьшения зависит от конкретных параметров системы. Если уменьшение достаточно медленное, $r(\rho)$ достигает нуля при конечном ρ , это означает, что у системы есть центр. Центр может быть регулярным только если отсутствуют горизонты (иначе это была бы регулярная ЧД типа 1, а это невозможно в силу теоремы о глобальной структуре), и тогда вместо ЧД получается частицеподобное решение.

Если предполагать отсутствие сингулярностей, другие возможности поведения $r(\rho)$ — это $r(\rho) \rightarrow r_0 = \text{const}$ и $r(\rho) \rightarrow \infty$ при $\rho \rightarrow -\infty$. (Заметим, что никакое колебательное поведение $r(\rho)$ невозможно в силу (3.155), согласно которому $r'' \geq 0$.)

В первом случае, согласно (3.157),

$$A' \approx -2\rho/r_0^2 \rightarrow +\infty, \quad A \approx -\rho^2/r_0^2 \quad \text{при } \rho \rightarrow -\infty.$$

Таким образом, эта “ r_0 -асимптотика” расположена в T -области. Радиус r_0 связан с предельным значением потенциала: $V \rightarrow -1/r_0^2$. Меняя обозначение $-\rho \rightarrow T$ и $t \rightarrow x$ (так как ρ здесь стало временной координатой, а бывшее время t — пространственной), можно написать асимптотический вид получившейся метрики Кантовского — Сакса в виде

$$ds^2 \approx \frac{r_0^2}{T^2} dT^2 - \frac{T^2}{r_0^2} dx^2 - r_0^2 d\Omega^2 \quad \text{при } T \rightarrow \infty. \quad (3.158)$$

Это существенно анизотропная модель, в которой нет расширения в двух угловых направлениях и имеет место экспоненциальное (в терминах физического времени $\tau \sim \log T$) расширение в

третьем направлении x . С точки зрения наблюдателей, расположенных на больших значениях $\rho > 0$, эта вселенная находится за горизонтом событий ЧД.

В случае $r \rightarrow \infty$ наиболее интересная возможность — это $r \sim |\rho|$ при $\rho \rightarrow -\infty$. Полагая $r \approx -a\rho$, $a = \text{const} > 0$, из (3.157) и (3.154) получаем

$$A \approx 1/a^2 - Ca^2\rho^2, \quad V \approx 3Ca^2, \quad C = \text{const}. \quad (3.159)$$

Легко проверить, что $C = 0$ приводит к метрике Минковского при больших отрицательных ρ (где темп хода часов другой, нежели при больших $\rho > 0$, если $a \neq 1$), метрика антидеситтеровская при $C < 0$ и деситтеровская при $C > 0$. В случаях $C \leq 0$ горизонты отсутствуют, так как иначе функция $B(\rho)$ имела бы минимум при конечном ρ , что невозможно в силу (3.157). Следовательно, возможные решения с $C \leq 0$ описывают проходимость кротовые норы, и детали их геометрии зависят от конкретного выбора $V(\phi)$.

Наконец, при $C > 0$ получаем решение с деситтеровской асимптотикой, описывающей (так как мы находимся в Т-области) изотропное расширение или сжатие. С точки зрения внешнего наблюдателя, расположенного при больших $\rho > 0$, это ЧД, однако возможный исследователь, падающий в такую ЧД, теперь имеет шанс на выживание и новую жизнь в расширяющейся и постепенно изотропизирующейся Вселенной Кантовского — Сакса. Конкретный профиль метрики ЧД и режим изотропизации после пересечения горизонта зависят от выбора $V(\phi)$.

Поскольку, согласно теореме о глобальной структуре, асимптотически плоская конфигурация может иметь только один горизонт, полученная ЧД неизбежно имеет такую же глобальную структуру, как пространство-время Шварцшильда, однако вместо сингулярности $r = 0$ на диаграмме Картера-Пенроуза имеет место деситтеровская бесконечность $r = \infty$.

Простой пример можно получить, полагая [86]

$$r = (\rho^2 + b^2)^{1/2}, \quad b = \text{const} > 0 \quad (3.160)$$

и используя метод обратной задачи. Уравнение (3.157) дает

$$B(\rho) = \frac{A(\rho)}{r^2(\rho)} = \frac{c}{b^2} + \frac{1}{b^2 + \rho^2} + \frac{\rho_0}{b^3} \left(\frac{b\rho}{b^2 + \rho^2} + \arctg \frac{\rho}{b} \right), \quad (3.161)$$

где $c = \text{const}$. Тогда уравнения (3.155) и (3.154) приводят к выражениям для $\phi(\rho)$ и $V(\rho)$:

$$\phi = \pm \sqrt{2} \arctg(\rho/b) + \phi_0, \quad (3.162)$$

$$V = -\frac{c}{b^2} \frac{r^2 + 2\rho^2}{r^2} - \frac{\rho_0}{b^3} \left(\frac{3b\rho}{r^2} + \frac{r^2 + 2\rho^2}{r^2} \arctg \frac{\rho}{b} \right), \quad (3.163)$$

где функция $r = r(\rho)$ задана выражением (3.160). В частности,

$$B(\pm\infty) = -\frac{1}{3}V(\pm\infty) = \frac{2bc \pm \pi\rho_0}{2b^3}. \quad (3.164)$$

Выбирая в (3.162) без потери общности знак плюс и $\phi_0 = 0$, получаем для $V(\phi)$ ($\psi := \phi/\sqrt{2}$):

$$V(\phi) = -\frac{c}{b^2}(3 - 2\cos^2\psi) - \frac{\rho_0}{b^3} [3\sin\psi\cos\psi + \psi(3 - 2\cos^2\psi)]. \quad (3.165)$$

Поведение решения зависит от двух констант интегрирования: c сдвигает график функции $B(\rho)$ вверх или вниз, а ρ_0 указывает максимум $B(\rho)$. Как $r(\rho)$, так и $B(\rho)$ — четные функции, если $\rho_0 = 0$, в противном случае симметрия $B(\rho)$ теряется. Из асимптотической плоскостности при $\rho = +\infty$ следует $2bc = -\pi\rho_0$, а шварцшильдовская масса, определенная обычным образом, есть $m = \rho_0/3$.

При том же условии асимптотической плоскостности при $\rho_0 = m = 0$ получаем простейшую конфигурацию кротовой норы, а именно кротовую нору Эллиса (4.1) [141]: $A \equiv 1$, $V \equiv 0$. При $\rho_0 < 0$, согласно (3.164), получаем кротовую нору с $m < 0$ и АдС-метрикой на противоположном конце, соответствующей космологической постоянной $V_- < 0$. При $\rho_0 > 0$, когда $V_- > 0$,

получаем регулярную ЧД с $m > 0$ и деситтеровской асимптотической далеко за горизонтом, в точности соответствующую нашему описанию черной вселенной.

Радиус горизонта $r(\rho_h)$ можно получить, решая трансцендентное уравнение $A(\rho_h) = 0$, где $A(\rho)$ задано выражением (3.161). Он зависит от параметров m и $b = \min r(\rho)$ и не может быть меньше величины b , которая также играет роль скалярного заряда: $\psi \approx \pi/2 - b/\rho$ при больших ρ . Поскольку $A(0) = 1 + c$, горловина $\rho = 0$ расположена в R-области, если $c > -1$, т.е., если $3\pi m < 2b$, на горизонте, если $3\pi m = 2b$ и в T-области за горизонтом, если $3\pi m > 2b$. Полученные соотношения между m и b показывают (и это, вероятно, справедливо в общем случае), что если масса ЧД доминирует над скалярным зарядом, то горловина невидима для удаленного наблюдателя, и ЧД выглядит из статической области приблизительно так же, как и в обычных случаях в ОТО.

Как следует из уравнений (3.153) и (3.154), потенциал V стремится к константе и, более того, $dV/d\phi \rightarrow 0$ на обоих концах области определения ρ . Это общее свойство всех классов регулярных решений, указанных в [86]. Точнее, *регулярная конфигурация скалярного поля требует потенциала, имеющего как минимум две точки с нулевым наклоном (не обязательно обе они являются экстремумами) при разных значениях ϕ .*

Подходящие потенциалы — это, например, $V = V_0 \cos^2(\phi/\phi_0)$ и потенциал сомбреро $V = (\lambda/4)(\phi^2 - \eta^2)^2$, где V_0 , ϕ_0 , λ , η — константы. Для плоской бесконечности $\rho = +\infty$, конечно, требуется $V_+ = 0$, а деситтеровская асимптотика может соответствовать максимуму V , так как фантомные поля, вообще говоря, стремятся взобраться на склон потенциала, а не скатиться с него, что очевидно из уравнения (3.153). В соответствии с этим в работе Фараони [145] где рассматривались пространственно плоские фантомные космологические модели, показано, что если потенциал $V(\phi)$ ограничен сверху величиной $V_0 = \text{const} > 0$, то деситтеровское решение с этим V_0 есть глобальный аттрактор. Весьма

вероятно, что этот вывод справедлив и для моделей Кантовского — Сакса после изотропизации.

Темп деситтеровского расширения на больших временах полностью определяется соответствующим значением потенциала $V_- > 0$ (которое, в наших обозначениях согласно (3.152), совпадает с эффективной космологической постоянной Λ на больших временах), а не такими деталями конкретного решения как, например, шварцшильдовская масса, соответствующая плоской асимптотике.

Можно заключить, что решения в виде черных вселенных являются типичными решениями уравнений Эйнштейна с источниками в виде фантомных скалярных полей с подходящими потенциалами.

Существование решений в виде черных вселенных приводит к идее, что наша Вселенная могла появиться в результате фантомно-доминированного коллапса с образованием ЧД в некоей “материнской” вселенной, испытав переход через горизонт и последующую изотропизацию (например, за счет рождения частиц). Известно, что наблюдениями не исключено, что наша Вселенная может описываться моделью Кантовского — Сакса [43], если ее изотропизация произошла достаточно рано (при красных смещениях $z \gtrsim 1000$). В принципе описанный сценарий дает новый механизм размножения вселенных в дополнение к известному механизму, связанному с хаотической инфляцией.

Глава 4

Кротовые норы

4.1. Понятие кротовой норы

В этой главе мы обсудим еще один вид конфигураций, связанных с сильным искривлением пространства-времени — кротовые норы (КН). Прежде всего уточним терминологию.

Используемый в русскоязычной литературе термин “кротовая нора” — несколько вольный перевод с английского введенного Дж.А. Уилером [305] термина “wormhole”. Буквальный перевод — “червоточина” — звучит по-русски менее привлекательно, но, к сожалению, нередко используется переводчиками фантастических произведений, герои которых с помощью кротовых нор легко и без особых эмоций преодолевают межзвездные расстояния или тысячелетние отрезки времени.

Под КН обычно понимают сравнительно узкие “мосты”, или “перемычки” между сравнительно большими или бесконечными по протяженности областями пространства-времени. Возможный вид геометрии КН иллюстрируется рисунком 4.1, где представлены двумерные образы вместо трехмерных. У каждой из изображенных на этом рисунке поверхностей есть участок в виде сравнительно тонкой трубки, соединяющей либо плоские “вселенные” (а), либо сферические (b) (конфигурация типа гантели),

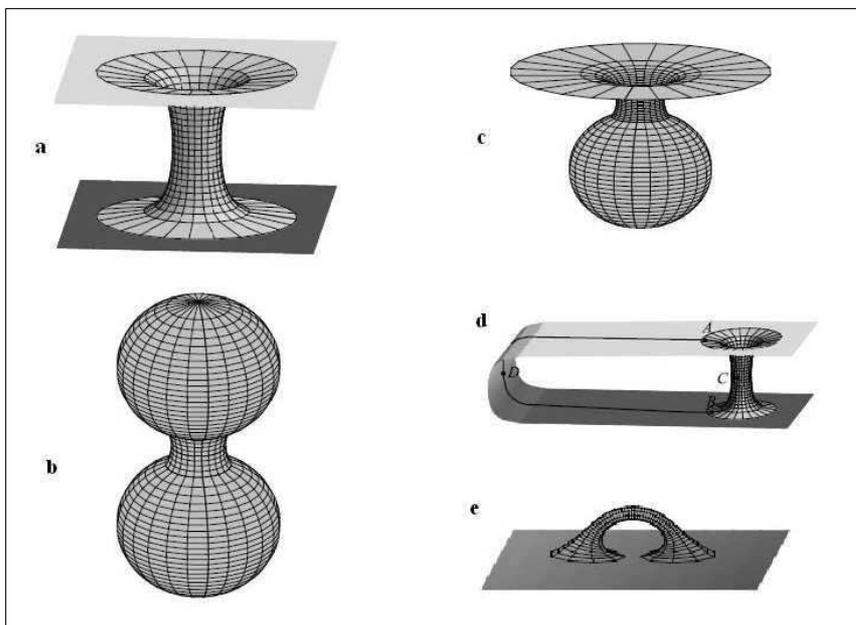


Рис. 4.1. Различные конфигурации кротовых нор — двумерные аналоги

либо плоскую со сферической (с) (конфигурация типа висящей капли) — либо разные области одной и той же “вселенной”, как на поверхностях d и e.

Мы видим, что понятие КН носит, по сути дела, интуитивный характер и опирается на слова типа “сравнительно узкий” и т.п., однако в нем есть элемент, допускающий строгое определение. Это — горловина, определяемая как двумерная поверхность минимальной площади, не допускающая дальнейшего непрерывного стягивания. На рисунке 4.1 горловины — это одномерные сечения минимального радиуса, которые условно изображают двумерные сферы минимальной площади. Впрочем, очевидно, что в общем случае КН не обязана быть сферически-симметричной, а ее горловина не обязана быть сферой.

В случае статических сферически-симметричных пространств, описываемых метрикой общего вида (3.1), несложно сформулировать общие требования на метрические коэффициенты, при выполнении которых можно считать, что данная метрика описывает КН.

Условие существования двух обширных областей, разделенных горловиной, означает, что функция $r \equiv e^\beta(u)$, равная при каждом значении u радиусу координатной сферы $u = \text{const}$, $t = \text{const}$, должна иметь регулярный минимум $r = r_{\min} > 0$ при некотором значении u и принимать значения $r \gg r_{\min}$ вдали от этого минимума.

Второе условие — отсутствие горизонтов: требование беспрепятственной двусторонней проходимости КН принципиально отличает их от черных дыр. Для метрики (3.1) это требование имеет вид $g_{tt} \equiv e^{2\gamma} > 0$ во всем диапазоне значений радиальной координаты.

Простейший пример метрики КН

$$ds^2 = dt^2 - du^2 - (u^2 + a^2)d\Omega^2, \quad a > 0 \quad (4.1)$$

отличается от метрики плоского пространства-времени лишь появлением слагаемого a^2 . Но геометрия при этом меняется радикально: при $a = 0$ имеем пространство Минковского, координата u меняется от нуля до бесконечности, $u = 0$ представляет собой регулярный центр, и продолжение на отрицательные значения u не имеет смысла. Если же $a > 0$, значение $u = 0$ есть регулярный минимум функции $r(u) = \sqrt{u^2 + a^2}$, т.е. горловина, ничто не мешает рассматривать отрицательные u , и $u \rightarrow -\infty$ — точно такая же пространственная бесконечность, как $u \rightarrow +\infty$. Трехмерная геометрия соответствует рисунку 4.1а и не зависит от времени.

Для иллюстрации важности второго условия сделаем небольшой экскурс в историю.

Исследователи задумались о том, чего можно ожидать при сильном искривлении пространства-времени, сразу после появления ОТО. Уже в год рождения ОТО (1916) Карл Шварцшильд

нашел свое знаменитое решение (см. (3.20)), описывающее, по принятой сейчас терминологии, простейшую черную дыру. Менее известна датированная тем же годом работа Людвиг Фламма [147], в которой замечено, что решение Шварцшильда описывает что-то вроде моста, или связки, между двумя мирами или двумя частями одного мира, т.е. именно то, что современные авторы называют кротовой норой.

Конкретизируем это наблюдение, переходя от координат кривизн в метрике (3.20) к изотропным координатам (3.2) с помощью преобразования

$$r = u \left(1 + \frac{m}{2u} \right)^2. \quad (4.2)$$

В результате метрика Шварцшильда примет вид

$$ds^2 = \frac{(2 - m/u)^2}{(2 + m/u)^2} dt^2 - \left(1 + \frac{m}{2u} \right)^4 (du^2 + u^2 d\Omega^2), \quad (4.3)$$

где координата $u \approx r$ при $r \rightarrow \infty$ и принимает значение $u = m/2$ на горизонте $r = 2m$. Удобно ввести безразмерную переменную $y = 2u/m$, тогда

$$ds^2 = \frac{(y - 1)^2}{(y + 1)^2} dt^2 - \frac{m^2}{4} \frac{(y + 1)^4}{y^4} (dy^2 + y^2 d\Omega^2), \quad (4.4)$$

и горизонту соответствует значение $y = 1$. R -область $r > 2m$ теперь описывается как область $y > 1$. Легко убедиться, однако, что метрика (4.4) не изменяется при инверсии $y \mapsto 1/y$. Иначе говоря, область $0 < y < 1$ представляет собой еще одну R -область, точно такую же, как исходная, значение $y = 0$ соответствует еще одной пространственной бесконечности, такой же, как исходная ($y = \infty$), а значение $y = 1$, где координатные сферы имеют минимальную площадь, равную $16\pi m^2$, описывает горловину в пространственном сечении исследуемого четырехмерного многообразия. Если бы геометрия ограничивалась пространственным сечением, были бы все основания говорить

о КН. Однако всё меняется из-за равенства $g_{tt} = 0$ при $y = 1$: это, как мы знаем из предыдущей главы, горизонт, благодаря которому четырехмерная геометрия непродолжаема в статической системе отсчета. Для более полного описания можно перейти, например, к координатам Крускала, в которых выявляются свойства полной геометрии пространства-времени. Из полученной в результате полной картины следует, что метрика (4.4) описывает лишь две R-области, представленные на диаграмме Крускала правым и левым квадрантами. (Для сравнения: метрика (3.20) при $r > 2m$ описывает лишь одну из R-областей.) Сфера $y = 1$, кажущаяся при взгляде на формулу (4.4) границей между R-областями, оказывается при полном описании парой различных поверхностей, между которыми имеется целая T-область. Парадокс объясняется, в частности, тем, что горизонт $r = 2m$ (он же — $y = 1$) *не принадлежит* ни одной из R-областей, а следовательно единство их описания метрикой (4.4) лишь кажущееся, и координатные области $y > 1$ и $y < 1$, соответствующие разным R-областям пространства-времени, следует рассматривать отдельно.

В связи с метриками типа (4.4), содержащими горизонты, иногда говорят о “непроходимых КН”. По нашему мнению, в таких случаях во избежание путаницы следует говорить о совсем другом классе релятивистских объектов — черных дырах, тем более что, как мы видели при построении диаграмм Картера — Пенроуза, все черные дыры с простыми горизонтами (и вообще с горизонтами нечетного порядка) содержат такие же пары R-областей, как и черная дыра Шварцшильда, и о них с тем же успехом (точнее, с той же степенью бесполезности) можно было бы говорить как о “непроходимых КН”.

Заметим, впрочем, что некоторые семейства решений уравнений гравитации могут содержать метрики, описывающие объекты обоих классов при разных значениях параметров [83].

4.2. Кротовая нора как машина времени

Работы по физике КН — разумеется, чисто теоретические, других пока нет — можно условно разделить на два дополняющих друг друга направления. В одних исследователи, заранее предполагая существование КН, обсуждают, что из этого может получиться. В других они (нередко это те же самые авторы) пытаются определить, при каких условиях — естественных или искусственных — КН могут возникнуть и каковы тогда будут их свойства и особенности.

Остановимся на вопросах “первого направления”, оставив на время в стороне вопрос о том, из чего “могут быть сделаны” КН. Будем говорить о наиболее интересном варианте КН, когда она соединяет различные области одного и того же пространства-времени. Для начала предположим, что где-то в космическом пространстве имеется стабильная КН, а вне ее пространство-время плоское или почти плоское, где, как обычно, работают законы специальной теории относительности. Тогда просто в силу симметрии пространства Минковского очевидно, что нора может быть не только межгалактическим тоннелем, но и ускорителем, и даже машиной времени. В самом деле, у норы две плоские асимптотики. В каждой из них можно выбрать систему отсчета (СО) (для простоты полагаем — инерциальную), в которой устье КН покоится. Но ниоткуда не следует, что это одна и та же СО для обоих устьев! Далее, пусть даже это одна и та же СО и в областях, где расположены устья, каким-либо разумным образом синхронизованы часы. Опять-таки, ниоткуда не следует, что, пройдя сквозь КН, путешественник попадет в ту же эпоху по этим согласованным часам — ведь метрика Минковского инвариантна относительно временных сдвигов. Другими словами, даже если нам откуда-либо известно, куда ведет данная (допустим, случайно попавшаяся в космосе) КН, мы, вообще говоря, не знаем, “в когда” она ведет.

Инвариантность метрики Минковского относительно бустов

— преобразований от одной инерциальной СО к другой — ведет к тому, что на выходе из КН путешественник может оказаться в движении с любой досветовой скоростью относительно исходной СО и (что важнее, если устья далеки друг от друга) относительно окружающих тел.

Таким образом, в зависимости от положения устьев, причем не просто в пространстве, а в пространстве-времени зависят возможная роль и функция КН.

Интересна и другая постановка вопроса: предположим, мы научились создавать КН здесь и сейчас, с вполне макроскопическими параметрами, так что от одного устья к другому можно пройти сквозь нору за считанные секунды, а сами устья расположены относительно близко друг от друга в пространстве — скажем, для наглядности, где-то в околоземном пространстве. Можно ли превратить такую нору в машину времени? В статье Торна, Морриса и Юртсевер [228] показано, как это сделать: идея заключается в том, чтобы одно из устьев — обозначим его буквой А — оставить на месте, а другое, Б (а оно должно в принципе вести себя как обычное массивное тело) разогнать до скоростей, сравнимых со скоростью света, а потом вернуть обратно и затормозить рядом с А. Тогда, за счет известного эффекта СТО — замедления хода часов на движущемся теле по сравнению с неподвижным — на устье Б пройдет меньше времени, чем на устье А. Насколько меньше, зависит от достигнутой скорости и от продолжительности путешествия. Здесь, по сути дела, проявляется известный парадокс близнецов в СТО: близнец, вернувшийся из полета к звездам с релятивистскими скоростями, оказывается младше брата-домоседа. Если разница во времени между устьями составляет, к примеру, полгода, то, находясь возле устья А в январе 2008 года, мы сможем увидеть сквозь КН зелень прошедшего лета-2007 — и реально в это лето вернемся, пройдя нору насквозь! Приблизимся снова к воронке А (она, как мы, договорились, где-то неподалеку) и еще раз пройдем сквозь нору от А к Б — попадем прямиком в прошлогодний снег января 2007-го. И

так сколько угодно раз. Проходя нору в обратном направлении, т.е. ныряя в устье Б, мы совершим скачок на полгода в будущее. Итак, осуществив одну-единственную манипуляцию с одним из устьев, мы получаем машину времени, которой можно пользоваться сколько угодно раз. Естественно, в предположении, что нора устойчива или что мы сами в состоянии поддерживать ее работоспособность.

Приведем начало статьи [228], оно говорит само за себя. “Обычно физики-теоретики задают вопрос: — Каковы законы физики?, или: Что эти законы предсказывают для нашей Вселенной? В данной работе мы вместо этого спрашиваем: Какие ограничения накладывают законы физики на деятельность сколько угодно развитых цивилизаций? Это приводит к некоторым интригующим размышлениям о природе самих законов.”

Вскоре после работ группы Торна (1988) появилась статья Новикова и Фролова [148], где показано, что даже без специальных усилий КН во внешнем гравитационном поле неизбежно с течением времени превращается в машину времени. В самом деле, предположим, что одно из устьев (А) стабильной КН находится вблизи поверхности нейтронной звезды, другое (Б) — вдали от нее. Тогда темп хода часов вблизи от А медленнее, чем в районе локализации Б, эта разница накапливается и появляются замкнутые временноподобные траектории, проходящие сквозь нору. Машина времени готова.

Возможное существование КН серьезно повлияло бы и на физику черных дыр — см. обсуждение в той же статье [148]. Так, ничто не мешает предположить, что устье А КН находится вне черной дыры, а Б — внутри ее горизонта. Такую нору вполне можно использовать и для спасения упавшего в черную дыру космонавта, и для исследования внутренних свойств самой дыры. Строго говоря, горизонт черной дыры тогда перестает быть горизонтом событий.

Итак, если в нашем распоряжении есть стационарная КН макроскопических размеров с устьями, не слишком удаленны-

ми друг от друга, есть и машина времени.

Из приведенных рассуждений следует важный вывод, что законы физики (в число которых мы здесь включаем и понятие о гравитации как искривлении пространства-времени) в принципе не запрещают появления КН, а следовательно, появления замкнутых временноподобных и световых траекторий, нарушающих принцип причинности. В результате в работах по теоретической физике всерьез обсуждаются парадоксы, которым, казалось бы, место лишь в фантастической литературе. Как насчет встречи с другим (третьим, десятым) экземпляром самого себя? А известный “парадокс бабушки” — возможность забраться в прошлое и убить собственную бабушку еще в детстве (или, в менее кровавом варианте — просто жениться на юной бабушке, отбив ее у юного дедушки)? И откуда тогда взялся сам герой? Если всц это и можно примирить с нашими представлениями о мире и, в конце концов, с элементарным здравым смыслом, то с большим трудом.

Один из известных способов преодоления указанных трудностей опирается на гипотезу параллельных миров, называемую также гипотезой Мультиверса, тесно связанную с предложенной Эвереттом многомировой интерпретацией квантовой теории [128, 144]. Как известно, квантовые процессы носят вероятностный характер, и любой из них может иметь несколько исходов. По Эверетту, все возможные исходы реализуются, но каждый в своей собственной вселенной. Таким образом, каждое квантовое явление (а это, между прочим, все взаимодействия частиц в любой части Вселенной) увеличивает число параллельных миров. При всей экзотичности и явной, вроде бы избыточной громоздкости, такая интерпретация логически непротиворечива и довольно естественно объясняет не только квантовые парадоксы, но и упомянутый “парадокс бабушки”: прилетая в прошлое, путешественник фактически сотворяет новую вселенную, где, быть может, и есть похожая бабушка (точнее — в тот момент девушка), но вовсе не та, что впоследствии родила его маму...

Другой возможный выход (более экономный, но, пожалуй, более трудный для восприятия) — предложенный Новиковым “принцип самосогласованности” [237], согласно которому “локально реализуется то, что самосогласовано глобально”. Это означает [8], что события во временной петле влияют друг на друга по замкнутому циклу, не только прошлое определяет будущее, но и будущее участвует в формировании прошлого, что, как ни странно, не ведет к нарушению принципа причинности.

Есть, впрочем, и радикально-консервативная точка зрения, согласно которой временных петель в природе нет и не может быть — что-то вроде хронологической цензуры. Как любая цензура, она вряд ли может рассчитывать на симпатии многих наших соотечественников и авторов в том числе. А если говорить серьезно, эта гипотеза, как и любая другая, нуждается в доказательствах на основе еще более фундаментальных принципов. (Нетрудно понять, что как и любую гипотезу утверждающую, что чего-то не может быть, эксперимент может только опровергнуть, но не доказать.)

4.3. Кротовые норы как решения уравнений гравитации

Работы второго направления более многочисленны и не менее интересны: к ним относится поиск конкретных моделей КН и исследование их специфических свойств, которые и определяют их реализуемость или нереализуемость, что с этими норами можно делать и как использовать.

Прежде всего, если КН, или — шире — пространства, содержащие временные петли, и не запрещены полностью, они должны быть в природе достаточно редкими, иначе они бы давно уже наблюдались. Следовательно, если физическая теория претендует на описание реальности, она должна содержать механизм, по меньшей мере до крайности затрудняющий образование макроскопических КН.

ОТО, вне всякого сомнения, претендует на описание реальности. Однако в ней найдено немало решений, описывающих КН и другие пространства с замкнутыми временными петлями; но все они, как правило, по тем или иным причинам признаются либо нереалистическими, либо, так сказать, неопасными. Так, весьма интересное решение уравнений Эйнштейна указал Гедель [151]: это однородная стационарная Вселенная, вращающаяся как целое. Она содержит замкнутые временноподобные траектории, однако расчет показывает, что минимальная временная протяженность такой петли много больше времени существования Вселенной.

Важные ограничения следуют из самой структуры уравнений Эйнштейна. Идея их получения состоит в том, что, подставляя в левую часть уравнений метрику, обладающую теми или иными свойствами, можно найти явные ограничения на возможный вид правой части (то есть ТЭИ), совместимый с этими свойствами. Проследим, как это делается, на знакомом нам примере статических сферически-симметричных пространств.

4.3.1. Сферически-симметричные кротовые норы. Общие свойства

Будем, как и раньше, исходить из общего вида (3.1) статической сферически-симметричной метрики

$$ds^2 = e^{2\gamma} dt^2 - e^{2\alpha} du^2 - e^{2\beta} d\Omega^2, \quad (4.5)$$

и рассмотрим некоторые общие свойства статических сферически-симметричных КН, представляя ТЭИ в виде

$$T_{\mu}^{\nu} = \text{diag}(\rho, -p_r, -p_{\perp}, -p_{\perp}), \quad (4.6)$$

где ρ , p_r , p_{\perp} — соответственно, плотность, радиальное давление и поперечное давление, которые, вообще говоря, могут быть произвольными функциями радиальной координаты u . Это наиболее общий вид ТЭИ, совместимый с метрикой (4.5).

Воспользуемся двумя различными радиальными координатами.

В квазиглобальных координатах¹¹ ($\alpha + \gamma = 0$) метрика имеет вид (3.30),

$$ds^2 = A(u)dt^2 - \frac{du^2}{A(u)} - r^2(u)d\Omega^2, \quad (4.7)$$

а условия на горловине КН $u = u_{\text{th}}$, выражающие отсутствие горизонта и конечный минимум радиуса r , имеют вид $A > 0$, $r > 0$, $r' = 0$, $r'' > 0$. Тогда два уравнения Эйнштейна (см. (3.5), (3.6))

$$R_0^0 - R_1^1 = 2A \frac{r''}{r} = -\varkappa(\rho + p_r), \quad (4.8)$$

$$G_1^1 = \frac{1}{r^2} [-1 + A'rr' + Ar'^2] = \varkappa p_r \quad (4.9)$$

приводят к следующим условиям на горловине:

$$\rho + p_r < 0, \quad p_r < 0. \quad (4.10)$$

Заметим, что на знак плотности ограничения нет. Условия (4.10) носят локальный характер и не зависят от предположений о свойствах системы вдали от горловины. Ясно, что оба неравенства весьма необычны для макроскопической материи; правда, для полевой материи второе из них легко нарушается, например, для радиального электрического поля имеем, согласно (3.14), $p_r = -\rho$.

Нетрудно убедиться, что первое из неравенств (4.10) нарушает одно из важнейших условий, которым удовлетворяет большинство известных видов материи, а именно — световое энергетическое условие (СЭУ). В общем виде это условие выражается неравенством

$$T_{\mu\nu}\xi^\mu\xi^\nu \geq 0, \quad \forall \xi^\mu : \xi^\nu\xi_\nu = 0. \quad (4.11)$$

¹¹В этом разделе обозначаем квазиглобальную координату буквой u вместо ρ , так как символом ρ обозначена плотность.

В метрике (4.5) для светоподобного вектора

$$\xi^\mu = (e^{-\gamma}, e^{-\alpha}, 0, 0)$$

это условие принимает вид

$$T_0^0 - T_1^1 \geq 0. \quad (4.12)$$

Для тензора (4.6) оно записывается в виде $\rho + p_r \geq 0$, в противоположность (4.10).

Материя, нарушающая СЭУ, получила название экзотической. Вывод о том, что именно такая материя необходима для существования КН, не ограничивается статикой и сферической симметрией и носит весьма общий характер: он справедлив для горловин произвольной пространственной конфигурации, как стационарных, так и зависящих от времени [168, 169].

Теперь воспользуемся координатами кривизн, в которых метрику можно записать в виде

$$ds^2 = e^{2\gamma(r)} dt^2 - e^{2\alpha(r)} dr^2 - r^2 d\Omega^2. \quad (4.13)$$

Отождествляя метрики (4.7) и (4.13), несложно убедиться, что на горловине, если она существует, выполняются условия $|\gamma| < \infty$, $e^{-2\alpha} \equiv B(r) = 0$, $dB/dr \neq 0$. Одно из уравнений Эйнштейна (3.5) имеет вид

$$-G_0^0 = \frac{1}{r^2}(1 - e^{-2\alpha}) + \frac{2\alpha' e^{-2\alpha}}{r} = \varkappa\rho \quad (4.14)$$

и может быть проинтегрировано: при подстановке $e^{-2\alpha} = 1 - 2m(r)/r$ оно принимает вид $2m'(r) = \varkappa\rho r^2$ и, следовательно,

$$m(r) = 4\pi G \int_{r_0}^r \rho r^2 dr, \quad (4.15)$$

где r_0 — константа интегрирования. Функция $m(r)$ называется массовой функцией; в случае асимптотически плоской конфигурации с регулярным центром она действительно дает массу

Шварцшильда при интегрировании от 0 до больших значений радиуса, где метрика приближенно шварцшильдовская.

В случае асимптотически плоской КН, интегрируя в (4.15) от горловины $r = r_{\text{th}}$ до больших значений радиуса, получаем

$$r_{\text{th}} = 2m - \varkappa \int_{r_{\text{th}}}^{\infty} \rho r^2 dr, \quad (4.16)$$

так как на горловине $B(r) \equiv 1 - 2m(r)/r = 0$. Вспоминая, что $2m(\infty) = 2m$ есть шварцшильдовский радиус, соответствующий массе m , приходим к выводу, что *радиус горловины КН больше шварцшильдовского радиуса, соответствующего массе КН на ее плоской асимптотике, если плотность материи, порождающей нору, отрицательна и меньше шварцшильдовского, если плотность положительна.*

Это действительно важное ограничение: если мы хотим получить КН не астрономических масштабов, а с горловиной размером порядка метров или километров, то придется учитывать, что метры и километры — это масштаб радиусов Шварцшильда для планетарных и звездных масс (для Земли радиус Шварцшильда порядка 1 см, для Солнца — около 3 км). Чтобы устье КН не обладало мощным полем тяготения, соответствующим массе звезды или планеты, для построения такой норы нужна не просто экзотическая материя, а материя с отрицательной плотностью.

4.3.2. Кротовые норы с безмассовыми скалярными полями

В этом и последующих разделах мы рассмотрим несколько примеров точных решений уравнений гравитации, описывающих КН.

КН в решении анти-Фишера

Простейший пример — рассмотренное ранее решение анти-Фишера (3.101) при $\varepsilon = -1$: в случае $k < 0$ оно описывает КН

с двумя пространственными асимптотиками при $u = 0$ и $u = \pi/|k|$. Метрика имеет вид

$$\begin{aligned} ds^2 &= e^{-2hu} dt^2 - \frac{k^2 e^{2hu}}{\sin^2(ku)} \left[\frac{k^2 du^2}{\sin^2(ku)} + d\Omega^2 \right] = \\ &= e^{-2hu} dt^2 - e^{2hu} [d\rho^2 + (k^2 + \rho^2) d\Omega^2], \end{aligned} \quad (4.17)$$

где гармоническая координата u выражается через квазиглобальную координату ρ , заданную на всей числовой оси \mathbb{R} , формулой $|k|u = \text{arctg}(\rho/|k|)$.

При $m > 0$ КН притягивает окружающую пробную материю на первой асимптотике ($\rho \rightarrow \infty$) и отталкивает на второй ($\rho \rightarrow -\infty$) и наоборот в случае $m < 0$. При $m = 0$ получается простейшее решение, называемое КН Эллиса (хотя фактически в работе Х. Эллиса [141] обсуждались анти-фишеровские КН с произвольной массой).

Горловина КН находится при значении $\rho = h$ и имеет размер

$$r_{\text{th}} = (h^2 + k^2)^{1/2} \exp\left(\frac{h}{k} \text{arctg} \frac{h}{k}\right). \quad (4.18)$$

В недавней статье [152] показано, что антифишеровские КН неустойчивы относительно малых возмущений и что характерное время роста возмущений имеет порядок r_{th}/c , т.е. времени распространения светового луча на расстояние порядка размера горловины.

Известны аналоги метрики (4.17) с ненулевым электрическим зарядом [74], которые можно получить из (3.95), (3.96) при $\varepsilon = -1$. Зависимость их устойчивости или неустойчивости от величины заряда пока не исследована.

КН с конформным скалярным полем

В случае конформного скалярного поля с лагранжианом (3.73), (3.76) существует семейство решений уравнений поля вида (3.103)–

(3.104), с метрикой [74]

$$ds^2 = (y + y_0)^2 \left[\frac{dt^2}{(y + y_1)^2} - h^2 \frac{(y + y_1)^2}{y^4} (dy^2 + y^2 d\Omega^2) \right], \quad (4.19)$$

$$y_0 = \text{th } hu_0; \quad y_1 = \text{cth } hu_1 \quad (4.20)$$

и скалярным полем

$$\phi = \sqrt{6} \frac{1 + yy_0}{y + y_0}. \quad (4.21)$$

В электронейтральном случае $Q = 0$ имеем $y_1 = 1$; пространственной бесконечности соответствует $y = \infty$, где скалярное поле $\phi = \phi_\infty = \sqrt{6}y_0$.

Метрика (4.20) описывает КН при $y_0 > 0$; при этом координата y задана при всех $y > 0$. Нетрудно проверить, что при $y \rightarrow 0$, как и при $y \rightarrow \infty$, метрика становится плоской — имеет вторую пространственную бесконечность, где $e^{2\gamma} \rightarrow (y_0/y_1)^2$ и $\phi \rightarrow (1/y_0)\sqrt{6}$. Таким образом, получена КН с горловиной (минимумом r) при $y = \sqrt{y_0 y_1}$.

Метрика (3.97) (в других обозначениях) и ее обобщения на другие скалярные поля с неминимальной связью (лагранжиан (3.73), $f(\phi) = 1 - \xi\phi^2$, $h(\phi) = 1$, где $\xi > 0$ — произвольная постоянная) обсуждались Барсело и Виссером [53, 54] как примеры возникновения КН даже со скалярными полями с нормальным знаком кинетической энергии.

Однако у всех КН, поддерживаемых скалярными полями с неминимальной связью и, в более общем случае, КН, возникающих за счет конформных продолжений в СТТ [77] есть отличительная особенность — смена знака эффективной гравитационной постоянной на переходной сфере, на которой функция $f(\phi)$ в лагранжиане (3.73) проходит через нулевое значение. В результате вторая пространственная бесконечность таких КН расположена в области с отрицательной эффективной гравитационной постоянной. Реалистичность таких конфигураций довольно сомнительна: они, очевидно, не могут связывать отдаленные области одной вселенной, а их дальнейшее устье должно находиться в

мире с очень необычными с нашей точки зрения законами физики.

Кроме того, в статьях [88, 89] была установлена неустойчивость по меньшей мере КН вида (3.104), либо электронейтральных, либо с небольшими электрическими зарядами (по сравнению со скалярным зарядом) относительно сферически-симметричных возмущений, и есть основания полагать, что такая неустойчивость носит общий характер, так как она связана с поведением полей вблизи поверхности перехода $f = 0$.

4.3.3. Построение кротовых нор с помощью решения следа уравнений Эйнштейна

В этом разделе демонстрируется значительная степень произвола в решениях уравнений Эйнштейна в виде КН, если заранее не ограничивать используемый вид экзотической материи. Мы покажем, что даже если для простоты ограничиться материей с нулевым следом ТЭИ, решения в виде КН образуют множество, параметризованное одной произвольной функцией плюс константа интегрирования [102].

Будем исходить из общей статической сферически-симметричной метрики в координатах кривизн

$$ds^2 = e^{2\gamma(r)} dt^2 - e^{2\alpha(r)} dr^2 - r^2 d\Omega^2, \quad (4.22)$$

где $d\Omega^2 = d\theta^2 + \sin^2 \theta d\phi^2$ — линейный элемент на единичной сфере.

При нулевом следе ТЭИ справедливо уравнение $R = 0$,¹² которое можно записать как линейное уравнение первого порядка

¹²Уравнение (4.23) интересно еще и тем, что это единственное однозначно определенное уравнение гравитационного поля в мире на бране типа Рэндалл и Сундрума [253], так как полная система уравнений для гравитации на бране [276] содержит вклад, зависящий от 5-мерного тензора Вейля, имеющий нулевой след. Поэтому метрики КН, полученные при помощи описанных здесь алгоритмов, могут рассматриваться как метрики вакуумных КН в мире на бране [102]. Основные понятия теории мира на бране см. в главе 8.

относительно функции $F(r) := r e^{-2\alpha}$:

$$F_r(2 + r\gamma_r) + F(2r\gamma_{rr} + 2r\gamma_r^2 + 3\gamma_r) = 2, \quad (4.23)$$

где индекс r означает d/dr . Его общее решение имеет вид

$$F(r) = \frac{2e^{-2\gamma+3\Gamma}}{(2+r\gamma_r)^2} \int (2+r\gamma_r) e^{2\gamma-3\Gamma} dr, \quad (4.24)$$

где

$$\Gamma(r) = \int \frac{\gamma_r dr}{2+r\gamma_r}. \quad (4.25)$$

Выбирая произвольную гладкую функцию $\gamma(r)$, получаем $F(r)$ из (4.24) и, после фиксации констант интегрирования, метрика полностью известна.

Из решения (4.24) метрики КН (а также ЧД [83]) строятся алгоритмически с помощью задания производящей функции $\gamma(r)$ с нужными свойствами.

Отметим для справки, что во многих статьях о КН, начиная с [228], функция $e^{2\alpha}$ в координатах кривизн записывается в виде $[1 - b(r)/r]^{-1}$, где $b(r)$ называют “функцией формы” (shape function). Это название, на наш взгляд, не вполне удачно, так как $b(r)$ несет информацию о профиле КН в весьма опосредованном виде; скорее, функцией формы следовало бы назвать $r(l)$ в формуле (4.26). Фактически (см. (4.15)) $b(r) = 2m(r)$, где $m(r)$ — массовая функция. Введенная здесь функция $F(r)$ равна $r - b(r) = r - 2m(r)$.

Выясним, как именно следует выбирать функцию $\gamma(r)$ и константу интегрирования в решении (4.24), чтобы получить метрику КН.

Координата r , удобная для решения уравнения (4.23), не является допустимой координатой во всем пространстве КН, так как величина r в нем должна иметь по меньшей мере один минимум, соответствующий горловине, и решение в терминах r поэтому распадается по крайней мере на две ветви. В качестве до-

пустимой координаты можно взять, например, гауссову координату l (истинную длину в радиальном направлении), связанную с r соотношением $l = \int e^\alpha dr$, и метрика тогда имеет вид

$$ds^2 = e^{2\gamma(l)} dt^2 - dl^2 - r^2(l) d\Omega^2. \quad (4.26)$$

Мы ищем статические дважды асимптотически плоские решения для проходимых КН. Соответственно, требуем:

- 1) $l \in \mathbb{R}$; $r \approx |l| \rightarrow \infty$ и $\gamma = \text{const} + O(r^{-1})$ при $l \rightarrow \pm\infty$;
- 2) обе функции $r(l) > 0$ и $\gamma(l)$ должны быть гладкими (по меньшей мере C^2) во всей области $l \in \mathbb{R}$.

Эти условия гарантируют отсутствие сингулярностей и горизонтов (которые соответствуют $\gamma \rightarrow -\infty$). Они также означают, что функция $r(l)$ имеет по меньшей мере один минимум $r_{\min} > 0$ (горловину) при некотором конечном l . Возвращаясь к функциям переменной r , видим, что на плоской асимптотике $e^\alpha \rightarrow 1$ и $F(r) \approx r$.

Предположим без потери общности, что минимум $r(l)$ (горловина) находится при $l = 0$. Тогда $r(0) = r_0 > 0$, $r_l(0) = 0$ и (в общем случае) $r_{ll}(0) > 0$, где индекс l обозначает d/dl . Вблизи $l = 0$ имеем $r - r_0 \sim l^2$, следовательно, метрическая функция $e^{2\alpha(r)}$ ведет себя как $(r - r_0)^{-1}$, и $F(r) = r e^{-2\alpha} \sim r - r_0$. Иными словами, простой ноль $F(r)$ является признаком горловины КН при условии, что функция $\gamma(r)$ при том же значении r сохраняет гладкость и конечность.

С другой стороны, производная $\gamma_l(0)$ может быть нулевой (что обязательно имеет место, если КН симметрична относительно своей горловины) или ненулевой. Если $\gamma_l(0) = 0$, получаем $\gamma_r(r_0) < \infty$. Если, напротив, $\gamma_l(0) \neq 0$, то около r_0 имеем $\gamma_r \sim 1/|l| \sim 1/\sqrt{r - r_0}$, так что

$$\gamma(r) \approx \gamma(r_0) + k\sqrt{r - r_0}, \quad k > 0. \quad (4.27)$$

Здесь нельзя полагать $k < 0$, так как тогда получилось бы, что выражение $2 + r\gamma_r$ меняется от значения 2 на пространственной

бесконечности до $-\infty$ при $r = r_0$, так что $2 + r\gamma_r = 0$ при некотором $r > r_0$, порождая сингулярность в (4.24).

Теперь мы можем выделить класс метрик, описывающих симметричные КН (КН1) и класс метрик, возможно, описывающих асимметричные КН (КН2) на основе решения (4.24).

(КН1). Задаем функцию $\gamma(r)$, гладкую в области $r_0 \leq r < \infty$, $r_0 > 0$, таким образом, что $\gamma(\infty) = 0$, $\gamma_r(r_0) < \infty$, и $2 + r\gamma_r > 0$ во всей области. Фиксируем константу интегрирования в (4.24), выполняя интегрирование от r_0 до r . Тогда функции $\gamma(r)$ и $F(r)$ определяют КН с горловиной при $r = r_0$, симметричную относительно этой горловины.

В самом деле, по построению, $F(r) \sim r - r_0$ вблизи r_0 . Вводя новую координату x соотношением $r = r_0 + x^2$, получаем $e^{2\alpha} dr^2 \sim (r - r_0)^{-1} dr^2 = 4dx^2$, что приводит к регулярной метрике, в которой все коэффициенты являются четными функциями от $x \in \mathbb{R}$. Как $x \rightarrow +\infty$, так и $x \rightarrow -\infty$ суть плоские асимптотики.

Каждая функция $\gamma(r)$, выбранная таким способом, порождает семейство симметричных КН с нулевой скалярной кривизной, параметризованное радиусом горловины r_0 , который произволен в области, где функция $\gamma(r)$ регулярна и $2 + r\gamma_r > 0$.

Другой алгоритм применим к функциям $\gamma(r)$, которые ведут себя согласно (4.27).

(КН2-а). Задаем функцию $\gamma(r)$, гладкую в области $r_0 \leq r < \infty$, $r_0 > 0$, такую, что $\gamma(\infty) = 0$, $2 + r\gamma_r > 0$ во всей области, и что вблизи r_0 выполнено условие (4.27). Тогда при подходящих значениях константы интегрирования в (4.24) сфера $r = r_0$ является горловиной, и решение гладко продолжается за нее.

В самом деле, решение (4.24) можно переписать в виде

$$F(r) = \frac{e^{-2\gamma+3\Gamma}}{(1 + \frac{1}{2}r\gamma_r)^2} \left[\int_{r_0}^r (1 + \frac{1}{2}r\gamma_r) e^{2\gamma-3\Gamma} dr + C \right]. \quad (4.28)$$

Пусть $C > 0$. Тогда $F(r)$ ведет себя около r_0 как $r - r_0 =: x^2$, и $\gamma = \gamma(r_0) + kx + O(x^2)$. Метрика ведет себя гладко при $r = r_0$ ($x = 0$) в терминах новой координаты x и может быть продолжена за эту сферу. Нельзя, однако, гарантировать, что это продолжение приведет к еще одной пространственной бесконечности и даст асимметричную КН, так как продолжение функций $\gamma(x)$ и $F(x)$ на отрицательные x может привести к горизонту или сингулярности.

При выборе $C \leq 0$ в (4.28) получаем еще две ситуации:

(КН2-б). Если $C < 0$, то $F(r_0) < 0$; вспоминая, что $F \sim r$ при больших r , видим, что $F(r) = 0$ при некотором $r = r_1 > r_0$, где γ_r конечно, мы возвращаемся к обстоятельствам, описанным в алгоритме КН1 и получаем симметричную КН с горловиной при $r = r_1$.

(КН2-в). Если $C = 0$, то около r_0 имеем $F(r) \sim (r - r_0)^{3/2}$, и метрику при $r = r_0$ регуляризует другая подстановка: $r - r_0 = \xi^4$. В результате получаем

$$\gamma = \gamma(r_0) + k\xi^2 + \text{другие четные степени } \xi.$$

Это снова симметричная КН, но здесь радиус r как функция допустимой координаты $\xi \in \mathbb{R}$ ведет себя вблизи горловины как $r_0 + \text{const} \cdot \xi^4$.

4.3.4. Альтернативная гравитация и вакуум как источники кротовых нор

Приведенные в предыдущем разделе алгоритмы показывают множественность и разнообразие метрик КН даже в достаточно узком классе статических сферически-симметричных пространств. Это, однако, не снимает вопроса о возможных реальных материальных источниках КН, удовлетворяющих — при той же симметрии — условиям (4.10).

Как мы видели, обычные классические (нефантомные) поля к КН не приводят, так как все они удовлетворяют СЭУ (4.11).

Этот результат получен в теории Эйнштейна. Он, однако, допускает обобщение на большой класс альтернативных теорий гравитации, в которых уравнения гравитации после некоторого конформного отображения сводятся к уравнениям ОТО с дополнительным нефантомным скалярным полем. Таковы скалярно-тензорные теории (СТТ) гравитации при условии $\varepsilon = +1$ (см. раздел 3.7, формулы (3.73)–(3.75)), довольно популярные в последнее время $F(R)$ -теории, в которых вместо лагранжиана ОТО ($\text{const} \cdot R$) гравитационный лагранжиан берется в виде некоторой функции $F(R)$ (они будут обсуждаться в главе “Многомерие”), а также класс теорий, объединяющий СТТ и $F(R)$ -теории, в которых лагранжиан содержит функцию двух переменных $F(\phi, R)$. Преобразование от исходных уравнений относительно физической метрики $g_{\mu\nu}$ (так называемой картины Йордана) к уравнениям ОТО со скалярными полями (картина Эйнштейна) получены в общем виде в статьях [298] (для СТТ), [57, 58, 136, 142, 184, 221, 222, 310] для $F(R)$ -теорий и в [220] для $F(\phi, R)$ -теорий.

В самом деле, пусть метрика $g_{\mu\nu}$ в исходном многообразии \mathbb{M}_J (картина Йордана) и метрика $\bar{g}_{\mu\nu}$ в многообразии \mathbb{M}_E (картина Эйнштейна) связаны конформным отображением

$$g_{\mu\nu} = f(x)\bar{g}_{\mu\nu}, \quad (4.29)$$

где $f(x)$ — всюду регулярная функция пространственно-временных координат, ограниченная как сверху, так и снизу некоторыми положительными константами. Согласно [168], в ОТО для существования статической горловины (определенной как минимальная 2-поверхность в 3-пространстве) необходимо, чтобы материальный источник в уравнениях Эйнштейна нарушал СЭУ. Следовательно, если СЭУ не нарушено, не могут существовать не только КН, но даже горловины. Далее, при сделанных предположениях преобразование (4.29) всегда переводит плоскую асимптотику в \mathbb{M}_E в плоскую асимптотику в \mathbb{M}_J и обратно (хотя соответствующие шварцшильдовские массы могут различаться из-за непостоянства f). Если предположить, что имеется

дважды асимптотически плоская КН в \mathbb{M}_J , то каждой ее плоской асимптотике соответствует аналогичная асимптотика в \mathbb{M}_E , и в силу гладкости преобразования мы получаем КН в \mathbb{M}_E , в противоречии со сказанным выше. Таким образом, статические асимптотически плоские КН отсутствуют и в \mathbb{M}_J .

Это простое рассуждение [87, 101] не требует никаких предположений о пространственной симметрии и работает при весьма общих предположениях об исходной теории. Оно допускает обобщение и на динамические КН [101], так как и они в рамках ОТО требуют нарушения СЭУ, хотя в динамическом случае введение понятия горловины сталкивается с некоторыми сложностями [169].

Следует подчеркнуть, что горловины в картине Йордана не запрещены, так как в ней могут быть локальные нарушения энергетических условий. Однако, как мы видели, КН как глобальных конфигураций при этом не существует.

При ослаблении ограничений на вид функции $f(x)$ можно получить КН в картине Йордана в некоторых теориях (как мы видели выше на примере конформного скалярного поля), однако при этом, как выясняется, они обладают довольно патологическими свойствами, например на второй плоской асимптотике эффективная гравитационная постоянная либо стремится к бесконечности [79], либо имеет неверный знак (см. раздел 4.3.2).

Итак, без явного введения фантомных полей получить более или менее реалистические КН в классической теории поля оказывается довольно затруднительно. Однако достаточно многообещающе с точки зрения поисков КН выглядят квантовые явления.

Так, в главе 1 упоминалась концепция пространственно-временной пены, существующей на планковских масштабах за счет больших хаотических квантовых флуктуаций геометрии и топологии [305]. Элементами пены, помимо прочих образований, являются и кротовые норы. В обычных условиях наличие пространственно-временной пены незаметно на макроскопиче-

ских масштабах, но на самых ранних этапах эволюции Вселенной, во время сверхбыстрого расширения — инфляции (см. последующие главы) — некоторые из КН получают шанс расширяться до макроразмеров. А. Кирилловым (см., напр., [191, 192] и сделанные там ссылки) предложена идея о том, подобные КН размерами порядка килопарсек образуют целую хаотическую сеть во Вселенной, а однородное изотропное фридмановское пространство-время получается в результате усреднения. Показано, что неоднородности этой пенообразной космической структуры могут наблюдаться как “темная материя” (тем самым решая известную проблему недостатка массы в галактиках и скоплениях галактик для объяснения их динамических свойств [186, 242]); высказаны и соображения о том, как та же концепция может привести к решению проблемы темной энергии, иначе говоря, объяснить наблюдаемое в настоящую эпоху ускоренное расширение Вселенной. Впрочем, сами авторы признают [193], что пока в этой картине остается множество нерешенных вопросов (каково распределение КН по размерам и массам? почему они до сих пор не наблюдаются астрономами? участвуют ли они в настоящее время в расширении Вселенной? и т.д.). Едва ли не главный вопрос — всц тот же: какой материальный источник поддерживает существование КН больших размеров после их отрыва от планковских масштабов?

Возможный ответ на этот вопрос — поляризация вакуума квантовых полей. Как известно, согласно квантовой теории поля, вакуум — состояние поля с наименьшей энергией, в котором нет реальных частиц. Однако это состояние — не абсолютный покой, а хаотическое рождение и исчезновение пар всевозможных виртуальных частиц, в своей совокупности образующих и некую плотность энергии (вообще говоря, не малую и произвольного знака), и давление. То и другое зависит от конфигурации пространства, в которой рассматривается вакуум, и от присутствующей в нем материи. Это явление, которое и называется поляризацией вакуума, реально существует и многократно подтверждено

экспериментом. Например, хорошо известный эффект Казимира — притяжение двух близко расположенных плоских металлических пластин в пустоте — объясняется тем, что свойства вакуума электромагнитного поля различны в пространстве между пластинами и во всем остальном пространстве.

Расчет свойств физического вакуума — дело весьма трудоемкое, но некоторые полученные результаты выглядят обнадеживающе с точки зрения существования КН. Так, в отличие от макроскопической материи, вакуум вовсе не обязан подчиняться обычным энергетическим условиям, и в результате, как выяснилось [167, 201], поляризация вакуума в принципе способна поддерживать существование статической КН произвольных размеров. Правда, недавно показано [104], что так называемая полуклассическая материя, пригодная для построения КН, обладает особым рода неустойчивостью за счет квантовых флуктуаций: нора, появившись и будучи предоставленной сама себе, как бы пойдет рябью и в конце концов разрушится (но что возникнет в результате разрушения, остается неясным). Данная неустойчивость не носит всеобщего характера и, разумеется, не относится к любым мыслимым КН. Более того: если мы говорим о принципиальных, с точки зрения физики, возможностях “сколь угодно развитой цивилизации”, то вполне можно допустить, что, будучи заинтересованной в межзвездном транспорте, она запустит тот или иной механизм отрицательной обратной связи, гасящий возмущения.

4.4. Возможные наблюдательные проявления кротовых нор

В последние годы вопрос о возможном существовании КН во Вселенной стал рассматриваться достаточно серьезно, и причиной этому, очевидно, стало уже не раз упомянутое в этой книге открытие ускоренного расширения Вселенной. Ускорение сменило замедление сравнительно недавно по космологическим меркам

— $3 - 4 \cdot 10^9$ лет назад — и продолжает возрастать. Как известно, если эволюция изотропной Вселенной доминирована космологической постоянной (соответствующие плотность и давление связаны равенством $\varepsilon + p = 0$), то она расширяется по деситтеровскому закону, с постоянным ускорением. Чтобы ускорение росло со временем, необходимо предполагать $\varepsilon + p < 0$, что, в частности, нарушает световое энергетическое условие. Такой материи, получившей, как известно, название темной энергии, должно быть около 70 процентов в настоящую эпоху, и эта доля постоянно увеличивается: в отличие от обычного вещества, плотность которого уменьшается с увеличением объема, фантомная темная энергия ведет себя прямо противоположным образом — с расширением ее плотность растет.

Но именно такого рода экзотическая материя, каковы бы ни были ее происхождение и физическая природа, годится в качестве строительного материала для КН.

В связи с этим обсуждается гипотеза [183] о том, что некоторые активные галактические ядра, квазары и другие компактные астрофизические объекты могут быть входами в КН. Одна из интересных возможностей состоит в том, что вход в КН может быть источником монополярного магнитного поля. Для этого нет необходимости в существовании изолированных магнитных полюсов (монополей): силовые линии магнитного поля, как и электрического, могут просто пронизывать КН (магнитные и электрические “заряды без заряда” по Уилеру [305]).

Среди обсуждаемых эффектов, в которых может проявиться специфика КН по сравнению с ЧД и другими компактными объектами, можно упомянуть: (1) наблюдение отдаленных областей Вселенной сквозь горловину (в предположении, что горловина достаточно широка и заполняющее ее вещество прозрачно) — например, есть вероятность увидеть один и тот же объект мимо КН и сквозь КН с разными красными смещениями; (2) специфические виды ускорения заряженных частиц и формы струй вещества за счет радиальных магнитных полей [183]; (3) осо-

бые характеристики орбит пробных тел вблизи устьев КН и, в частности, возможность их осцилляций вокруг горловины, если в ней располагается минимум гравитационного потенциала; (4) эффект отклонения световых лучей (линзирование) вблизи КН, отличающийся от аналогичного эффекта вблизи обычных массивных тел и ЧД.

Численные оценки подобных эффектов делаются, но их предсказательная сила крайне мала из-за неопределенности природы экзотической материи, необходимой для образования КН. В самом деле, как мы видели в разделе 4.3.3, даже при условии равной нулю скалярной кривизны возможная геометрия КН включает произвольную функцию, а в общем случае — две произвольные функции. Весь этот произвол имеет место даже в предположении статичности и сферической симметрии, не говоря уже о более сложных и реалистических ситуациях. Кроме того, если экзотическая материя, образующая КН, занимает конечный объем в пространстве, то область вне его будет, как всегда, описываться вакуумными решениями уравнений гравитации (например, решением Шварцшильда) с присущими им вполне обычными характеристиками.

Впрочем, если какие-либо эффекты в астрономических наблюдениях окажутся объяснимыми с помощью КН, такая интерпретация будет без сомнения весьма интересной и интригующей.

Глава 5

Этапы развития Вселенной

5.1. Космологический принцип и уравнения Эйнштейна

При обсуждении многих проблем, касающихся свойств Вселенной в целом или ее части, доступной наблюдениям, можно абстрагироваться от отличительных свойств ее отдельных областей и заполняющей их материи. Это выражается в так называемом *космологическом принципе*, согласно которому все точки Вселенной и все направления в ней равноправны, иными словами, пространство-время однородно и изотропно. В силу уравнений Эйнштейна распределение материи тогда тоже должно быть однородным и изотропным. Космологический принцип достаточно хорошо согласуется с наблюдениями: области размерами ≥ 100 Мпк действительно мало отличаются друг от друга.

Предположение об однородности и изотропии пространства и распределенной в нем материи сильно упрощает расчеты и облегчает анализ физических процессов. Соответствующие уравнения очень кратко обсуждались в главе 2. Остановимся на них подробнее.

В системе имеются динамические переменные, связанные с материей (плотность, давление и т.д.) и с гравитацией (компо-

ненты метрического тензора $g_{\mu\nu}$). Они подчиняются уравнениям Эйнштейна

$$G_{\mu}^{\nu} \equiv R_{\mu}^{\nu} - \frac{1}{2}\delta_{\mu}^{\nu}R = -8\pi GT_{\mu}^{\nu} - \Lambda\delta_{\mu}^{\nu}, \quad (5.1)$$

где R_{μ}^{ν} — тензор Риччи, $R \equiv R_{\alpha}^{\alpha}$ — скалярная кривизна; обе эти величины выражаются через метрический тензор, его первые и вторые производные по координатам пространства-времени; тензор энергии-импульса (ТЭИ) T_{μ}^{ν} описывает свойства материи, заполняющей пространство — см. главу 2. В уравнениях учтен также и так называемый Λ -член, с космологической постоянной Λ , который, по-видимому, наблюдается экспериментально. Уравнения (5.1) могут быть записаны в другой, эквивалентной форме

$$R_{\mu}^{\nu} = -8\pi G \left(T_{\mu}^{\nu} - \frac{1}{2}\delta_{\mu}^{\nu}T \right) + \delta_{\mu}^{\nu}\Lambda, \quad (5.2)$$

где $T \equiv T_{\alpha}^{\alpha}$. В силу свернутых тождеств Бьянки (2.43) и уравнений Эйнштейна ТЭИ удовлетворяет равенствам (2.49)

$$\nabla_{\nu}T_{\mu}^{\nu} = 0, \quad (5.3)$$

которые имеют смысл законов сохранения (см. главу 2 и примечание к формуле (2.49)) и из которых выводятся уравнения движения материи.

В силу предположений об однородности и изотропии пространства квадрат интервала может быть записан в виде

$$ds^2 = g_{\mu\nu}dx^{\mu}dx^{\nu} = dt^2 - a^2(t) \left[\frac{dr^2}{1 - kr^2} + r^2d\Omega^2 \right], \quad (5.4)$$

где $d\Omega^2 = d\theta^2 + \sin^2\theta d\varphi^2$ — элемент длины на единичной сфере (как и в предыдущих главах 3 и 4). Здесь сделан определенный выбор координат: синхронное (истинное космологическое) время t и безразмерные сферические координаты r , θ , φ . Поведение гравитационного поля определяется масштабным фактором $a(t)$

(имеющим размерность длины) и параметром k , принимающим три значения в зависимости от знака кривизны 3-пространства: $k = +1$ для сферического, $k = 0$ для плоского и $k = -1$ для гиперболического пространства.

В метрике (5.4) ненулевые компоненты тензоров Риччи и Эйнштейна имеют вид

$$\begin{aligned} R_0^0 &= \frac{3\ddot{a}}{a}, & R_i^i &= \frac{1}{a^2}(a\ddot{a} + 2\dot{a}^2 + 2k), \\ G_0^0 &= -\frac{3}{a^2}(\dot{a}^2 + k), & G_i^i &= -\frac{1}{a^2}(2a\ddot{a} + \dot{a}^2 + k), \end{aligned} \quad (5.5)$$

где по подчеркнутому индексу нет суммирования.

В силу свойств симметрии тензора G_μ^ν , с учетом уравнений Эйнштейна, тензор T_μ^ν автоматически имеет вид

$$T_\mu^\nu = \text{diag}(\rho, -p, -p, -p), \quad (5.6)$$

где ρ и p — некоторые функции временной координаты. Нетрудно убедиться, что если в качестве материи рассматривается идеальная жидкость, то выбранные координаты принадлежат сопутствующей системе отсчета (СО) жидкости, а величины ρ и p следует интерпретировать как ее плотность и давление.

В самом деле, ТЭИ идеальной жидкости есть

$$T_\mu^\nu = (p + \rho)u_\mu u^\nu - p\delta_\mu^\nu, \quad (5.7)$$

где ρ и p — соответственно, плотность и давление жидкости, а u^μ — 4-скорость ее элемента. В сопутствующей СО (т.е. для наблюдателя, неподвижного относительно жидкости) 4-скорость имеет вид $u^\mu = (1, 0, 0, 0)$, и ТЭИ имеет вид (5.6).

Из (5.5) очевидно, что имеется два различных уравнения Эйнштейна; удобно выбрать для рассмотрения уравнение с левой частью G_0^0 , так как оно содержит лишь первую производную от масштабного фактора a , и след $R = 8\pi GT + 4\Lambda$:

$$3\frac{\dot{a}^2}{a^2} + \frac{3k}{a^2} = 8\pi G\rho + \Lambda, \quad (5.8)$$

$$\frac{6}{a^2}(a\ddot{a} + \dot{a}^2 + k) = 8\pi G(\rho - 3p) + 4\Lambda. \quad (5.9)$$

Еще одно уравнение первого порядка следует из закона сохранения (5.3):

$$\frac{d\rho}{da} = -\frac{1}{a}3(p + \rho),$$

или

$$\frac{d\rho}{dt} = -3H(p + \rho), \quad (5.10)$$

где величина $H := \dot{a}/a$, характеризующая скорость расширения, называется параметром Хаббла.

Легко убедиться, что уравнение (5.9) является следствием (5.8) и (5.10). Таким образом, имеем два независимых уравнения (5.8) и (5.10) для трех неизвестных функций $a(t)$, $\rho(t)$ и $p(t)$.

Очевидно, полный вид решения уравнений зависит от свойств материи, описываемых уравнением состояния — связью давления и плотности, $p = p(\rho)$. Часто эту связь можно выбрать в самом простом линейном виде

$$p = w\rho, \quad (5.11)$$

где $w = \text{const}$ имеет различные значения для разных видов материи:

$w = -1$ соответствует космологической постоянной,

$w = 0$ — пылевидному веществу (газу невзаимодействующих частиц),

$w = 1/3$ — неупорядоченному излучению или ультрарелятивистскому газу,

$w = 1$ — веществу максимальной жесткости, совместимой с причинностью, в котором скорость звука равна скорости света.

Рассмотрим некоторые геометрические свойства пространств ФРУ с метрикой (5.4) и произвольным масштабным фактором $a(t)$.

Пусть один наблюдатель находится в точке $r = r_0 = 0$, а второй — в точке с некоторым $r \neq 0$. Тогда последнее слагаемое в

(5.4) позволяет вычислить площадь сферы $r = \text{const}$, на которой находится второй наблюдатель в некоторый момент времени t : $S = 4\pi a^2(t)r^2$. Итак, даже если второй наблюдатель покоится, т.е. $r = \text{const}$, физический радиус сферы меняется пропорционально масштабному фактору $a(t)$:

$$R(t) \equiv a(t)r. \quad (5.12)$$

Наблюдаемые скопления галактик как раз и являются для нас примерами таких “вторых наблюдателей”. Координатный радиус такого скопления, r , фиксирован, в то время как физический радиус $R(t)$, с наблюдателем в центре сферы, растет со временем. Экспериментальным подтверждением увеличения физического радиуса служит уменьшение частоты света, испущенного вторым наблюдателем, при приеме первым (красное смещение).

Расстояния между точками пространства можно измерять разными способами, и результаты измерения (даже в пространстве Минковского) зависят от выбора СО, а в случае метрики, зависящей от времени, даже от выбора пространственного сечения (т.е. способа синхронизации часов) в рамках одной и той же СО. Мы будем пользоваться понятием мгновенного физического расстояния, т.е., расстояния в фиксированный момент времени в метрике (5.4). Такое расстояние измеряется множеством наблюдателей в СО, неподвижной в метрике (5.4) с часами, синхронизованными по космологическому времени t . При измерении вдоль радиуса r физическое расстояние между соседями, разделенными координатной разностью dr , есть $a(t)dr/\sqrt{1 - kr^2}$. Суммарное расстояние

$$R_{\text{inst}} = a(t) \int_0^r \frac{dr}{\sqrt{1 - kr^2}}, \quad (5.13)$$

измеренное всеми такими наблюдателями в один и тот же момент времени, равно мгновенному физическому радиусу.

Если ввести вместо r координату χ равенством

$$\chi = \int_0^r \frac{dr}{\sqrt{1 - kr^2}} = \begin{cases} \arcsin r, & k = 1, \\ r, & k = 0, \\ \operatorname{Arsh} r, & k = -1, \end{cases} \quad (5.14)$$

получим в терминах χ альтернативную запись метрики (5.4)

$$ds^2 = dt^2 - a^2(t) [d\chi + r^2(\chi)d\Omega^2],$$

$$r(\chi) = \begin{cases} \sin \chi, & k = 1, \\ \chi, & k = 0, \\ \operatorname{sh} \chi, & k = -1 \end{cases} \quad (5.15)$$

и крайне простое выражение для мгновенного расстояния (вдоль радиуса, при фиксированных значениях θ и φ) между наблюдателями $r = r(\chi_1)$ и $r = r(\chi_2)$:

$$R_{\text{inst}} = a(t)(\chi_2 - \chi_1). \quad (5.16)$$

Рассмотрим движение светового сигнала от центра $r = 0$ к произвольной сфере, отмеченной меткой r . В этом случае движение происходит вдоль радиуса ($d\theta = d\varphi = 0$). Из условия $ds = 0$, согласно (5.4) и (5.15), имеем

$$\int_{t_0}^{t_1} \frac{dt}{a(t)} = \int_0^r \frac{dr}{\sqrt{1 - kr^2}} = \chi. \quad (5.17)$$

Физическое расстояние, проходимое светом за интервал времени (t_0, t_1) , есть, разумеется, просто $t_1 - t_0$ (или $c(t_1 - t_0)$, если явно писать скорость света c), так как скорость света — универсальная постоянная. Однако представляет интерес другая величина — мгновенное физическое расстояние $R_{\text{hor}}(t_1, t_0)$ от центра до той точки (с радиальной координатой r_1 или χ_1), куда сигнал, испущенный при $t = t_0$, пришел в момент t_1 и измеренное в момент t_1 . На него влияет то, что за время, пока шел сигнал,

Вселенная успела расшириться.¹³ Такое расстояние в силу (5.17) есть

$$R_{\text{hor}}(t_1, t_0) = a(t_1)\chi_1 = a(t_1) \int_{t_0}^{t_1} \frac{dt}{a(t)}. \quad (5.18)$$

Эта величина называется радиусом горизонта частиц в момент t_1 для частиц, испущенных в момент t_0 . (Часто, опуская слово “радиус”, говорят “горизонт частиц” или даже просто “горизонт” в тех случаях, когда это не вызывает путаницы.)

Следует отметить, что понятие “горизонт частиц” имеет несколько определений различной степени строгости. Одно из них — это “расстояние, на которое уходит свет в течение некоторого отрезка времени”. Например, радиус горизонта частиц нашей Вселенной составляет порядка 10^{28} см.

Напомним также, что, кроме понятия горизонта частиц в космологии, в теории гравитации введен целый ряд различных понятий горизонтов, которые применяются прежде всего в физике черных дыр (см. гл. 3). Например, горизонт событий иногда определяют как геометрическое место точек, которых мог бы достигнуть свет из фиксированной удаленной точки за бесконечно большой интервал времени. Типичным примером такого горизонта служит горизонт событий ЧД Шварцшильда.

Некоторые решения для масштабного фактора

Рассмотрим временную зависимость масштабного фактора $a(t)$ для некоторых видов материи с уравнением состояния (5.11) в однородном изотропном пространстве с метрикой ФРУ (5.4). Напомним, что об однородности пространства в современную эпоху

¹³Расстояние, которое *прошел* сигнал, и расстояние, на которое он *удалился* в расширяющейся Вселенной, — совершенно разные величины. Поясним это на аналогии. Вы входите на движущийся вверх эскалатор, не спеша по нему поднимаетесь и сделали, допустим, 10 шагов — но за то же время вы удалились от нижней точки отнюдь не на 10 шагов, а на все 40 за счет скорости эскалатора!

можно говорить лишь имея в виду усреднение плотности вещества на масштабах порядка сотни мегапарсек.

Основными уравнениями для последующего анализа являются уравнения (5.8), (5.10). Как будет видно из дальнейшего, масштабный фактор во многих случаях (включая модели начальной стадии эволюции) много больше единицы, и слагаемым k/a^2 в уравнении (5.8) можно пренебречь. Тогда уравнение (5.8) приобретает вид

$$\dot{a}^2 = \frac{8\pi G}{3} \rho a^2. \quad (5.19)$$

Уравнение (5.10) с учетом (5.11) принимает вид

$$\frac{d\rho}{da} = -3\gamma \frac{\rho}{a}, \quad \gamma := w + 1, \quad (5.20)$$

с очевидным решением

$$\rho = C a^{-3\gamma} = C a^{-3(w+1)}. \quad (5.21)$$

Константа C находится из начальных условий: $C = \rho_{\text{in}} a_{\text{in}}^{3\gamma}$, где индекс “in” относится к начальному моменту рассматриваемой стадии эволюции Вселенной. Переходы между стадиями в реальности не могут быть резкими, но этим эффектом часто пренебрегают.

Подставляя выражение (5.21) в уравнение (5.19), нетрудно найти временную зависимость масштабного фактора:

$$a(t) = \left[a(t_{\text{in}})^{3\gamma/2} + \frac{2}{3\gamma} \sqrt{\frac{8\pi G}{3}} C (t - t_{\text{in}}) \right]^{2/(3\gamma)}.$$

Подставляя константу C , получаем следующую формулу для масштабного фактора для различных значений w или γ :

$$a(t) = a_{\text{in}} \left[1 + \frac{2}{3\gamma} \sqrt{\frac{8\pi G \rho_{\text{in}}}{3}} (t - t_{\text{in}}) \right]^{2/(3\gamma)}. \quad (5.22)$$

Влияние свойств материи отражено в параметре γ .

Общепринято деление эволюции Вселенной на 5 стадий: инфляция, разогрев (иногда говорится “рехитинг” или “прехитинг”), стадия доминантности излучения, стадия доминантности материи и современная стадия ускоренного расширения, называемая также вторичной инфляцией. Каждая из стадий, в зависимости от преобладания того или иного вида материи, приближенно характеризуется своим параметром γ .

Поскольку масштабный фактор быстро растет со временем, первым слагаемым в скобках чаще всего можно пренебречь. В результате получаем компактную формулу

$$a(t) \simeq a_{\text{in}} \left[\frac{2}{3\gamma} H_{\text{in}} \cdot (t - t_{\text{in}}) \right]^{2/(3\gamma)}. \quad (5.23)$$

Здесь $H_{\text{in}} \equiv H(t_{\text{in}}) = \sqrt{8\pi G\rho_{\text{in}}/3}$ — значение параметра Хаббла в начале данной стадии.

Временную зависимость параметра Хаббла нетрудно получить:

$$H(t) \equiv \frac{\dot{a}(t)}{a(t)} = \frac{2}{3\gamma t}. \quad (5.24)$$

(Отметим, что параметр Хаббла первое время после открытия расширения Вселенной назывался постоянной Хаббла, так как рассматривались лишь времена, близкие к современной эпохе, на которых величина H почти не менялась.)

Все стадии, кроме инфляционных, хорошо описываются формулами, полученными выше. Инфляционные стадии, приближенно описываемые метрикой де Ситтера, имеют ряд важных особенностей. Прежде чем обсуждать их, полезно рассмотреть основные свойства пространства де Ситтера.

5.2. Пространство де Ситтера

В начале прошлого века де Ситтер изучал свойства искривленного пространства, не содержащего материи. Единственное, что от-

личало его от привычного пространства Минковского — это присутствие ненулевого Λ -члена в уравнении (5.8). Своеобразные свойства этого пространства долгое время представляли лишь академический интерес. Но сейчас становится понятно, что начальная, инфляционная и конечная, современная, стадии развития нашей Вселенной неплохо описываются именно с помощью пространства де Ситтера. Поэтому полезно и поучительно будет посвятить некоторое время его изучению.

Этот тип пространств кратко обсуждался в главе 2, сейчас мы рассмотрим его подробнее. Поскольку материя отсутствует, $p = \rho = 0$, уравнение (5.8) сильно упрощается:

$$\dot{a}^2 - H^2 a^2 = -k, \quad H^2 \equiv \frac{\Lambda}{3} \quad (5.25)$$

(в предположении $\Lambda > 0$) и легко интегрируется:

$$k = 0: \quad a(t) = a_0 e^{\pm Ht}, \quad a_0 = \text{const} \\ \text{(пространственно-плоская Вселенная);} \quad (5.26)$$

$$k = 1: \quad a(t) = H^{-1} \text{ch}[H(t - t_0)], \quad t_0 = \text{const} \\ \text{(замкнутая Вселенная);} \quad (5.27)$$

$$k = -1: \quad a(t) = H^{-1} \text{sh}[H(t - t_0)], \quad t_0 = \text{const} \\ \text{(гиперболическая Вселенная).} \quad (5.28)$$

Выбор констант (т.е. выбор начала отсчета времени), такой, что $a_0 = H^{-1}$, $t_0 = 0$, а также знака плюс в (5.26) (это соответствует расширению Вселенной) приводит к удобному общему выражению для асимптотик всех трех выражений для $a(t)$ при больших временах:

$$a(t) = H^{-1} e^{Ht}, \quad t \gg H^{-1}. \quad (5.29)$$

Воспользуемся этим выражением для анализа и более глубокого понимания общих формул, полученных выше. Пусть два наблюдателя находятся в точках пространства с фиксированными значениями $r_1 = 0$ и $r_2 = r$, то есть оба они покоятся в той СО, в

которой записана метрика. Тем не менее, измеряемое ими физическое расстояние, согласно выражениям (5.12) и (5.29), растет со временем экспоненциально.

Важно также разобраться с расстояниями, на которые уходит свет в пространстве де Ситтера. Рассмотрим плоскую Вселенную ($k = 0$) с масштабным фактором $a(t) = H^{-1}e^{Ht}$ как самый простой случай, сохраняющий все необходимые свойства (к тому же, асимптотики всех трех вариантов одинаковы). Для плоских моделей введенные нами безразмерные координаты r и χ совпадают. Аналогично (5.17), получаем выражение

$$\Delta r(t, t') = \int_t^{t'} d\tau/a(\tau) = e^{-Ht} - e^{-Ht'} \quad (5.30)$$

для координатного расстояния, покрываемого световым сигналом, стартовавшим в момент t , к моменту t' . Следовательно, максимальное значение Δr , достижимое для света, испущенного в момент t , есть

$$\Delta r_{\text{hor}} \equiv \Delta r(t, t' \rightarrow \infty) = e^{-Ht}. \quad (5.31)$$

Значит, горизонт находится на конечном координатном расстоянии, и более того, чем позже был испущен свет, тем оно меньше. В координатном смысле размер горизонта уменьшается со временем.

С точки зрения физических расстояний ситуация другая: за время $(t' - t)$ свет удаляется на расстояние

$$R_{\text{phys}}(t, t') = a(t')\Delta r(t, t') = H^{-1} \left[e^{H(t'-t)} - 1 \right].$$

При фиксированном начальном моменте t расстояние растет экспоненциально, что разительно отличается от привычной ситуации в пространстве Минковского. При этом, согласно уравнению

$$R_{\text{hor}}(t) = R_{\text{phys}}(t, t' \rightarrow \infty) = H^{-1} e^{H(t'-t)}, \quad (5.32)$$

размер горизонта стремится к бесконечности.

Понятие горизонта является чрезвычайно важным в космологии, поскольку определяет динамические процессы в ранней Вселенной. В инфляционный период, вскоре после рождения Вселенной, пространство можно было приближенно рассматривать как деситтеровское, при этом размер горизонта составлял порядка 10^{-27} см. Именно таков был размер причинно связанных областей. Вспоминая, что комптоновская длина электрона — порядка 10^{-11} см, нетрудно видеть, что само наличие горизонта привносит специфику в физические процессы. На современном этапе преобладает темная энергия, что, скорее всего, означает вхождение в пространство де Ситтера. В этом случае современный размер горизонта составляет $\sim 10^{28}$ см, и бóльшие расстояния ненаблюдаемы.

Реально ли расширение в пространстве де Ситтера?

Рассмотрим еще раз метрику де Ситтера в виде (5.4) с плоским 3-пространством ($k = 0$):

$$ds^2 = dt^2 - H^{-2} e^{2Ht} (dr^2 + r^2 d\Omega^2), \quad (5.33)$$

где $H = \sqrt{\Lambda/3}$. Масштабный фактор меняется со временем, а с ним и расстояния между точечными объектами, неподвижными относительно космологической СО, в которой записана данная метрика.

В то же время преобразованием координат метрику можно свести к статическому виду — см. (3.18):

$$ds^2 = (1 - H^2 r^2) dt^2 - (1 - H^2 r^2)^{-1} dr^2 - r^2 d\Omega^2.$$

В этой метрике расстояния между телами, фиксированными в данной СО, не меняются. Существует и такая СО, в которой расстояния уменьшаются со временем — например, так будет, если взять знак минус в формуле (5.26).

Вернемся к метрике (5.33) и рассмотрим движение пробных частиц. Уравнение геодезической имеет вид

$$\frac{d^2 x^i}{ds^2} + \Gamma_{\mu\nu}^i \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} = 0, \quad (5.34)$$

Поскольку символ Кристоффеля $\Gamma_{tt}^i = 0$, одним из решений этого уравнения является $x^i(t) = \text{const}$. Координаты точек (и их разности) не зависят от времени. С другой стороны, физические измерения ($R_{\text{phys}}(t) = e^{Ht} R_{\text{phys}}(0)$) очевидно указывают на рост расстояний между покоящимися частицами.

Как же понять, расширяется ли Вселенная на самом деле, при том что пространство де Ситтера максимально симметрично и все его точки-события совершенно равноправны? Это, казалось бы, означает, что переход от одного момента времени к другому ничего не меняет и никакого расширения нет.

Простое объяснение кажущегося парадокса заключается в том, что расширяется не четырехмерное пространство-время, а 3-пространство той или иной СО. Более того, мы только что убедились, что линии времени $x^i = \text{const}$ той СО, в которой метрика имеет вид (5.33), являются геодезическими — а расстояния R_{phys} растут — следовательно, этот пучок (конгруэнция) геодезических расширяется. Другие конгруэнции геодезических в том же пространстве-времени могут сжиматься — это зависит от задания начальных данных в уравнении (5.34).

Можно привести еще один аргумент, подтверждающий реальность расширения пространства в метрике (5.33). Рассмотрим систему пробных частиц, однородно распределенных во всем пространстве и неподвижных в координатах (5.33), т.е. с мировыми линиями $x^i = \text{const}$. Важно, что они слабо, но взаимодействуют друг с другом (например, гравитационно). Если мы обнаружим, что взаимодействие между ними уменьшается, то это однозначно свидетельствует, что расстояния между частицами увеличиваются, то есть Вселенная действительно расширяется.

Вспомним, что в этих же координатах размер горизонта (5.31) уменьшается со временем по закону e^{-Ht} , тогда как ко-

ординатное расстояние между частицами фиксировано. Следовательно, для любых двух частиц существует момент времени, когда размер горизонта станет меньше расстояния между ними, что будет означать отсутствие причинной связи, то есть с этого момента взаимодействие между пробными частицами отсутствует. Иначе говоря, первоначально ненулевое взаимодействие между частицами должно со временем стремиться к нулю — Вселенная расширяется (а точнее, расширяется 3-пространство рассматриваемой системы отсчета).

Можно провести аналогичное рассуждение в терминах физических расстояний — результат будет, естественно, тот же.

5.3. Инфляция

Идея инфляции — сверхбыстрого расширения Вселенной на ранней стадии — была выдвинута в начале 80-х годов прошлого века. Успехи инфляционной парадигмы в объяснении наблюдаемых свойств Вселенной сделали ее идеи общепринятыми. Сейчас известно множество инфляционных сценариев, и выделить из них тот, который был реализован на практике, представляется сложной задачей. Число инфляционных моделей постоянно растет. Ниже обсуждаются свойства инфляционных моделей, построенных на базе идеи так называемой хаотической инфляции.

Прошли десятилетия с того момента, когда стало понятно, что хорошо известное расширение Вселенной, начиная с “горячей” (радиационно-доминированной) стадии, оставляет много вопросов, и стало очевидно, что Вселенная нетривиально развивалась до горячей стадии. Полный перечень проблем можно найти в [210], как и в большинстве учебников по космологии. Основные из них обсуждаются в разделе 5.6.

Путь решения этого набора проблем был заложен в работах [6, 158, 282].

Через некоторое время стало понятно, что инфляция в состоянии объяснить основной набор наблюдательных данных. Успе-

хи инфляционного сценария были столь впечатляющими, что он включен в Стандартную космологическую модель в качестве основной составляющей.

Наиболее простой и распространенный способ описания инфляционного периода заключается в следующем. Предполагается существование некоторого скалярного поля (инфлатона), которое эволюционирует совместно с гравитационным полем, им порождаемым. При некоторых условиях на потенциал этого поля, которые обсуждаются ниже, возникает ситуация, напоминающая деситтеровскую. Значит, размер пространства под горизонтом растет экспоненциально быстро, что и является главной особенностью инфляционного периода.

Плотность лагранжиана системы, состоящей из скалярного поля и гравитации, обычно записывается в виде

$$L = \sqrt{-g} \left\{ \frac{1}{16\pi G} R + \frac{1}{2} g^{\mu\nu} \partial_\mu \varphi \partial_\nu \varphi - V(\varphi) \right\}, \quad (5.35)$$

где $g = \det(g_{\mu\nu})$, G — гравитационная постоянная.

Забегаая вперед, скажем, что инфляционный процесс эффективно осуществляется при большой плотности энергии скалярного поля. Последняя может возникнуть благодаря квантовым флуктуациям. Оценим размер флуктуации, используя соотношение неопределенностей

$$\Delta E \Delta t \sim 1 (= \hbar) \quad (5.36)$$

Для простоты будем считать, что $V(\varphi) = m^2 \varphi^2 / 2$ (m — масса поля φ), а фоновое значение поля есть $\varphi_0 = 0$, т.е. рассматриваем низкие энергии. Флуктуация возникает в причинно связанной области, что накладывает ограничение на ее максимально возможный пространственный размер $\Delta l \sim \Delta t$ флуктуации. Энергия системы, описываемой лагранжианом (5.35), имеет вид

$$E = \int d^3x \sqrt{-g} \left\{ \frac{1}{2} g^{\mu\nu} \partial_\mu \varphi \partial_\nu \varphi + V(\varphi) \right\}. \quad (5.37)$$

Здесь скалярная кривизна R отброшена, поскольку оценка проводится только для скалярного поля φ . Учитывая все предыдущее, оценим интеграл (5.37) и из равенства (5.36) найдем пространственный размер флуктуации

$$\Delta l^3 \Delta t \left[\frac{1}{2} \left(\frac{\Delta \varphi}{\Delta t} \right)^2 + \frac{1}{2} m^2 \Delta \varphi^2 \right] \sim \Delta \varphi^2 \Delta l^2 [1 + m^2 \Delta l^2] \sim 1.$$

Следовательно, амплитуда флуктуации по порядку величины есть

$$\Delta \varphi \sim \frac{1}{\Delta l \sqrt{1 + m^2 \Delta l^2}}.$$

Это общее выражение. На его основе можно оценить энергию флуктуации с заданным пространственным размером, например порядка нарушения симметрии электрослабого и сильного взаимодействий, $\Delta l \sim 1/M_{\text{GUT}} \sim 1000/M_{\text{Pl}}$. Особенно простое выражение получится, если сделать естественное предположение $m \ll M_{\text{GUT}}$. В этом случае флуктуация потенциальной энергии есть

$$\Delta V \sim m^2 M_{\text{GUT}}^2,$$

а кинетической —

$$\Delta E \sim M_{\text{GUT}}^4.$$

Мы приходим к пониманию того факта, что квантовые флуктуации полей в окружающем нас пространстве непрерывно создают области с повышенной плотностью энергии. С точки зрения стороннего наблюдателя, время жизни такой флуктуации мало. В рассмотренном примере это порядка 10^{-40} с. Пространственный размер области, занятой флуктуацией, составляет $\sim 10^{-30}$ см. Будучи очень малым, он, тем не менее много больше планковского размера, что позволяет использовать уравнения Эйнштейна в их стандартном виде для описания процессов внутри этой области. Точка зрения внутреннего наблюдателя рассматривается ниже.

Уравнение скалярного поля следует из (5.35):

$$\partial_\mu(\sqrt{-g}g^{\mu\nu}\partial_\nu\varphi) + \sqrt{-g}V'(\varphi) = 0. \quad (5.38)$$

Метрику предполагаем в виде (5.4). В соответствии с однородностью пространства предполагаем также и однородность распределения скалярного поля φ , т.е. $\varphi = \varphi(t)$. Уравнение (5.38) упрощается:

$$\ddot{\varphi} + 3H\dot{\varphi} + V'(\varphi) = 0, \quad H \equiv \dot{a}/a. \quad (5.39)$$

Еще одно уравнение получается, если учесть, что плотность энергии скалярного поля есть $\rho = \frac{1}{2}\dot{\varphi}^2 + V(\varphi)$. Тогда уравнение (5.8) в виде

$$H^2 = \frac{8\pi G}{3} \left(\frac{1}{2}\dot{\varphi}^2 + V(\varphi) \right) \quad (5.40)$$

есть второе уравнение системы для динамических переменных $\varphi(t)$, $a(t)$. В данном случае можно считать, что слагаемое $\Lambda/3$ в (5.8) либо равно нулю, либо уже включено в определение потенциала V .

Ключевым моментом для инфляционного процесса является “медленное” изменение скалярного поля φ (инфлатона). В этом случае поведение системы (5.39), (5.40) происходит как бы в пространстве де Ситтера, даже если $\Lambda = 0$. Действительно, по аналогии с обычной механикой материальной точки, медленное движение осуществляется, если слагаемое, ответственное за трение, $3H\dot{\varphi}$, велико, т.е.

$$3H|\dot{\varphi}| \gg |\ddot{\varphi}|. \quad (5.41)$$

Это позволяет еще более упростить систему уравнений. Действительно, используя (5.41), запишем уравнение (5.39)

$$3H\dot{\varphi} + V'(\varphi) \simeq 0. \quad (5.42)$$

Следовательно, приходим к неравенству

$$V'(\varphi) \sim 3H\dot{\varphi} \gg \ddot{\varphi}.$$

Домножая первый и последний члены на $\dot{\varphi}$ и интегрируя, получаем искомое неравенство

$$\dot{\varphi}^2 \ll V(\varphi). \quad (5.43)$$

Это неравенство означает, что кинетическая энергия мала по сравнению с потенциальной, а значит, последняя во время инфляции меняется незначительно, $V \simeq \text{const}$ и, кроме того, в силу уравнения (5.40), параметр Хаббла также почти постоянен:

$$H \equiv \frac{\dot{a}}{a} \simeq \sqrt{\frac{8\pi G}{3} V(\varphi)}. \quad (5.44)$$

Решение уравнения (5.44), $a(t) \propto \exp(Ht)$, означает (приблизительно) экспоненциальный рост масштабного фактора, а значит, и физических расстояний, аналогично пространству де Ситтера. Это и не удивительно, поскольку (приблизительно) постоянный потенциал инфлатона φ может быть интерпретирован как космологическая постоянная.

Динамика скалярного поля, связанного с гравитацией, гораздо более интересна по сравнению с динамикой пространства де Ситтера. В частности, неравенство (5.43) обязательно со временем нарушается, поскольку поле φ , хоть и медленно, но движется к минимуму потенциала. Когда потенциал станет достаточно мал, параметр Хаббла H будет также мал. Следовательно, “трением” в уравнении (5.39) можно пренебречь, что означает начало быстрого движения инфлатона к минимуму потенциала и окончание инфляционной стадии. Однако сейчас проследим за процессами на инфляционной стадии. Динамика инфлатона φ определяется уравнением (5.42), а масштабный фактор a — уравнением (5.44). Последнее решается в общем виде:

$$a(t) = a(t_{\text{in}}) \exp \left[\int_{t_{\text{in}}}^t H(\varphi) dt \right]. \quad (5.45)$$

Если инфляция началась в момент $t = 0$, а закончилась к моменту t_e , то начальная пространственная область увеличи-

лась в $a(t_e)$ раз. Обычно используется значение параметра Хаббла $H_e \sim 10^{13}$ GeV, согласующееся с наблюдательными данными. Для подавляющего большинства оценок достаточно считать $H(\varphi) \approx H_e = \text{const}$. В этом случае масштабный фактор на инфляционной стадии выглядит следующим образом:

$$a(t) \simeq a(t_{\text{in}}) \exp[H_{\text{end}}(t_{\text{end}} - t_{\text{in}})]. \quad (5.46)$$

Здесь H_{end} — значение параметра Хаббла на завершающем этапе инфляционного периода, t_{end} — момент завершения инфляции.

Также полезным оказалось понятие числа е-фолдов (e-folds), N , определяемое как

$$N \equiv \ln \left[\frac{a(t)}{a(t_{\text{in}})} \right] \quad (5.47)$$

Другими словами, если известно число N , то это означает, что к данному моменту времени Вселенная расширилась в e^N раз.

Если инфляция началась, когда скалярное поле имело значение φ_{in} , то число N вычисляется следующим образом:

$$N \equiv \ln \left[\frac{a(t)}{a(t_{\text{in}})} \right] = \int_{t_{\text{in}}}^t H dt = \int_{\varphi_{\text{in}}}^{\varphi} H \frac{d\varphi}{\dot{\varphi}} = - \int_{\varphi_{\text{in}}}^{\varphi} \frac{3H(\varphi)^2 d\varphi}{V'(\varphi)}. \quad (5.48)$$

Рассмотрим подробнее условие медленного скатывания. Чтобы определить более явно, при каких условиях это возможно, проделаем следующее. Из уравнения (5.42) в форме

$$\dot{\varphi} = - \frac{V'(\varphi)}{3H(\varphi)}$$

выразим вторую производную:

$$\ddot{\varphi} = \frac{V'(\varphi)}{3H^2(\varphi)} H'(\varphi) \dot{\varphi} - \frac{V''(\varphi)}{3H(\varphi)} \dot{\varphi}.$$

Чтобы условие медленного скатывания выполнялось, оба слагаемых в правой части должны быть малы по сравнению с $3H\dot{\varphi}$.

Вспоминая выражение (5.44) для параметра Хаббла, получаем следующие условия:

$$\varepsilon \equiv \frac{M_{\text{Pl}}^2}{16\pi} \frac{V'(\varphi)^2}{V(\varphi)^2} \ll 1; \quad \eta \equiv \frac{M_{\text{Pl}}^2}{8\pi} \left| \frac{V''(\varphi)}{V(\varphi)} \right| \ll 1. \quad (5.49)$$

Тензор энергии-импульса (ТЭИ) скалярного поля $\varphi(t)$ в метрике (5.4), как это следует из симметрии задачи, совпадает с ТЭИ идеальной жидкости в сопутствующей СО (см. раздел 5.1), $T_{\mu}^{\nu} = \text{diag}(\rho, -p, -p, -p)$, причем для скалярного поля плотность энергии ρ и давление p определяются формулами

$$\rho = \frac{1}{2}\dot{\varphi}^2 + V(\varphi) \quad (5.50)$$

$$p = \frac{1}{2}\dot{\varphi}^2 - V(\varphi). \quad (5.51)$$

В период инфляции $\frac{1}{2}\dot{\varphi}^2 \ll V(\varphi)$, что и означает приближенно деситтеровское пространство с соответствующим поведением масштабного фактора.

5.4. Послеинфляционные стадии

5.4.1. Постинфляционный разогрев Вселенной

Следующая после инфляции стадия эволюции Вселенной, пожалуй, наиболее сложна для анализа. Даже ее название (*reheating*), произошедшее от английского “reheating”, что буквально переводится как “повторный нагрев”, не совсем правильно отражает действительность (так как, скорее всего, не было никакого первичного нагрева). Именно во время этой стадии происходит активное рождение частиц высоких энергий и их термализация, что на языке статистической физики означает нагрев только что рожденной плазмы.

Инфляционный период заканчивается, когда “трение” становится малым, т.е. при $H_{\text{end}} \lesssim m_{\text{infl}}$ (m_{infl} — масса инфлатона).

Инфлатонное поле начинает быстро осциллировать около минимума потенциала согласно уравнению (5.39). Именно это переменное во времени движение порождает частицы. Постепенно энергия осцилляций поля переходит в энергию частиц. Используя формулы для давления $p = \frac{1}{2}\dot{\varphi}^2 - V(\varphi)$ и плотности энергии $\rho = \frac{1}{2}\dot{\varphi}^2 + V(\varphi)$, это уравнение легко преобразовать к знакомому виду

$$\frac{d\rho}{dt} = -3H(t)(p + \rho). \quad (5.52)$$

Дополнительное условие (уравнение состояния) $\rho + p = \gamma\rho$, справедливо лишь приближенно, после усреднения по периоду колебаний. При этом параметр γ зависит от модели:

$$\gamma \simeq \frac{\int_0^T (\rho + p) dt}{\int_0^T \rho dt} = \frac{\int_0^T \dot{\varphi}^2 dt}{\int_0^T [\frac{1}{2}\dot{\varphi}^2 + V(\varphi)] dt} = \frac{\int_0^{\varphi_{\max}} \dot{\varphi} d\varphi}{\int_0^{\varphi_{\max}} d\varphi [\frac{1}{2}\dot{\varphi}^2 + V(\varphi)] / \dot{\varphi}}. \quad (5.53)$$

Дальнейшее упрощение этой формулы возможно, если предположить, что за один период колебаний Вселенная расширится незначительно, что реалистично. Тогда, согласно закону сохранения энергии, имеем $\dot{\varphi} = \sqrt{2[V(\varphi_{\max}) - V(\varphi)]}$. Вблизи своего минимума потенциал аппроксимируется параболой $V(\varphi) \cong \lambda\varphi^2$, уравнение (5.53) сильно упрощается, и величина γ на стадии рехитинга оказывается равной единице,

$$\gamma \cong 1. \quad (5.54)$$

Определим зависимость масштабного фактора от времени. Его начальное значение на стадии рехитинга совпадает со значением в конце инфляции, $a(t_{\text{in}}) = a_{\text{infl}}(t_{\text{end}}) = H(\varphi_U)^{-1} e^{N_U}$. Предполагается, что видимая часть Вселенной образовалась за N_U е-фолдов до окончания инфляции. Начальный размер Вселенной был равен обратному параметру Хаббла $H(\varphi_U)^{-1}$, а значение

инфлатона равнялось φ_U . Плотность энергии в начале стадии рехитинга совпадала с плотностью энергии в конце инфляции, $\rho(t_{\text{in}}) = \rho_{\text{infl}}(t_{\text{end}}) = V(\varphi_{\text{end}})$. Тогда формула (5.23) записывается в виде

$$a_{\text{reh}}(t) = \left(\frac{2}{3}H_{\text{end}}\right)^{2/3} H(\varphi_U)^{-1} e^{N_U} t^{2/3}. \quad (5.55)$$

Напомним, что обычно предполагается $H_{\text{end}} \simeq H(\varphi_U) \simeq 10^{13}$ ГэВ.

Во время быстрых осцилляций инфлатонное поле порождает частицы, передавая им часть своей энергии. Предположим, что плотность энергии взаимодействия инфлатона φ с фермионами ψ и другими скалярными частицами χ имеет обычный вид

$$L_{\text{int}} = h\varphi\bar{\psi}\psi + g\varphi\chi^2,$$

где g, h — константы связи. Тогда скорость изменения плотности энергии инфлатонного поля есть

$$\dot{\rho} = \frac{d}{dt} \left[\frac{1}{2}\dot{\varphi}^2 + V(\varphi) + \Delta L \right] = \dot{\varphi}\ddot{\varphi} + \dot{\varphi}V'(\varphi) + \dot{\Delta}L,$$

где ΔL — слагаемое, ответственное за уменьшение плотности энергии инфлатона при его распаде на другие, вторичные частицы.

Считаем обратную реакцию излучения на динамику инфлатона малой [134]. Тогда в первом приближении, согласно уравнениям движения, получаем

$$\ddot{\varphi} \simeq -3H\dot{\varphi} - V'.$$

Подставляя в предыдущее выражение, находим скорость изменения плотности энергии инфлатона

$$\dot{\rho} \simeq -3H(\rho + p) + \dot{\Delta}L.$$

Теперь смысл последнего слагаемого ясен — это скорость изменения плотности энергии инфлатона за счет распада на другие частицы,

$$\Delta \dot{L} = -\Delta \dot{\rho},$$

где $\Delta \dot{\rho}$ — скорость изменения плотности вторичных частиц за счет распада инфлатона. Раскладывая в ряд Тейлора по производным поля φ , имеем

$$\Delta \dot{\rho}(\dot{\varphi}) = \Delta \dot{\rho}(0) + A\dot{\varphi} + \Gamma\dot{\varphi}^2 + \dots \simeq \Gamma\dot{\varphi}^2.$$

Первый член разложения равен нулю, поскольку покоящееся поле не излучает, второй — потому что при осцилляциях инфлатона среднее от производной равно (приближенно) нулю. Учтем также, что давление и плотность энергии колеблющегося инфлатонного поля связаны обычным соотношением $p = w\rho$. Тогда

$$\dot{\rho} \simeq -(3H + \Gamma)\dot{\varphi}^2 = -(3H + \Gamma)(p + \rho) \simeq -(3H + \Gamma)(w + 1)\rho.$$

Вспомянув, что $H = \dot{a}/a$, получаем зависимость плотности энергии инфлатона от времени

$$\rho = \rho_{end} \left(\frac{a(t)}{a(t_{end})} \right)^{-3(w+1)} \exp \{ -\Gamma(w+1)(t - t_{end}) \}, \quad (5.56)$$

где индекс “end” относится к моменту окончания инфляции. Физически это выражение вполне понятно. Второй множитель означает уменьшение плотности энергии инфлатонного поля благодаря расширению пространства, а третий — за счет его распада на частицы.

Что происходит при этом с плотностью энергии рождающихся частиц материи? Скорость изменения плотности энергии релятивистских частиц ρ_m записывается следующим образом:

$$\begin{aligned} \dot{\rho}_m &= -3H(p_m + \rho_m) + \Gamma\dot{\varphi}^2 = \\ &= -3H(w_m + 1)\rho_m + \Gamma(w + 1)\rho = \\ &= 4H\rho_m + \Gamma(w + 1)\rho. \end{aligned} \quad (5.57)$$

Здесь учтено, что $w_m = 1/3$ для релятивистских частиц. Очевидна конкуренция двух слагаемых — уменьшение плотности за счет расширения пространства (первое слагаемое) и увеличение плотности за счет распада инфлатона.

Релятивистские частицы быстро термализуются. В момент начала осцилляций инфлатона доминирует приток энергии в плазму частиц, температура растет. Со временем интенсивность осцилляций уменьшается и начинает доминировать эффект расширения пространства, приводящий к охлаждению плазмы. Оценим температуру плазмы, при которой ее плотность максимальна. Полагая $\rho_m \sim \rho$, получаем из предыдущего уравнения $H \sim \Gamma$. Поскольку параметр Хаббла и космологическое время связаны соотношениями $t \sim 1/H$ (см. (5.24)), а температура связана со временем как $T \sim \sqrt{M_{\text{Pl}}/t}$ (см. (5.63)), приходим к следующей оценке:

$$T_{\text{reh}} \sim \sqrt{M_{\text{Pl}}/t} \sim \sqrt{M_{\text{Pl}}H} \sim \sqrt{M_{\text{Pl}}\Gamma}.$$

Большая неопределенность в расчете величины Γ не позволяет однозначно определить температуру нагрева среды. Обычно предполагается, что она варьируется в широких пределах от 10^4 до 10^{12} ГэВ.

5.4.2. Стадия доминантности излучения

До сих пор нам удавалось выразить плотность энергии в полевых терминах, что позволило решить уравнение (5.42) и найти зависимость масштабного фактора от времени. Но в конце стадии рехитинга осцилляции инфлатона затухли, породив высокоэнергичные частицы. Наличие плазмы частиц указывает на желательность введения понятия температуры, чтобы использовать стандартные результаты статистической физики и термодинамики. Понятие температуры является строгим, лишь если система находится в равновесии. В нашем же случае это, очевидно, не так из-за расширения Вселенной.

Тем не менее, релаксационные процессы проходят так быстро, что в каждый момент времени состояние Вселенной близко к равновесному. Для проверки предположим, что это так и посмотрим, не приходим ли мы к противоречию. Введем температуру T плазмы в момент времени t . Ниже будет показано, что космологическое время и температура связаны соотношением $t \sim M_{\text{Pl}}/T^2$ (см. формулу (5.63)). Установление равновесия происходит при столкновениях частиц. Характерное время столкновения электрона и фотона $t_{\gamma e}$ есть $t_{\gamma e} \sim 1/(n\sigma v)$, где $n \sim T^3$ — плотность электронов, $\sigma \sim \alpha^2/T^2$ — комптоновское сечение и $v \simeq 1$ — скорость электронов. Понятием температуры можно пользоваться, только если $t_{\gamma e} \ll t$. Подставляя оценки величин, приведенные в этом абзаце, получим условие применимости понятия температуры в расширяющейся Вселенной в виде $T \ll \alpha^2 M_{\text{Pl}} \sim 10^{17}$ ГэВ. Известно, что температура после окончания инфляции вряд ли превышала 10^9 ГэВ. Следовательно, равновесие в среде устанавливается очень быстро, по крайней мере за счет комптоновского рассеяния, что оправдывает введение понятия температуры.

Для нахождения явного вида масштабного фактора $a_{\text{RD}}(t)$ на радиационно-доминированной стадии используем общую формулу (5.22) с $\gamma = 4/3$. Это значение следует из связи давления и плотности энергии $p = \rho/3$ для среды, состоящей из ультрарелятивистских частиц, и формулы (5.20).

Начальными условиями для данной стадии являются: $t_{\text{in}} = t_{\text{reh}}$, $\rho_{\text{in}} = \rho(t_{\text{reh}})$ и $a_{\text{in}} = a(t_{\text{reh}})$, где индекс 'reh' относится к окончанию предыдущей стадии, рехитинга. Окончательное выражение для масштабного фактора приведено в (5.79).

Рассмотрим, как меняется температура вещества со временем. Прежде всего напомним, что плотность энергии ультрарелятивистских частиц связана с температурой T следующим образом:

$$\rho = \frac{\pi^2}{30} g_* T^4, \quad (5.58)$$

где g_* — число сортов частиц с учетом их статистического веса.

Далее установим важную связь между температурой и масштабным фактором. С одной стороны, на данной стадии $\gamma = 4/3$ и, согласно (5.21), имеем $\rho \sim a^{-4}$. С другой стороны, $\rho \sim T^4$. Следовательно,

$$a(t) = \text{const}/T(t). \quad (5.59)$$

Температура уменьшается обратно пропорционально масштабному фактору.

Покажем, что энтропия после стадии рехитинга (когда создавались частицы и, следовательно, росла энтропия) остается постоянной. Действительно, используя известную связь плотности энтропии и температуры

$$s = \frac{2\pi^2}{45} g_* T^3, \quad (5.60)$$

нетрудно получить

$$S \sim sa^3(t) = \text{const}. \quad (5.61)$$

Другая полезная формула, вытекающая из (5.59), также важна для оценок в космологических моделях:

$$T(t) = \frac{a(t_{\text{reh}})}{a(t)} T_{\text{reh}}. \quad (5.62)$$

Наконец, манипулируя формулами (5.58), (5.59) и (5.79), нетрудно получить явную зависимость температуры среды, заполняющей Вселенную, от времени:

$$T = \left(\frac{45}{32\pi^2 g_*} \right)^{1/4} \sqrt{\frac{M_{\text{Pl}}}{t}}. \quad (5.63)$$

Подчеркнем, что эта формула справедлива только на радиационно-доминированной стадии.

5.4.3. Стадия доминантности вещества

С расширением Вселенной температура среды уменьшается, массивные частицы становятся нерелятивистскими, а длина волны фотонов увеличивается. Наступает момент, когда энергия покоя частиц сравнивается с их кинетической энергией. Начиная с этого момента, наступает стадия доминантности вещества. С течением времени давление p становится настолько малым, что его полагают равным нулю, и соответственно параметр $\gamma = 1$ в уравнении (5.20). В полной аналогии с предыдущими стадиями, масштабный фактор $a_{\text{MD}}(t)$ может быть найден из формулы (5.22) с $\gamma = 1$. Начальные условия для данной стадии таковы:

$$t_{\text{in}} = t_{\text{RD}}, \quad \rho_{\text{in}} = \rho(t_{\text{RD}}) = \frac{\pi^2}{30} g_* T_{\text{RD}}^4, \quad a_{\text{in}} = a_{\text{RD}}(t_{\text{RD}}),$$

где t_{RD} соответствует окончанию радиационно-доминированной стадии.

Интересно, что формула (5.59) продолжает быть верной и на этой стадии с одной оговоркой: теперь T означает температуру фотонов. Последние называют реликтовыми.

Перед тем как приступить к доказательству (5.59), обсудим еще один важнейший период в жизни Вселенной — так называемый период рекомбинации. Дело в том, что при высоких температурах электроны и протоны не в состоянии образовывать атомы водорода из-за его малой энергии связи. Вселенная заполнена плазмой заряженных частиц, интенсивно взаимодействующих с фотонами. Это, кстати, означало, что испущенные фотоны совершали нечто вроде броуновского движения. Вселенная была непрозрачна для электромагнитного излучения. Но в некоторый момент, когда температура существенно уменьшилась, атомы водорода стали образовываться. Будучи нейтральными, они слабо взаимодействовали с фотонами, позволяя последним распространяться на большие расстояния. Временной интервал, при котором атомы рекомбинировали, а среда становилась прозрачной, называется периодом рекомбинации. Все данные говорят о

том, что два периода — рекомбинации и равенства энергии материи и излучения — по времени примерно совпадали. В дальнейшем будем пренебрегать их небольшим с космологической точки зрения различием.

Итак, полагаем, что фотоны не взаимодействуют с материей после достижения температуры T_{rec} . До этого момента распределение фотонов по энергиям представляет собой распределение Бозе-Эйнштейна и к моменту рекомбинации имеет вид

$$dN(t_{\text{rec}}) = V_{\text{rec}} \frac{E_{\text{rec}}^2}{\pi^2} \frac{dE_{\text{rec}}}{\exp(E_{\text{rec}}/T_{\text{rec}}) - 1}. \quad (5.64)$$

Здесь dN — число фотонов с энергиями между E_{rec} и $E_{\text{rec}} + dE_{\text{rec}}$ внутри объема V_{rec} . Поскольку после рекомбинации взаимодействие фотонов с окружающей средой считается малым, энергия фотонов уменьшается только за счет общего расширения пространства,

$$E(t) = p(t) = \frac{2\pi}{\lambda(t)} = \frac{2\pi}{\frac{a(t)}{a(t_{\text{rec}})} \lambda(t_{\text{rec}})} = \frac{a(t_{\text{rec}})}{a(t)} E_{\text{rec}}. \quad (5.65)$$

Кроме того, число фотонов есть величина сохраняющаяся, так что

$$dN(t) = dN(t_{\text{rec}}),$$

в то время как объем V , содержащий эти частицы, увеличивается по закону

$$V(t) = \left(\frac{a(t)}{a(t_{\text{rec}})} \right)^3 V_{\text{rec}}.$$

В результате получаем распределение фотонов в произвольный момент времени t

$$dN(t) = V(t) \frac{E(t)^2}{\pi^2} \frac{dE}{\exp \left[E(t) \frac{a(t)}{a(t_{\text{rec}})} / T_{\text{rec}} \right] - 1}. \quad (5.66)$$

Распределение продолжает иметь вид Бозе-Эйнштейна, причем

$$T(t) = \frac{a(t_{\text{rec}})}{a(t)} T_{\text{rec}}. \quad (5.67)$$

Наше утверждение доказано — закон (5.59) действительно имеет место как на радиационно доминированной стадии, так на стадии доминирования вещества. Неизвестная константа в (5.59) может быть определена из нормировки на данный момент времени: $\text{const} = a_0 T_0$ или на момент рекомбинации, как это сделано ранее.

5.4.4. Современная стадия ускоренного расширения (вторичной инфляции)

Наблюдательные данные все более уверенно указывают на то, что плотность темной энергии, субстанции, равномерно распределенном в пространстве, дает вклад в современную полную плотность энергии около 70%. Наиболее естественное предположение состоит в том, что эта величина не зависит от времени и связана с космологической постоянной Λ . Это, в свою очередь, означает, что в данную эпоху Вселенная приближенно описывается моделью де Ситтера, но с плотностью энергии, много меньшей, чем в инфляционный период. Более того, с дальнейшим расширением пространства плотность материи убывает, в то время как энергия, связанная с Λ -членом, остается постоянной. Следовательно, параметр Хаббла также стремится к постоянному значению

$$H_\Lambda \rightarrow \sqrt{\frac{8\pi}{3} \frac{\rho_\Lambda}{M_{\text{Pl}}^2}}, \quad (5.68)$$

а масштабный фактор — к функции

$$a(t) = H_\Lambda^{-1} e^{H_\Lambda(t-t_0)}. \quad (5.69)$$

Параметр Хаббла характеризует скорость расширения, а для описания изменения этой скорости (т.е. ускорения) используется еще один (безразмерный) параметр — параметр замедления $q(t) := -a\ddot{a}/\dot{a}^2$. Параметр q был введен в то время, когда считалось, что Вселенная расширяется с замедлением. Ускоренному расширению соответствуют значения $q < 0$.

Современные наблюдательные оценки этих параметров, хотя и различны у разных авторов, находятся приблизительно в следующих пределах:

$$H_0 \approx 0.71 \pm 0.04 \frac{\text{км}}{\text{с} \cdot \text{Мпк}}, \quad q_0 \approx -1 \pm 0.4, \quad (5.70)$$

где индекс “0” соответствует современной эпохе. В случае (5.69) получаем $q(t) = q_0 = -1$.

Обратная величина параметра Хаббла характеризует размер причинно связанной области. Оценка, основанная на наблюдательных данных, дает

$$H^{-1} \approx 10^{28} \text{ см} \approx 10^{62} M_{\text{Pl}}.$$

Именно таков размер видимой части Вселенной. Если параметр Хаббла действительно постоянен, то информация о более далеких областях Вселенной не будет нам доступна никогда.

При использовании формул типа (5.22) [например, приведенных ниже формул (5.77), (5.78), (5.79), (5.80)] для численных оценок возникает вопрос, связанный с конкретной величиной масштабного фактора a_{in} , начального для каждой стадии. Кроме того, не очевидно значение координатного расстояния r . Поэтому возникает проблема в вычислении, например, масштабного фактора $a(t_0)$ на данный момент времени t_0 , а следовательно, и физического расстояния между двумя объектами, равного

$$R_0 = a(t_0)r. \quad (5.71)$$

Именно эта величина измеряется наблюдателями. Проблему можно обойти, если рассматривать выражение (5.71) как способ избавиться от сопутствующей координаты r , нормируя масштабный фактор на его современное значение $a_0 = a(t_0)$. Тогда

расстояние между двумя точечными объектами в произвольный момент времени выражается следующим образом:

$$R(t) = \frac{a(t)}{a_0} R_0.$$

Например, расстояние на стадии доминантности вещества есть

$$R(t) = \left(\frac{t - t_{\text{RD}}}{t_0 - t_{\text{RD}}} \right)^{2/3} R_0 \simeq \left(\frac{t}{t_0} \right)^{2/3} R_0.$$

Время жизни Вселенной t_0 и ее размер a_0 уже неплохо определены.

Часто вводят также безразмерный масштабный фактор

$$a(t) \equiv \frac{a(t)}{a_0}, \quad 0 < a(t) \leq 1.$$

5.4.5. Будущее Вселенной: ожидается ли “Большой разрыв”?

Описание темной энергии с помощью Λ -члена в уравнениях Эйнштейна достаточно хорошо согласуется с наблюдениями в современную эпоху. Однако те же наблюдения оставляют довольно большой интервал допустимых величин параметра w , если описывать темную энергию как идеальную жидкость с уравнением состояния $p = w\rho$. При таком описании параметр w , вообще говоря, зависит от времени, а при $w = \text{const}$ закон сохранения приводит к формуле (5.21), $\rho \sim a^{-3(w+1)}$.

Легко убедиться, что ускоренное расширение, $q < 0$, требует отрицательного давления, $w < -1/3$. Тогда, при большом масштабном факторе a , член с k в уравнении (5.82) мал по сравнению с остальными слагаемыми, так что *поведение масштабного фактора при больших временах не зависит от пространственной кривизны*, и зависимость $a(t)$ может быть описана следующим образом:

- а) При $-1/3 > w > -1$, когда $\rho > |p|$ (выполняется условие энергодоминантности), происходит т.н. степенная инфляция:

$$a \sim t^{2/[3(w+1)]}, \quad q = -1 + \frac{3}{2}(w+1) > -1; \quad (5.72)$$

- б) при $w = -1$, что соответствует космологической постоянной $\Lambda > 0$, $\rho = \text{const} > 0$, получается экспоненциальная инфляция:

$$a \sim e^{Ht}, \quad H = \text{const}, \quad q = -1; \quad (5.73)$$

- в) при $w < -1$ с так называемой фантомной материей, имеет место гиперинфляция, заканчивающаяся сингулярностью за счет роста масштабного фактора

$$a \sim (t_* - t)^{-2/[3|w+1|]}, \quad q = -1 - \frac{3}{2}|w+1| < -1, \quad (5.74)$$

где t_* — момент сингулярности (Big Rip — Большой разрыв).

Как видно из оценок (5.70), с наблюдательной точки зрения вполне допустимы все три указанных варианта. Однако поскольку ускорение расширения Вселенной не просто имеет место, а возрастает, имеется некоторый перевес в сторону варианта в), а во многих работах обсуждаются космологические модели, в которых темная энергия “преодолеывает фантомный барьер”, т.е. в некоторый момент времени ($3 - 4 \cdot 10^9$ лет назад) переходит от $w > -1$ к $w < -1$.

В случае в) материя ведет себя весьма экзотически: в отличие от обычной материи, ее плотность не убывает, а растет с ростом объема, а в сингулярности обращаются в бесконечность a и ρ одновременно.

Катастрофический рост $a(t)$ приводит к росту всех расстояний — начиная от межгалактических, затем межзвездных, межпланетных — и кончая внутриатомными, т.е. непосредственно перед сингулярностью произойдет распад всего вещества и даже всех составных частиц.

Неизбежна ли такая печальная судьба нашей Вселенной, если подтвердится значение $w < -1$ в современную эпоху? Ответ, к счастью, отрицательный.

В самом деле, если $w = p/\rho$ зависит от времени, можно предположить, что темная энергия представлена скалярным полем ϕ с лагранжианом

$$L_s = \frac{1}{2}\varepsilon g^{\mu\nu} \phi_{,\mu} \phi_{,\nu} - V(\phi), \quad (5.75)$$

где $\varepsilon = \pm 1$, а $V(\phi)$ — потенциал самодействия ϕ . Это те же нормальные и фантомные скалярные поля, с которыми мы сталкивались в главах 3 и 4. При $\phi = \phi(t)$,

$$\rho = \frac{1}{2}\varepsilon\dot{\phi}^2 + V, \quad p = \frac{1}{2}\varepsilon\dot{\phi}^2 - V, \quad w = \frac{p}{\rho} = -1 + \frac{2\varepsilon\dot{\phi}^2}{2V + \varepsilon\dot{\phi}^2}. \quad (5.76)$$

Следовательно, нормальное скалярное поле ($\varepsilon = +1$) с положительным потенциалом V дает $w > -1$, а фантомное скалярное поле ($\varepsilon = -1$) с $V > 0$ приводит к $w < -1$. Однако, если при больших t скалярное поле достаточно быстро стремится к экстремуму потенциала, $V_{\text{ext}} > 0$, то $\dot{\phi} \rightarrow 0$ и $w \rightarrow -1$ при $t \rightarrow \infty$; V_{ext} ведет себя как эффективная космологическая постоянная, и соответственно для эволюции Вселенной получается де-ситтеровская асимптотика (5.73).

Известно, что нормальное скалярное поле стремится к минимуму потенциала (“скатывается” вдоль кривой $V(\phi)$), тогда как фантомное, наоборот, к максимуму (“взбирается” по кривой $V(\phi)$). Если $V(\phi)$ имеет максимум, к нему и будет стремиться фантомное скалярное поле в ходе космологической эволюции.

Для этого случая известен точный результат [145]: *если потенциал $V(\phi)$ ограничен сверху, то его максимум представляет собой глобальный аттрактор для космологических решений уравнений скалярного и гравитационного полей.*

5.5. Масштабный фактор в общем случае

Полезно собрать основные формулы для масштабного фактора на разных стадиях. Поскольку каждая стадия длится много дольше, чем предыдущая, т.е. $t \gg t_{\text{in}}$, используем более компактную, но приближенную формулу (5.23).

Инфляционная стадия:

$$a(t) \equiv a_I(t) = H_U^{-1} \exp(\int H dt) \approx H_e^{-1} e^{N_U t}. \quad (5.77)$$

Численные значения $N_U \approx 60$, $H_e \approx 10^{13}$ ГэВ согласуются с наблюдениями для большинства моделей.

Стадия рехитинга:

$$a(t) \equiv a_{\text{reh}}(t) = a_I(t_e) \left(\frac{2}{3} H_e\right)^{2/3} (t - t_e)^{2/3}. \quad (5.78)$$

Стадия радиационной доминантности:

$$a(t) \equiv a_{\text{RD}}(t) = a_{\text{reh}}(t_{\text{reh}}) \left[\frac{1}{2} H(t_{\text{reh}})\right]^{1/2} (t - t_{\text{reh}})^{1/2}. \quad (5.79)$$

Стадия доминантности материи:

$$a(t) \equiv a_{\text{MD}}(t) = a_{\text{RD}}(t_{\text{RD}}) \left[\frac{2}{3} H(t_{\text{RD}})\right]^{2/3} (t - t_{\text{RD}})^{2/3}. \quad (5.80)$$

Современное ускоренное расширение (вторичная инфляция), в предположении, что темная энергия описывается Λ -членом в уравнениях Эйнштейна:

$$a(t) = H_\Lambda^{-1} e^{H_\Lambda(t-t_0)}, \quad H_\Lambda = 3/\Lambda. \quad (5.81)$$

Все эти формулы носят приближенный характер, так как учитывают лишь один преобладающий вид источника гравитационного поля на каждом этапе.

Реальное содержание Вселенной — это смесь излучения, вещества и темной энергии, которая предположительно описывается космологической постоянной. Получим общее выражение для масштабного фактора, для чего используем уравнение (5.8):

$$\frac{\dot{a}^2}{a^2} + \frac{k}{a^2} = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3}. \quad (5.82)$$

Вклад от плотности темной энергии ρ_Λ выделен и записан в виде Λ -члена, $\Lambda/3 = (8\pi G/3)\rho_\Lambda$. Оставшаяся часть плотности энергии складывается из плотности вещества ρ_m и плотности излучения ρ_r ,

$$\rho = \rho_m + \rho_r. \quad (5.83)$$

Зависимость этих величин от масштабного фактора уже обсуждалась:

$$\rho_m(t) = \rho_m(t_0)/a^3(t), \quad \rho_r(t) = \rho_r(t_0)/a^4(t).$$

Подставляя эти выражения в уравнение (5.82), получаем уравнение для масштабного фактора

$$\frac{\dot{a}^2}{a^2} = \frac{8\pi G}{3} \left[\frac{\rho_m(t_0)}{a^3(t)} + \frac{\rho_r(t_0)}{a^4(t)} + \rho_\Lambda \right] - \frac{k}{a^2}. \quad (5.84)$$

Дополнительное условие выберем в виде $a(t_0) = 1$. Численное решение этого уравнения не представляет особой сложности. Преобразуем его, введя определения, часто встречающиеся на практике.

Прежде всего, определим критическую плотность вещества в современную эпоху

$$\rho_c \equiv \frac{3H_0^2}{8\pi G}. \quad (5.85)$$

где H_0 — современное значение параметра Хаббла. При таком значении плотности Вселенная должна была быть плоской ($k = 0$) при параметре $\Lambda = 0$. Вообще говоря, средняя плотность Вселенной $\rho \neq \rho_c$, поэтому удобно ввести безразмерный параметр плотности

$$\Omega_{\text{tot}} = \rho/\rho_c. \quad (5.86)$$

Введем относительные плотности Ω на данный момент времени. Например, для вещества, излучения и космологической постоянной имеем

$$\Omega_m = \frac{\rho_m(t_0)}{\rho_c}, \quad \Omega_r = \frac{\rho_r(t_0)}{\rho_c}, \quad \Omega_\Lambda = \frac{\rho_\Lambda(t_0)}{\rho_c}.$$

Выражение для Ω_k , параметра плотности, связанного с кривизной пространства, выглядит несколько по-другому:

$$\Omega_k = -\frac{k}{H_0^2}.$$

Теперь уравнение (5.84) выглядит так:

$$\frac{\dot{a}^2(t)}{a^2(t)} \equiv H(t) = H_0^2 [\Omega_k a^{-2}(t) + \Omega_m a^{-3}(t) + \Omega_r a^{-4}(t) + \Omega_\Lambda], \quad (5.87)$$

где $H_0 = H(t_0)$. Выбрав $t = t_0$ в выражении (5.87), получим правило суммы

$$\Omega_r + \Omega_m + \Omega_\Lambda + \Omega_k = 1.$$

Наконец, приведем зависимость от времени размера горизонта (в данном случае это означает расстояние, на которое уходит свет за время t):

$$R_{\text{hor}}(t) = a(t) \int_0^t \frac{dt'}{a(t')}.$$

На инфляционной стадии $a(t) = \text{const} \cdot e^{Ht}$, где H — параметр Хаббла, выраженный через плотность энергии и с хорошей точностью равный постоянной величине. Размер горизонта есть $R_{\text{hor}} = H^{-1}(1 - e^{-Ht})$ и при больших временах стремится к обратной величине параметра Хаббла, т.е. тоже к константе.

На других этапах имеем приближенно $a(t) = \text{const} \cdot t^\beta$, где постоянная $\beta < 1$. Тогда параметр Хаббла имеет вид $H(t) = \beta t^{-1}$, а размер горизонта $R_{\text{hor}} = t/(1 - \beta)$. Следовательно, $H^{-1} = R_{\text{hor}}(1 - \beta)/\beta$, т.е. обратный параметр Хаббла по порядку величины совпадает с размером горизонта.

5.6. Зачем нужен инфляционный период?

В сценариях рождения Вселенной, исследовавшихся до возникновения идеи инфляции, в которых первой стадией считалась стадия “горячей” Вселенной с преобладанием излучения, имеется несколько неразрешимых проблем. Перечислим наиболее существенные.

- 1) Начальная сингулярность. Изучение классических уравнений ОТО применительно к Вселенной как к целому приводит к выводу о том, что ее начальное состояние характеризовалось бесконечной плотностью. Идея инфляции не решает проблему полностью, хотя и указывает путь решения. По-видимому, начало нашей Вселенной положила одна из квантовых флуктуаций, когда классические уравнения, и тем более выводы из них, были неприменимы.
- 2) Наше пространство описывается с высокой точностью геометрией Евклида. Причины этого неочевидны.
- 3) Наша Вселенная вмещает в себя порядка 10^{22} звезд. Как мог возникнуть объект такой массы и объема? Почему он существует так долго?

- 4) Реликтовое излучение с высокой степенью изотропно. Почему так малы колебания его плотности энергии?
- 5) Как появились неоднородности плотности энергии, послужившие основой для формирования галактик?
- 6) Почему в нашей Вселенной преобладает вещество, в то время, как законы природы почти симметричны относительно замены частиц на античастицы?
- 7) На ранней стадии, при больших плотностях энергии, должны образовываться топологически стабильные состояния, такие, например, как магнитный монополю. Их количество и масса достаточно велики, чтобы они не остались незамеченными [315]. Однако эти объекты до сих пор не обнаружены.

Обсудим основные проблемы. Первая из них — проблема плоскостности Вселенной.

5.6.1. Проблема плоскостности

Напомним понятие критической плотности, определяемой как плотность вещества, при которой параметр кривизны $k = 0$ при $\Lambda = 0$. Из уравнения (5.8) непосредственно следует

$$\rho_{\text{crit}}(t) \equiv \frac{3}{8\pi} M_{\text{Pl}}^2 H^2(t) \quad (5.88)$$

(Напомним, что $G = 1/M_{\text{Pl}}^2$ и $\dot{a}/a = H$). Поскольку постоянная Λ мала по крайней мере на радиационно-доминированной стадии, пренебрежем ей в уравнении (5.8), после чего получим

$$\left| \frac{\rho_{\text{crit}}(t) - \rho(t)}{\rho_{\text{crit}}(t)} \right| = \dot{a}^{-2}(t) \quad (5.89)$$

для $k = \pm 1$. Современное значение плотности энергии мало отличается от критического, [281], так что неравенство

$$\left| \frac{\rho_{\text{crit}}(t_0) - \rho(t_0)}{\rho_{\text{crit}}(t_0)} \right| = \dot{a}^{-2}(t_0) < 1 \quad (5.90)$$

надежно установлено. Из уравнений (5.89), (5.90) получаем

$$\left| \frac{\rho_{\text{crit}}(t) - \rho(t)}{\rho_{\text{crit}}(t)} \right| = \left| \frac{\rho_{\text{crit}}(t_0) - \rho(t_0)}{\rho_{\text{crit}}(t_0)} \right| \frac{\dot{a}^2(t_0)}{\dot{a}^2(t)} < \frac{\dot{a}^2(t_0)}{\dot{a}^2(t)}.$$

В отсутствие инфляции имеется лишь две стадии — радиационно-доминированная стадия и стадия доминирования вещества. Зависимость от времени масштабного фактора на обеих стадиях представима в виде

$$a(t) = \text{const} \cdot t^n, \quad n = 1/2 \text{ или } 2/3 < 1. \quad (5.91)$$

Следовательно,

$$\left| \frac{\rho_{\text{crit}}(t) - \rho(t)}{\rho_{\text{crit}}(t)} \right| < \frac{\dot{a}^2(t_0)}{\dot{a}^2(t)} \sim \left(\frac{t}{t_0} \right)^{2(1-n)} \ll 1. \quad (5.92)$$

Тогда, если зависимость (5.91) верна, то в момент рождения Вселенной, т.е. при $t \ll t_0$, плотность вещества должна была с необъяснимой точностью равняться критической. Единственным выходом из положения является нарушение закона (5.91), по крайней мере, на начальной стадии. Именно это и обеспечивает процесс инфляции.

Пусть Вселенная проходит инфляционную стадию с масштабным фактором (5.45) $a(t) = H^{-1} \text{ch}(Ht)$. Параметр Хаббла в время инфляции практически постоянен, $H \simeq \text{const}$. Тогда зависимость отношения (5.89) от времени есть

$$\left| \frac{\rho_{\text{crit}}(t) - \rho(t)}{\rho_{\text{crit}}(t)} \right| = \text{sh}^{-2}(Ht).$$

Эта функция экспоненциально убывает на инфляционной стадии, что и объясняет ее малость в постинфляционный период, см. (5.92).

5.6.2. Начальный размер Вселенной

Если сейчас Вселенная имеет размер $L_U \sim 10^{28}$ см, то в момент t ее размер составлял

$$L(t) = \frac{a(t)}{a(t_0)} L_U = \frac{T_0}{T} L_U. \quad (5.93)$$

Оценим наименьший размер Вселенной. Из формулы (5.93) следует, что это имело место при наибольшей температуре. Классическое описание справедливо при энергиях, а следовательно, и температурах, меньших планковской, $T \leq M_{\text{Pl}} \sim 10^{32}$ К. Подставляя в формулу (5.93) это значение, а также современную температуру реликтового излучения $T_0 \sim 2,7$ К, получаем минимальный размер Вселенной в момент рождения

$$L_{\text{min}} \sim 10^{-4} \text{ см.}$$

Остается загадкой, почему Вселенная родилась, имея такой большой, по сравнению с планковским, размер.

Механизм инфляции как раз и позволяет Вселенной быстро расширяться до указанного размера.

Кроме того, если уж Вселенная родилась, имея размер $\sim 10^{-4}$ см на планковском масштабе, то в этот момент размер причинно связанной области был порядка 10^{-33} см. Это значит, что Вселенная состояла из $\sim 10^{87}$ причинно не связанных областей. Совершенно неясно, почему плотность энергии во всех областях одинакова с точностью порядка 10^{-5} . В этом суть проблемы горизонта.

Предположение о существовании инфляционного периода помогает решить и другие проблемы, такие, например, как проблема монополей [315].

5.6.3. Причинные связи во время инфляции и после нее

Инфляционная стадия обладает интереснейшей особенностью, которая выделяет ее из других стадий. Поскольку понимание

этой особенности необходимо для понимания большинства последующих процессов, остановимся на ней подробнее.

Рассмотрим две пространственные точки в покое в космологической СО. Пусть в начальный момент t_{in} расстояние между ними равнялось ($l(t_{\text{in}})$) и обе находились в одной причинно связанной области — “под горизонтом”. Поскольку размер причинно связанной области определяется величиной, обратной параметру Хаббла H , то в первый момент выполняется неравенство

$$l(t_{\text{in}}) < H^{-1}(t_{\text{in}}).$$

В дальнейшем расстояние между частицами растет вместе с масштабным фактором $a(t)$

$$l(t) = a(t)r, \quad (5.94)$$

где координатное расстояние r со временем не меняется, а параметр Хаббла есть по определению $H(t) \equiv \dot{a}(t)/a(t)$. Используя эти формулы, нетрудно найти зависимость от времени соотношения $l(t)/H(t)^{-1}$. Именно это отношение расстояния между частицами к размеру горизонта в зависимости от времени и есть предмет нашего интереса:

$$\frac{l(t)}{H^{-1}(t)} = \dot{a}(t)r. \quad (5.95)$$

В инфляционный период масштабный фактор есть $a(t) \propto e^{H_0 t}$, а параметр Хаббла приблизительно постоянен, $H(t) \simeq H_0 = \text{const}$. Соотношение (5.95) принимает вид

$$\left(\frac{l(t)}{H(t)^{-1}} \right)_{\text{infl}} \propto e^{H_0 t},$$

так что расстояние между двумя частицами растет экспоненциально быстро по сравнению с размером горизонта. Даже если эти частицы родились в причинно связанной области, найдется такой момент времени t_1 , когда они окажутся в причинно не связанных областях, для которых $l(t_1) > H^{-1}(t_1)$. Считается, что в этот момент расстояние между частицами “пересекло горизонт”.

Итак, родившиеся частицы оказываются далеко друг от друга, в причинно не связанных областях, именно за счет свойств пространства во время инфляции.

Что же происходит по окончании инфляции? Масштабный фактор ведет себя как

$$a(t) \propto t^\beta,$$

где параметр β находится в интервале $0 < \beta < 1$ для любой постинфляционной стадии [см. (5.79), (5.80)]. Вычислим интересное нас соотношение согласно формуле (5.95):

$$\left(\frac{l(t)}{H(t)^{-1}} \right)_{\text{power}} \propto t^{\beta-1}, \quad 0 < \beta < 1.$$

Очевидно, что размер горизонта $H(t)^{-1}$ растет со временем быстрее, чем расстояние между двумя частицами. В некоторый момент времени t_2 размер горизонта сравнивается с расстоянием между этими частицами, а затем и превосходит их. Частицы снова оказываются в причинно связанной области. Все сказанное относится, конечно, не только к расстоянию между частицами, но и к любым другим явлениям, слабо влияющим на метрику, например флуктуациям полей.

5.7. Основные свойства расширяющегося пространства

5.7.1. Красное смещение

Здесь и далее масштабный фактор нормирован так, что $a(t_0) = 1$.

Пусть в момент t источником испущен световой сигнал с длиной волны $\lambda(t)$, который принимается на Земле с длиной волны λ . В момент испускания масштабный фактор равен $a(t)$, а в момент поглощения, естественно, единице.

Красное смещение определяется как

$$z = \frac{\Delta\lambda}{\lambda} = \frac{\lambda - \lambda(t)}{\lambda}. \quad (5.96)$$

Очевидно,

$$z = \frac{1 - a(t)}{a(t)} \Rightarrow a(t) = \frac{1}{1 + z}. \quad (5.97)$$

Будем считать Вселенную плоской, т.е. $k = 0$. Координатное $r(t)$ и физическое $R(t)$ расстояния до источника есть

$$r(t) = \int_t^{t_0} \frac{dt'}{a(t')}, \quad (5.98)$$

$$R(t) = a(t)r(t) = \frac{1}{1 + z}r(t). \quad (5.99)$$

Нас, конечно, будет интересовать физическое расстояние (5.99). Для его определения потребуется выражение для $a(t)$ или связанное с ним $H(z)$. Подробнее,

$$R(t) = a(t) \int_t^{t_0} \frac{dt'}{a(t')} = \frac{1}{1 + z(t)} \int_t^{t_0} [1 + z(t')] dt'.$$

Замена переменных $t' \rightarrow z$ приводит к выражению

$$dz = -\frac{\dot{a}}{a(t')^2} dt' = -\frac{H(t')}{a(t')} dt' = -H(z)[1 + z(t')] dt'.$$

Зависимость $H(z)$ известна: согласно (5.87)

$$H(z) = H_0 \sqrt{\Omega_r(1 + z)^4 + \Omega_m(1 + z)^3 + \Omega_\Lambda}.$$

Величина $\Omega_k = 0$, так как мы рассматриваем плоскую Вселенную с $k = 0$. Окончательно, расстояние до объекта, излучившего свет, связано с красным смещением следующим образом:

$$R(t) = \frac{1}{1 + z} \int_0^z \frac{dz'}{H(z')}.$$

Ну, а красное смещение z определяется из наблюдений по сдвигу частот испущенного и поглощенного света по формуле (5.96).

Таким образом, в расширяющейся Вселенной удастся узнать расстояние до излучающего объекта. Можно также определить и его скорость удаления от наблюдателя за счет космологического расширения:

$$v(t) = \dot{R}(t) = \dot{a}(t)r(t) = H(z)R(t) = \frac{H(z)}{1+z} \int_0^z \frac{dz}{H(z)}.$$

Момент излучения t также однозначно связан с красным смещением z . Эту связь легко установить для каждого из основных периодов, используя выражения (5.78), (5.79), (5.80) и соотношение (5.97).

Задача. Наблюдателями зарегистрировано излучение квазара с $z = 6.41$. Определить, когда и на каком расстоянии до Земли был излучен свет от источника.

Заметим, что кроме космологического красного смещения, как известно, существует и другое — гравитационное — красное смещение, являющееся одним из классических наблюдаемых эффектов релятивистской гравитации и связанное с замедлением хода времени вблизи тяготеющих масс. К космологии оно прямого отношения не имеет, но должно учитываться при интерпретации астрономических наблюдений.

5.7.2. Фотометрическое расстояние

В наблюдениях важную роль играет понятие фотометрического расстояния d_L . Пусть объект излучает в единицу времени энергию E в виде фотонов или гравитационных волн. Наблюдатель на Земле регистрирует поток энергии F . Поскольку в расширяющейся Вселенной длина волны меняется во время движения от источника к наблюдателю, кажущаяся светимость объекта есть $L = E/a(t)$. Это просто эффект уменьшения энергии фотонов из-за увеличения их длины волны. Фотометрическое расстояние

естественно определить через соотношение

$$F = \frac{L}{4\pi d_L^2}.$$

Тогда связь энергии, излученной источником и поглощенной приемником на Земле в момент t_0 , есть

$$F dt_0 \cdot 4\pi r^2 = E dt = L a(t) dt. \quad (5.100)$$

Докажем, что имеет место связь

$$\frac{dt}{dt_0} = a(t). \quad (5.101)$$

Действительно, для светового сигнала, испущенного в момент t и принятого в настоящий момент времени t_0 наблюдателем на координатном расстоянии r , имеем

$$\int_t^{t_0} \frac{dt'}{a(t')} = r.$$

То же самое через интервал времени dt :

$$\int_{t+dt}^{t_0+dt_0} \frac{dt'}{a(t')} = r.$$

Сравнивая два выражения для r , получаем (5.101) и, следовательно,

$$F = \frac{La^2(t)}{4\pi r^2}.$$

Возвращаясь к (5.100), (5.101), получаем

$$d_L = \sqrt{\frac{L}{4\pi F}} = \sqrt{\frac{L}{4\pi F}} = \sqrt{\frac{L}{4\pi La^2(t)/(4\pi r^2)}} = \frac{r}{a(t)} = \frac{R_{\text{phys}}}{a^2(t)},$$

и окончательно находим связь фотометрического расстояния и красного смещения в виде

$$d_L = \frac{R_{\text{phys}}}{(1+z)^2}. \quad (5.102)$$

5.7.3. Скорость частиц в пространстве ФРУ

Свойства материи — ее плотность и уравнение состояния — влияют на глобальные свойства самого пространства и, в частности, на динамику его расширения, т.е. на зависимость масштабного фактора от времени. Но обратное тоже верно: оказывается, расширение пространства приводит к уменьшению скорости частиц. Докажем это для массивной нерелятивистской частицы. Пусть имеется два наблюдателя, находящиеся на координатном расстоянии dr друг от друга. Считаем их неподвижными в сопутствующей системе отсчета, т.е. $dr = \text{const}$. Однако физически, согласно закону Хаббла, второй наблюдатель удаляется от первого со скоростью

$$dv_{\text{obs}} = H(t)dR$$

где dR — физическое расстояние между. Частица, имеющая скорость $V(t)$ относительно первого наблюдателя и пролетающая мимо него в момент времени t , долетит от него до второго наблюдателя за время

$$dt = dR/V(t).$$

Если бы пространство не расширялось, то скорость частицы $V(t)$ была бы постоянной. В расширяющейся Вселенной это не так. Через время dt скорость частицы измеряет второй наблюдатель, который движется относительно первого со скоростью dv_{obs} . Он определит, что скорость пролетающей мимо него частицы равна

$$V(t + dt) = V(t) - dv_{\text{obs}} = V(t) - H(t)dR.$$

Из этого равенства очевидно, что

$$\frac{dV}{dt} = -H(t)\frac{dR}{dt} = -H(t)V = -\frac{V}{a}\frac{da}{dt}.$$

Последнее равенство — просто учет определения параметра Хаббла. Окончательно, имеем следующее тривиальное уравнение:

$$dV/V = -da/a,$$

решение которого имеет вид

$$V(t) = \text{const}/a(t). \quad (5.103)$$

Константа определяется начальными условиями, как обычно.

Итак, в расширяющейся Вселенной скорость свободной частицы уменьшается со временем. Законы Ньютона справедливы только в пространстве Минковского, в котором $a(t) = 1$. Кроме того, в отличие как от ньютоновской механики, так и от СТО, в природе существует выделенная система отсчета — та, относительно которой реликтовое излучение изотропно.

Глава 6

Инфляционные модели

Инфляционная парадигма развивается уже более 20 лет, начиная с работ [158, 282, 283]. С ее помощью успешно решаются основные проблемы космологии нашей Вселенной, начиная с самой ранней стадии ее образования и кончая стадией образования Галактик. Существование инфляционного периода в истории развития нашей Вселенной представляется неизбежным, поскольку позволяет объяснить множество наблюдательных фактов [35, 210]. Первые механизмы инфляции [158, 283] основывались на самосогласованных уравнениях скалярного и гравитационного полей. Они наиболее экономным образом решали ключевые проблемы теории “большого взрыва” — проблему горизонта, проблему плоскостности и т.д. Ключевым моментом являлось то, что взаимодействующие гравитационное и скалярное поля при некоторых условиях приводили к экспоненциальному росту первоначально малых областей пространства. Скалярное — “инфлатонное” — поле в конечном итоге распадалось на фермионы, что являлось причиной нагрева Вселенной до температур $\sim 10^9$ ГэВ [213].

Тем не менее, простейшие инфляционные модели не могли убедительно объяснить весь набор наблюдательных данных. Например, предсказания этих моделей относительно флуктуаций температуры космического фонового излучения не противоречат действительности лишь при достаточно сложных формах

потенциала инфлатонного поля. В основном это чрезвычайная малость параметров потенциала инфлатона, скалярного поля, ответственного за саму возможность осуществления инфляционного процесса. С другой стороны, его взаимодействие с полями материи должно быть достаточным, чтобы в постинфляционный период эффективно породить наблюдаемое число барионов и лептонов.

Дальнейшее развитие теории привело к появлению большого числа инфляционных моделей, в которых вводятся дополнительные поля, такие, например, как гибридная инфляция [212], инфляция на псевдо-намбу-голдстоуновском поле [132]. Большинство этих моделей основывается на взаимодействии классических составляющих различных полей. В то же время в природе реально существует большое число сортов полей, взаимодействие с которыми должно приводить к новым явлениям во время инфляции. Так, например, эффекты диссипации приводят к уменьшению вероятности фазовых переходов. Некоторые инфляционные модели в качестве одного из основных элементов действительно включают взаимодействие классического инфлатонного поля с создаваемыми им частицами различных сортов. На этом эффекте базируется “теплая” модель (warm inflation); обратное влияние рождающихся частиц на динамику инфлатонного поля рассматривается в работе [134].

6.1. Квадратичная инфляция

Рассмотрим инфляцию с квадратичным потенциалом как наиболее перспективную. Известно, что инфлатонное поле, как и множество других скалярных полей, появляется в многомерных теориях естественным образом. Форма потенциала может иметь сложный вид и сильно зависит от выбранной теории. Нас же будет интересовать область потенциала вблизи минимума, поскольку наша современная Вселенная сформировалась на последней стадии инфляции, в процессе ее завершения. Но вбли-

зи минимума любой разумный потенциал может быть разложен в ряд Тейлора и аппроксимирован квадратичным слагаемым. Итак, выберем простейший, квадратичный потенциал в качестве примера:

$$V(\varphi) = \frac{1}{2}m^2\varphi^2. \quad (6.1)$$

Теперь общие уравнения раздела 5.3 могут быть конкретизированы. В частности, система уравнений (5.39), (5.40) решается аналитически. Решение имеет вид

$$\varphi(t) = \varphi_{\text{in}} - \frac{mM_{\text{Pl}}}{\sqrt{12\pi}}t \quad (6.2)$$

Напомним, что это решение справедливо, пока выполняется неравенство (5.41). Для квадратичного потенциала оно сводится к неравенству

$$\varphi \gg \frac{M_{\text{Pl}}}{2\sqrt{3\pi}}. \quad (6.3)$$

Полезно оценить значение поля, при котором условия медленного скатывания нарушаются, т.е. когда $\epsilon \simeq 1$, $\eta \simeq 1$, и, следовательно, инфляция заканчивается. Нетрудно видеть, что для рассматриваемого квадратичного потенциала это происходит при

$$\varphi_{\text{end}} \approx \frac{1}{2\sqrt{\pi}}M_{\text{Pl}}. \quad (6.4)$$

Таким образом, поле φ (инфлатон) должно находиться на планковском масштабе, чтобы возникли условия для инфляционной стадии.

Оценим, во сколько раз вырос пространственный размер флуктуации поля. Для этого надо найти число e -фолдов N , поскольку размер Вселенной возрастает в e^N раз. Число же e -фолдов легко получается из формулы (5.48) в нашем частном случае квадратичного потенциала:

$$N = \frac{2\pi}{M_{\text{Pl}}^2}(\varphi_{\text{in}}^2 - \varphi_{\text{end}}^2). \quad (6.5)$$

Из наблюдений известно, что $N \approx 60$. Теперь мы можем оценить значение инфлатона в момент рождения современного горизонта, т.е. в момент образования той пространственной области, которую мы называем наблюдаемой Вселенной:

$$\varphi_{\text{in}} \approx M_{\text{Pl}} \sqrt{\frac{N}{2\pi}} \sim 3M_{\text{Pl}}. \quad (6.6)$$

Здесь учтено, что $\varphi_{\text{end}} \ll \varphi_{\text{in}}$ и что из наблюдений известно число е-фолдов $N \approx 60$. Теперь нетрудно оценить значение потенциала (6.1), имевшее место при рождении современного горизонта:

$$V(\varphi_{\text{in}}) = \frac{1}{2}m^2\varphi_{\text{in}}^2 \approx 4.5m^2M_{\text{Pl}}^2 \approx 4.5 \cdot 10^{-12}M_{\text{Pl}}^4.$$

Последнее равенство получено для обычно принимаемого значения массы инфлатона $m \approx 10^{-6}M_{\text{Pl}}$. Таким образом современный горизонт нашей Вселенной появился при энергиях, много меньше планковских.

Посмотрим каков размер горизонта, появившегося на планковском масштабе энергий, когда $V \sim M_{\text{Pl}}^4$. При этом начальное значение поля есть $\varphi_{\text{Planck}} \sim M_{\text{Pl}}^2/m \sim 10^6M_{\text{Pl}}$. Подставляя это значение в выражение (6.5), получим

$$N \sim 2\pi \cdot 10^{12}.$$

Следовательно, линейные размеры пространственной области, возникшей при планковском масштабе энергии, увеличились в $\exp(2\pi \cdot 10^{12})$ раз. Ее начальный размер был порядка планковского. Сразу после окончания инфляции пространственный размер такой области равнялся

$$L \sim \exp(2\pi \cdot 10^{12}) \cdot 10^{-33} \text{ см} \sim 10^{2 \cdot 10^{12}} \text{ см}.$$

6.2. Гибридная инфляция

Основным недостатком простейшей модели с квадратичным потенциалом является малость массы инфлатона по сравнению с

массой Планка, $m/M_{\text{Pl}} \simeq 10^{-6}$. Существуют возможности исправить ситуацию. Например, введение второго гипотетического скалярного поля. Одной из самых известных попыток такого рода является гибридная инфляция [212, 218].

Потенциал гибридной инфляции зависит от двух скалярных полей σ и χ и обычно записывается в виде

$$V(\chi, \sigma) = \varkappa^2 \left(M^2 - \frac{\chi^2}{4} \right)^2 + \frac{\lambda^2}{4} \chi^2 \sigma^2 + \frac{1}{2} m^2 \sigma^2, \quad (6.7)$$

где \varkappa, λ и m — параметры теории.

Подобная форма позволяет получить естественным образом медленное скатывание вдоль долины $\chi = 0$, $\sigma > \sigma_c$ и быстрые флуктуации обоих полей при окончании инфляции. Медленное скатывание необходимо для осуществления инфляционной стадии, а быстрые колебания полей — для эффективного производства частиц материи и нагрева Вселенной.

Динамика полей определяется классическими уравнениями

$$\begin{aligned} \ddot{\sigma} + 3H\dot{\sigma} + \frac{1}{2}\lambda^2\chi^2\sigma + m^2\sigma &= 0, \\ \ddot{\chi} + 3H\dot{\chi} - \varkappa^2\chi \left(M^2 - \frac{\chi^2}{4} \right) + \frac{1}{2}\lambda^2\chi\sigma^2 &= 0, \end{aligned} \quad (6.8)$$

где H — параметр Хаббла в этот период.

В контексте гибридной инфляции медленное скатывание означает медленное изменение поля σ вдоль долины $\chi = 0$, которое может иметь место при условии $\sigma > \sigma_c$. Для оценки можно считать, что медленное скатывание происходит при $m \ll H$. В этом случае можно пренебречь второй производной в классических уравнениях (6.8), что позволяет получить решение в виде

$$\sigma(t) = \sigma_{\text{in}} \exp\left(-\frac{m^2}{3H}t\right), \quad m \ll H.$$

Инфляция длится, пока поля не достигнут критической точки

$$\sigma_c = \sqrt{2} \frac{\varkappa}{\lambda} M$$

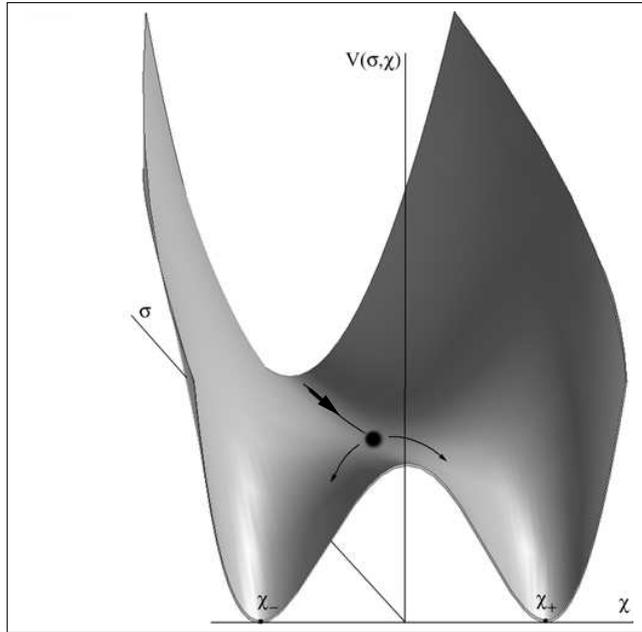


Рис. 6.1. Потенциал гибридной инфляции. Классическое движение полей осуществляется вдоль критической линии (толстая стрелка). После достижения критической точки возможны два направления движения — к минимумам χ_- и χ_+ .

(черная точка на рисунке рис. 6.2). После того как поле σ проходит это значение, движение вдоль линии $\chi = 0$, $\sigma < \sigma_c$ становится нестабильным, и поле χ быстро достигает одного из минимумов $\chi_{\pm} = \pm 2M$, $\sigma = 0$,

Судьба Вселенной зависит от начальных данных, χ_{in} , σ_{in} . Предположим, как обычно, что инфляционная стадия длится в течение времени, равного $N_U H^{-1}$ ($\hbar = c = 1$). Как обычно, выбираем $N_U \approx 60$. Тогда из условия $\sigma(t = N_U H^{-1}) = \sigma_c$ нетрудно

получить начальное значение поля σ :

$$\sigma_{\text{in}} = \sigma_c \exp\left(\frac{m^2}{3H^2} N_U\right).$$

Инфляция заканчивается интенсивными флуктуациями поля в области случайно выбранного минимума.

Тем не менее, и эта хорошо изученная картина имеет серьезный недостаток. Во время инфляционной стадии, когда поля σ, χ движутся вдоль линии $\chi = 0$, пространство разделяется на множество причинно не связанных областей. Благодаря квантовым флуктуациям скалярные поля в этих областях несколько отличаются друг от друга. Очевидно, подавляющее большинство областей содержат поле $\chi \neq 0$. Следовательно, в конце инфляции, по достижении полем σ критического значения σ_c , образуется колоссальное число (около 10^{78}) доменов. В половине из них, тех, которые содержат поле $\chi < 0$, начнется скатывание к минимуму $\chi_- = -2M$, в то время как в другой половине — к минимуму $\chi_+ = +2M$. После инфляции возникает вселенная, состоящая из хаотично распределенных доменов со значениями поля χ_+ или χ_- внутри. Соседние домены разделены полевыми стенками, поскольку при движении от χ_+ к χ_- необходимо побывать в точке ($\chi = 0, \sigma = 0$), доставляющей экстремум потенциала. Подобный период “стеночной доминантности” в эволюции Вселенной неприемлем, поскольку исключает существование Вселенной в ее нынешней форме. Соответственно и обычно предполагаемое движение “в среднем” вдоль линии $\chi = 0$ исключено. Это очень сильное ограничение, налагаемое на модель гибридной инфляции, может быть обойдено как в рамках самой модели, так и при ее модификации. Рассмотрим эти возможности [265].

Оказывается, в первоначальном виде гибридная инфляция может быть реализована лишь при одном условии: начальное значение поля χ во Вселенной, появившейся на инфляционной стадии, отвечало условию $\chi_{\text{in}} \neq 0$. Более того, в течение инфляционного периода поле χ не должно пересекать критическую

линию $\chi = 0$, чтобы избежать описанных выше проблем. Одним из условий медленного движения поля χ является

$$\eta \equiv \frac{V''_{\chi\chi}}{3H^2} \ll 1.$$

Нетрудно оценить эту величину для $\sigma \approx \sigma_{\text{in}}$, $\chi = 0$ и $H \approx \sqrt{8\pi/3}\varkappa M^2/M_{\text{Pl}}$, где M_{Pl} — масса Планка. В результате имеем

$$\begin{aligned} \eta &\equiv \frac{V''_{\chi\chi}}{3H^2} \simeq \frac{\frac{1}{2}\lambda^2\sigma_{\text{in}}^2 - \varkappa^2 M^2}{3H^2} \approx \\ &\approx \frac{\varkappa^2 M^2 \left[\exp\left(\frac{2m^2}{3H^2} N_U\right) - 1 \right]}{3H^2} \approx \frac{2N_U}{(8\pi)^2} \frac{M_{\text{Pl}}^4 m^2}{\varkappa^2 M^6}, \end{aligned}$$

что приводит к неравенству

$$\eta = \frac{6N_U}{(8\pi)^2} \frac{M_{\text{Pl}}^4 m^2}{\varkappa^2 M^6} \ll 1. \quad (6.9)$$

Объединяя его с выражением для температурных флуктуаций

$$\left(\frac{16\pi}{45}\right)^{1/2} \frac{\lambda \varkappa^2 M^5}{M_{\text{Pl}}^3 m^2} \sim \frac{\delta T}{T} \approx 10^{-5}, \quad (6.10)$$

получаем оценку для параметра λ :

$$\lambda \approx \frac{(8\pi)^2}{6N_U} \sqrt{\frac{45}{16\pi}} \frac{\delta T}{T} \eta \frac{M}{M_{\text{Pl}}}. \quad (6.11)$$

Очевидно, этот параметр мал при условии $M \gg M_{\text{Pl}}$, обычном для гибридной инфляции.

Заметим, что модель гибридной инфляции [212] была разработана как раз для того, чтобы избежать неестественно малых параметров. Однако уравнение (6.11) однозначно указывает, что $\lambda \ll 1$ при разумных значениях параметров, и, следовательно, проблема малого параметра остается нерешенной.

Если в инфляционный период среднее значение поля окажется слишком близко к критической линии $\chi = 0$, флуктуации поля внутри некоторых пространственных областей — доменов —

могут пересечь эту линию. В будущем эти домены будут заполнены полем $\chi = \chi_-$ (например) и погружены в пространство с другим значением поля $\chi = \chi_+$. Два минимума разделены замкнутой полевой стенкой, как описано выше. Число таких стенок сильно зависит от начальных условий в момент рождения Вселенной.

6.2.1. Образование черных дыр в модели гибридной инфляции

Хорошо известно, что флуктуации плотности энергии в ранней Вселенной приводили к формированию первичных черных дыр (ПЧД) [27, 236]. Последние не очень массивны, с массами примерно в интервале $M_{\text{BH}} \sim 10^{-5} - 10^{20}$ г в зависимости от конкретной модели. Существуют модели, которые приводят к образованию ПЧД с гораздо большими массами в интервале $10^{20} - 10^{40}$ г [133, 268, 313].

Предыдущее рассмотрение указывает на возможность образования массивных ПЧД также и в рамках гибридной инфляции. Прежде всего, оценим размер и энергию замкнутых стенок, обсуждавшихся выше. Пусть поле в объеме с характерным размером, определяемым параметром Хаббла (H^{-1}), пересекает критическую линию в момент, характеризуемый e -фолдом номер N до окончания инфляции. Это означает, что характерный размер стенки возрастет в e^N раз до окончания инфляции. В конечном итоге, сразу после окончания инфляции, объем, занятый флуктуацией, будет содержать поле в минимуме $\chi = \chi_+$, $\sigma = 0$, в то время, как $\chi = \chi_-$, $\sigma = 0$ вне его. Нетрудно сообразить, что этот объем будет окружен замкнутой стенкой, характеризующейся значением поля $\chi = 0$, $\sigma = 0$. Если размер области велик по сравнению с толщиной стенки, то локально ее можно приближенно считать плоской, разделяющей два минимума потенциала (6.7), $V(\chi, 0)$. Плоская стенка хорошо изучена и представляет

собой солитон [251]. Ее поверхностная плотность энергии есть

$$\mu = \frac{8\sqrt{2}}{3} \varkappa M^3. \quad (6.12)$$

В нашем случае толщина стенки выражается через микроскопические параметры и поэтому много меньше ее размера. Следовательно, использование формулы (6.12) является хорошим приближением. Таким образом, энергия E_{wall} , сосредоточенная в стенке сразу по окончании инфляции, есть

$$E_{\text{wall}} \approx 4\pi\mu (H^{-1}e^N)^2 = 4\sqrt{2} \frac{M_{\text{Pl}}^2}{\varkappa M} e^{2N}. \quad (6.13)$$

Численное значение N находится в интервале ($0 < N < N_U \approx 60$). Гравитационный радиус стенки легко выражается через уже введенные параметры:

$$r_g = \frac{2E_{\text{wall}}}{M_{\text{Pl}}^2} \approx \frac{8\sqrt{2}}{\varkappa M} e^{2N}.$$

Как правило, он много больше толщины стенки $d = 2\sqrt{2}/(\varkappa M)$ для любого e -фолда N . Это означает, что сжатие стенки из-за внутренних натяжений закончится образованием черной дыры с массой $M_{\text{BH}} \approx E_{\text{wall}}$ [268].

Оценим массу подобной ЧД для значений параметров $\varkappa = 10^{-2}$ и $M = 10^{16}$ ГэВ, обычно предполагаемых в гибридной инфляции. Так, если пересечение критической линии произошло при $N = 40$, то масса ЧД есть

$$M_{\text{BH}} \approx 3 \cdot 10^{59} \text{ ГэВ} \sim 100 M_{\odot},$$

где M_{\odot} — масса Солнца. Аналогичная оценка для ЧД наименьшей массы получается, если выбрать момент окончания инфляции, которому соответствует e -фолд $N = 1$:

$$M_{\text{BH, small}} \approx 10^6 M_{\text{Pl}}.$$

Таким образом, гибридная инфляция порождает ЧД в широком интервале масс $10^{25} \div 10^{59}$ ГэВ. Число ЧД зависит от

того, насколько близко среднее значение поля приближается к критической линии при своем классическом движении во время инфляции. А это, в свою очередь, зависит от начальных условий, при которых возникает Вселенная, и параметров гибридной инфляции.

В результате квантовых флуктуаций поле χ оказывается распределенным вокруг среднего значения с дисперсией $\langle \delta\chi \rangle \approx (H/2\pi)\sqrt{N_U}$. Если поле χ приблизится к этому значению во время инфляционной стадии, перепроизводство ЧД неизбежно. Чтобы избежать подобной ситуации, начальное значение этого поля должно удовлетворять неравенству

$$\chi_{\text{in}} \geq \frac{H}{2\pi} \sqrt{N_U} = \sqrt{\frac{2N_U}{3\pi}} \frac{\kappa M^2}{M_{\text{Pl}}}. \quad (6.14)$$

Вселенные, в которых не удовлетворяется это условие, будут содержать слишком много ЧД по сравнению с нашей.

6.2.2. Гибридная инфляция с наклоном

Рассмотрим модификацию гибридной инфляции. А именно, предположим, что потенциал (6.7) имеет добавочное слагаемое вида $\delta V(\sigma, \chi) = \alpha\chi$.

Пусть наша Вселенная родилась, имея размер H^{-1} с некоторым значением поля слева от критической линии, см. рис. 6.2. Тогда по окончании инфляции основная часть Вселенной окажется в минимуме χ_- , в то время как отдельные пространственные области окажутся в минимуме χ_+ . Эти области будут окружены замкнутыми полевыми стенками, как обсуждалось выше. Новым аспектом является то, что плотность энергии внутри областей $V(0, \chi_+)$ больше (если $\alpha > 0$), чем плотность энергии окружающего пространства $V(0, \chi_-)$. Следовательно, эти области будут отделены от космологического расширения и в конечном итоге образуют плотные объекты. Некоторые из них коллапсируют, образуя черные дыры. Ситуация выглядит сходной с той, что наблюдалась в симметричном случае. Различие заключается в том,

что избыточная плотность энергии ложного вакуума приводит к более быстрому формированию более массивных ЧД.

Другая возможность представляется более интересной. Если наша Вселенная образовалась справа от критической линии, основное пространство будет заполнено полем в ложном вакууме $\chi = \chi_+$. Некоторые области, тем не менее, будут содержать истинный вакуум $\chi = \chi_-$. Некоторые из таких “областей истинного вакуума” будут расширяться, поглощая окружающее пространство, заполненное ложным вакуумом. Ситуация очень похожа на конечную стадию фазового перехода первого рода, но с существенным отличием — области истинного вакуума образуются без фазового перехода, по крайней мере первого рода. Как следствие, эти области при рождении не обладают сферической симметрией.

Сразу после быстрого падения в состояния ($\chi = \chi_-$, $\sigma = 0$) и ($\chi = \chi_+$, $\sigma = 0$), всё пространство будет заполнено ложным вакуумом с плотностью энергии $\varepsilon \equiv V(0, \chi_+) - V(0, \chi_-)$, за исключением некоторого числа областей истинного вакуума. Это означает начало новой, вторичной деситтеровской стадии с низкой плотностью энергии, которая может длиться продолжительное время. Достаточно большие области истинного вакуума способны разрушить новую инфляционную стадию, в то время как малые — сжимаются и исчезают, оставляя после себя колебания инфлатонного поля. Найдем критический размер, начиная с которого области расширяются, а не сжимаются, для областей сферической формы. Полная энергия такой области есть сумма энергии стенки E_{wall} и энергии объема E_V :

$$E = E_{\text{wall}} + E_V, \quad E_{\text{wall}} = 4\pi\mu R^2, \quad E_V = -\frac{4\pi}{3}\varepsilon R^3. \quad (6.15)$$

Кроме того, предположим, что выполняется естественное соотношение

$$\varepsilon < V(0, 0) = \varkappa^2 M^4. \quad (6.16)$$

Критическое значение радиуса находится из соотношения

$$\frac{dE}{dR} = 8\pi\mu R - 4\pi\varepsilon R^2 = 0,$$

которое приводит к следующему результату:

$$R_{\text{crit}} = \frac{2\mu}{\varepsilon}. \quad (6.17)$$

Если радиус области больше критического, R_{crit} , область расширяется.

После образования областей истинного вакуума имеется два пути эволюции Вселенной. Если число таких областей мало, их расширение неспособно компенсировать инфляционное расширение окружающих областей ложного вакуума с плотностью энергии $-\varepsilon$. В противоположном случае области истинного вакуума, расширяясь, эффективно захватывают окружающее пространство, преобразуя энергию ложного вакуума в кинетическую энергию стенок. Чтобы оценить минимально необходимое число таких областей, найдем прежде всего отношение ξ критического радиуса к параметру Хаббла H_ε^{-1} ($H_\varepsilon \approx \sqrt{\frac{8\pi}{3} \frac{\varepsilon}{M_{\text{Pl}}^2}}$),

$$\xi = \frac{R_{\text{crit}}}{H_\varepsilon^{-1}} = \frac{2\mu}{\varepsilon} \sqrt{\frac{8\pi}{3} \frac{\varepsilon}{M_{\text{Pl}}^2}} = \frac{2^6 \sqrt{\pi}}{3^{3/2}} \frac{\varkappa M^3}{\sqrt{\varepsilon} M_{\text{Pl}}}.$$

Стоит подчеркнуть, что параметр Хаббла сразу после окончания основного инфляционного периода $H_\varepsilon \sim \sqrt{\varepsilon}$ меньше, чем H во время инфляции, $H \sim \varkappa M^2$, т.е. $H_\varepsilon < H$.

Если хотя бы один островок истинного вакуума с радиусом $R \geq R_{\text{crit}}$ образуется в причинно связанной области после окончания первичной инфляции, то, расширяясь, он поглотит всю эту область. Вселенная окажется в постинфляционной стадии. Размер островка должен быть меньше хаббловского размера H_ε^{-1} во время вторичной инфляции, если $\xi < 1$. Это имеет место, если

$$\varepsilon > \frac{2^{12} \pi}{3^3} \frac{\varkappa^2 M^6}{M_{\text{Pl}}^2}.$$

Это неравенство и ограничение (6.16) легко удовлетворяются при $M/M_{\text{Pl}} < 10^{-2}$. Оценим среднее число островков истинного вакуума n_b и приравняем его полному числу n_d причинно связанных доменов. Полный объем Вселенной $V_U \simeq [H^{-1} \exp(N_U)]^3$, если предположить, что переход от первой ко второй инфляционной стадии произошел быстро. Объем причинно связанных доменов во время вторичной инфляции есть H_ϵ^{-3} . Тогда полное число островков есть

$$n_b = n_d = \frac{V_U}{H_\epsilon^{-3}} = \left[\frac{H_\epsilon}{H} e^{N_U} \right]^3 = \left[\frac{\sqrt{\epsilon}}{\kappa M^2} e^{N_U} \right]^3.$$

Так как число е-фолдов $N_U \approx 60$, полное число островков истинного вакуума оказывается очень большим — всего на несколько порядков меньше, чем число причинно связанных областей. Лишь при этом условии вторичная инфляция может закончиться полным переходом в истинный вакуум. Такое количество островков может образоваться, если на первой инфляционной стадии классическое значение поля приблизится к критической линии или даже пересечет ее. Таким образом, модернизированный потенциал свободен от основного дефекта гибридной инфляции — перепроизводства полевых стенок. В данном случае чем больше островков появится по окончании первичной инфляции, тем быстрее истинный вакуум заполнит все пространство Вселенной.

Процесс распада ложного вакуума сопровождается столкновением полевых стенок, что приводит к интенсивному рождению частиц. Структура пространства после завершения первичной инфляции достаточно сложна. Она состоит из случайно распределенных областей истинного и ложного вакуумов, разделенных полевыми стенками. Тем не менее, можно оценить энергию, выделяемую при столкновении стенок. Для этого предположим, что Вселенная заполнена стенками с характерным размером $R \sim H_\epsilon^{-1}$. Расстояние l между стенками — того же порядка, $l \sim H_\epsilon^{-1}$. Тогда полная энергия E_{tot} стенок примерно равна

$$E_{\text{tot}} \sim \mu n_b H_\epsilon^{-2}.$$

Характерное время t_{coll} столкновения стенок примерно равно расстоянию между ними, $t_{\text{coll}} \sim H_\varepsilon^{-1}$. С другой стороны, известно [199], что при столкновении двух стенок происходит эффективное выделение энергии при осцилляциях стенок вокруг их общего центра масс. Таким образом, время выделения энергии $t_{\text{coll}} \sim H_\varepsilon^{-1}$ очень мало. Эта энергия выделяется в виде колебаний полей σ и χ . Масса квантов поля χ , $m_\chi = V''(\chi, 0) = \sqrt{3}\varkappa M$. Эти массивные частицы должны распадаться, если они связаны с фермионами, например, как $g\chi\bar{\psi}\psi$. Вероятность такого распада есть $\Gamma \sim g^2 m_\chi / 4\pi$. Рождающиеся релятивистские фермионы нагревают окружающую среду в течение всего периода t_{heat} их образования. Очевидно, этот период больше хаббловского времени H_ε^{-1} на тот момент:

$$t_{\text{heat}} \approx \Gamma^{-1} \sim 4\pi / (g^2 m_\chi) \sim 1 / (g^2 \varkappa M).$$

Полагая, что один островок истинного вакуума приходится на масштаб H_ε^{-1} , можно оценить плотность энергии ρ_{end} среды сразу по окончании инфляции:

$$\rho_{\text{end}} \sim \rho_{\text{wall}} \sim \mu H_\varepsilon^{-2} / H_\varepsilon^{-3}.$$

Плотность энергии фермионов (ψ -частиц) меняется со временем следующим образом:

$$\rho_\psi \approx \frac{\rho_{\text{end}}}{\left(1 + \frac{3}{2}\gamma\tau\right)^2} \left[1 - e^{-\tau/\tau_0}\right]. \quad (6.18)$$

Здесь $\gamma = 1$ для пыли и $\gamma = 4/3$ для газа ультрарелятивистских частиц, $\tau_0 = H_\varepsilon / \Gamma$ и время $t = H_\varepsilon^{-1}\tau$. Первый множитель в выражении (6.18) появляется из-за расширения Вселенной, а второй — из-за распада частиц квантов χ -поля.

Чтобы оценить температуру среды $T_{\text{heat}} \sim \rho_\psi^{1/4}$, выберем реалистические значения параметров: $\varepsilon = 10^{-4}\varkappa^2 M^4$, $\varkappa = 10^{-4}$ и $M = 10^{-4}M_{\text{Pl}}$. Очевидно, что рассмотренный механизм эффективно нагревает среду — температура $T_{\text{heat}} \approx 6 \cdot 10^{-9}M_{\text{Pl}}$ очень

высока, что может привести к перепроизводству нейтралино, если таковые существуют.

Итак, рассмотренный механизм квантовых флуктуаций указывает на то, что гибридная инфляция должна содержать малый параметр, чтобы выводы на ее основе не противоречили наблюдательным данным. Добавление небольшого наклона к потенциалу существенно облегчает ситуацию. Изменение формы потенциала приводит к нетривиальной динамике скалярного поля и, в частности, к новому механизму перехода от инфляционной стадии к обычному космологическому расширению в пространстве ФРУ с эффективным производством частиц материи. Механизм образования массивных первичных черных дыр [268] также имеет место в рамках гибридной инфляции.

6.3. Влияние массивных полей на процесс инфляции

В этом разделе мы обсудим влияние квантовых флуктуаций на классическую эволюцию инфлатонного поля. Целью анализа будет изучение обратного влияния облака виртуальных частиц, создаваемых инфлатонным полем, на движение самого поля. Подчеркнем, что здесь рассматриваются эффекты взаимодействия только с виртуальными частицами, которые никак не связаны с температурой окружающей среды. Известно, что такое облако, обладая инерционными свойствами, может приводить к заметно замедлению эволюции системы. Это явление хорошо изучено в теории твердого тела (поляронный эффект [202]). Весь анализ проводится в метрике Минковского, т.е. без учета гравитационных эффектов.

Инерционные свойства виртуального облака зависят не столько от вида потенциала взаимодействия и сорта частиц, сколько от величины константы взаимодействия и массы виртуальных частиц, поскольку именно последние определяют насыщенность облака частицами различных сортов. Очевидно также, что вклю-

чение новых полей лишь усилит эффект торможения при движении классического поля. Поэтому в дальнейшем рассматривается простейшая форма взаимодействия, которая позволяет получать аналитические результаты. Именно, рассмотрим, кроме инфлатонного поля φ , дополнительное скалярное поле χ и запишем действие в виде

$$S = \int d^4x \sqrt{-g} \left[\frac{1}{2} \varphi_{,\mu} \varphi^{,\mu} - V(\varphi) + \frac{1}{2} \chi_{,\mu} \chi^{,\mu} - \frac{1}{2} m_\chi^2 \chi^2 - \varkappa \chi u(\varphi) \right]. \quad (6.19)$$

Здесь $u(\varphi)$ — полином относительно φ со степенью не выше третьей для перенормируемых теорий. Первая степень дополнительного поля χ во взаимодействии необходима лишь для того, чтобы получить в конечном итоге обозримые аналитические результаты, справедливые для произвольной константы связи \varkappa . Взаимодействие такого рода возникает в суперсимметричных теориях и рассматривается в гибридных моделях инфляции [218]. В статье [134] подобная форма взаимодействия используется при изучении обратного влияния рождающихся частиц на классическое движение инфлатонного поля. Как указывалось выше, учет реальных взаимодействий, например, с фермионами, лишь усилит эффект торможения. Поскольку нас интересует взаимодействие с виртуальными χ -частицами, рассмотрим амплитуду перехода

$$A(\varphi_i, \chi_i = 0; \varphi_f, \chi_f = 0) = \int_{\varphi_i}^{\varphi_f} D\varphi \int_0^0 D\chi \exp[iS]. \quad (6.20)$$

При этом предполагается, что поле χ достаточно массивно и к рассматриваемому моменту находится около минимума своего эффективного потенциала. Проинтегрировав по полевой переменной χ , получим амплитуду вида (6.20) с эффективным действием

$$S_{\text{eff}} = \int d^4x \sqrt{-g} \left[\frac{1}{2} \varphi_{,\mu} \varphi^{,\mu} - V(\varphi) \right] + \frac{\varkappa^2}{2} \int d^4x \int d^4x' \sqrt{-g} u(\varphi(x)) G(x, x') u(\varphi(x')), \quad (6.21)$$

где $G(x, x')$ — функция Грина скалярного поля χ . Полученное выражение является точным, но из-за присутствия нелокального слагаемого, которое физически и интерпретируется как вклад от облака виртуальных χ -частиц, точные аналитические преобразования невозможны. Поэтому, как и в методе эффективного действия [175], разложим нелокальное слагаемое в (6.21) по степеням $x - x'$. При этом первые несколько членов перенормируют исходные параметры потенциала и полевую переменную φ (волновую функцию). Поскольку параметры потенциала все равно определяются физическими условиями на некотором энергетическом масштабе, нас будут интересовать новые слагаемые, которые не могут быть включены в перенормировки.

При явном разложении выражения (6.21) в ряд по степеням $x - x'$ потребуется интегрирование функции Грина, что, при учете гравитационных эффектов, является непростой задачей. Вместо этого воспользуемся уравнением на функцию Грина [68], записанным в виде

$$G(x, x') = \frac{1}{m_\chi^2} \delta(x - x') - \frac{1}{m_\chi^2} \frac{1}{\sqrt{-g}} \partial_\mu \sqrt{-g} \partial^\mu G(x, x'). \quad (6.22)$$

После подстановки этого выражения в (6.21) и пренебрежения высшими производными по полю φ получим

$$S_{\text{eff}} = \int d^4x \sqrt{-g} \left[\frac{1}{2} \varphi_{,\mu} \varphi^{,\mu} - V_{\text{ren}}(\varphi) - \frac{1}{\sqrt{-g}} \frac{\alpha^2}{2m_\chi^4} u(\varphi) \partial_\mu \sqrt{-g} \partial^\mu u(\varphi) \right] \quad (6.23)$$

с соответствующим уравнением движения

$$\partial_\mu \sqrt{-g} \partial^\mu \varphi + \sqrt{-g} V'_{\text{ren}}(\varphi) + \frac{\alpha^2}{m_\chi^2} u'_\varphi(\varphi) \partial_\mu \sqrt{-g} \partial^\mu u(\varphi) = 0 \quad (6.24)$$

и энергией вида

$$E = T_{00} = \frac{1}{2} (\partial_0 \varphi)^2 + V_{\text{ren}}(\varphi) + \frac{\alpha^2}{2m_\chi^2} (\partial_0 u)^2, \quad (6.25)$$

где введен безразмерный параметр $\alpha \equiv \varkappa/m_\chi$.

Перенормированный эффективный потенциал содержит сумму вкладов от поправок из-за взаимодействия со всеми существующими полями. Вклад от взаимодействия с полем χ легко получить в явном виде: $\delta V_\chi = -(\alpha^2/2)\varphi^4$. Недостатком первой модели хаотической инфляции с потенциалом вида $\lambda_{\text{gen}}\varphi^4$ являлась малость константы связи $\lambda_{\text{gen}} (\sim 10^{-13})$, при которой не возникало противоречий с наблюдательными данными. Это означает, что все слагаемые в выражении для λ_{gen} , включая и δV_χ , должны сокращаться с высокой степенью точности. Далее показано, что учет перенормировки кинетического члена позволяет, в частности, сильно ослабить условия на параметры теории, налагаемые наблюдательными данными.

Слагаемые, аналогичные последнему в выражении (6.23), возникают при перенормировке любой теории и обычно интерпретируются как “квантовые поправки к параметрам теории, зависящие от значения самого поля” [175]. В слабых полях вклад от этого слагаемого пренебрежимо мал. На инфляционной же стадии, при больших величинах полей, последнее слагаемое в уравнении (6.23) может оказаться существенным.

Классическое уравнение (6.24) может быть получено и другим путем, с использованием классических уравнений, минимизирующих исходное действие (6.19),

$$\begin{aligned} \frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}\partial^\mu\chi) + m_\chi^2\chi + \varkappa\varphi^2 &= 0, \\ \frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}\partial^\mu\varphi) + V'(\varphi) + 2\varkappa\varphi\chi &= 0. \end{aligned} \quad (6.26)$$

Здесь рассмотрен случай $u(\varphi) = \varphi^2$. Первое из уравнений (6.26) можно преобразовать к виду

$$\chi(x) = -\varkappa \int G(x, x')\varphi^2(x')dx'. \quad (6.27)$$

Подставляя (6.22) и (6.27) во второе уравнение системы (6.26),

приходим в том же приближении медленного движения к уравнению (6.24).

Отметим, что поправка к потенциалу $\delta V = -(\alpha^2/2)\varphi^4$ возникла из анализа классических уравнений (6.26). С другой стороны, точно такой же ответ может быть получен при вычислении первой квантовой поправки к потенциалу поля φ за счет взаимодействия с полем χ при нулевых 4-импульсах внешних линий, соответствующих квантам поля φ . Последнее слагаемое в выражении (6.24) обычно интерпретируется как “обратная реакция излучения” [134].

Во время инфляции поле считается однородным, $\varphi = \varphi(t)$, и уравнение (6.24) сильно упрощается. Запишем его для конкретного случая $u(\varphi) = \varphi^2$, учитывая также, что масштабный фактор a обычным образом выражается через параметр Хаббла H , $a = \exp(\int H dt)$:

$$\frac{d^2\varphi}{dt^2} + 3H\frac{d\varphi}{dt} + V'(\varphi) + \frac{4\alpha^2}{m_\chi^2} \left[3H\varphi^2\frac{d\varphi}{dt} + \varphi^2\frac{d^2\varphi}{dt^2} + \varphi\left(\frac{d\varphi}{dt}\right)^2 \right] = 0.$$

Учитывая медленное изменение поля φ со временем, отбросим слагаемые, пропорциональные $d^2\varphi/dt^2$ и $(d\varphi/dt)^2$, что оправдывается конечным результатом, после чего приходим к легко интегрируемому уравнению

$$\left(3H + \frac{12H\alpha^2}{m_\chi^2}\varphi^2 \right) \dot{\varphi} + V'(\varphi) = 0. \quad (6.28)$$

Решая это уравнение для потенциала вида $V(\varphi) = \lambda\varphi^n$ с учетом обычной связи параметра Хаббла и потенциала,

$$H = \sqrt{8\pi V(\varphi)/3}/M_{\text{Pl}},$$

приходим к выражению для неявной зависимости полевой переменной φ от времени

$$t = \frac{\sqrt{24\pi}}{M_{\text{Pl}}\lambda^{1/2}} \left[\frac{1}{n(2-n/2)} \left(\varphi_0^{2-n/2} - \varphi^{2-n/2} \right) + \right.$$

$$+ \frac{4}{n(4-n/2)} \frac{\alpha^2}{m_\chi^2} \left(\varphi_0^{4-n/2} - \varphi^{4-n/2} \right) \Big]. \quad (6.29)$$

Здесь первое слагаемое воспроизводит результат стандартной инфляционной модели, а второе учитывает взаимодействие с виртуальными χ -частицами.

Необходимо отметить, что, согласно уравнению (6.28), имеется промежуточная стадия инфляции, когда уже нельзя пренебрегать второй производной по полю φ в квадратных скобках. Тем не менее, наиболее интересной оказывается первая стадия, и именно она исследуется ниже. Дальнейшее рассмотрение проводится в предположении доминантности второго слагаемого в уравнениях (6.28),(6.29), т.е. при условии

$$2\alpha\varphi/m_\chi \gg 1. \quad (6.30)$$

Удобно для дальнейшего ввести величину

$$\varphi_c \equiv \frac{m_\chi}{2\alpha}, \quad (6.31)$$

представляющую собой пограничное значение поля и отделяющую первую стадию инфляции со сверхмедленным движением поля φ от более поздней стадии обычной инфляции. В обоих случаях динамика поля описывается простыми аналитическими выражениями.

В случае $\varphi > \varphi_c$ временная зависимость поля есть

$$\begin{aligned} \varphi(t) &= \left[\varphi_0^{4-n/2} - t/f_n \right]^{1/(4-n/2)}, \\ f_n &= \frac{8\sqrt{6\pi}}{n(4-n/2)} \frac{\kappa^2}{m_\chi^4 M_{\text{Pl}} \lambda^{1/2}}. \end{aligned} \quad (6.32)$$

Это выражение получено при условии “сверхмедленного скатывания”, которое, согласно уравнению (6.25), имеет не совсем обычный вид

$$\ddot{\varphi} \ll 12H\varphi^2 \dot{\varphi} \frac{\alpha^2}{m_\chi^2}. \quad (6.33)$$

Важно отметить, что, как и ожидалось, скорость изменения поля

$$\dot{\varphi} = \frac{m^2}{12\alpha^2\varphi^2} \cdot \frac{V'}{H},$$

полученная из выражения (6.28) с учетом (6.30), оказывается гораздо меньше обычного значения $\dot{\varphi} = V'/3H$. Первый этап инфляции заканчивается, когда перестает выполняться условие (6.33). Далее начинается период обычной инфляции, который длится, пока выполняется условие $\ddot{\varphi} \ll 3H\dot{\varphi}$. Как показывают оценки, второй этап оказывается очень коротким.

Определим величину квантовых флуктуаций, возникающих на инфляционной стадии. Флуктуации невзаимодействующих полей изучены достаточно хорошо [210]. В то же время, для полей с самодействием известны лишь оценки по порядку величины [131].

Определим величину квантовых флуктуаций, возникающих на первой стадии инфляции для потенциала $\lambda\varphi^4$. Проще всего это удается сделать, если ввести вспомогательное поле $\tilde{\varphi}$ и заменой $\tilde{\varphi} = (\alpha/m_\chi)\varphi^2$ преобразовать действие (6.23) к виду

$$S = \int dx \sqrt{-g} \left[\partial_\mu \tilde{\varphi} \partial^\mu \tilde{\varphi} - \frac{1}{2} \tilde{m}^2 \tilde{\varphi}^2 \right], \quad (6.34)$$

отвечающему свободному полю с массой $\tilde{m} \equiv m_\chi \sqrt{2\lambda}/\alpha$. Такая замена правомерна на рассматриваемой стадии инфляции, когда значения поля заведомо больше нуля. Величина флуктуаций массивного поля известна [103], $\Delta\tilde{\varphi} = \sqrt{3/(8\pi^2)} H^2/\tilde{m}$. Известно также и ограничение на массу поля $\tilde{\varphi}$, получаемое сравнением с измерениями COBE [64] флуктуаций плотности энергии, $\delta\rho/\rho \approx 6 \cdot 10^{-5}$ на масштабе современного горизонта: $\tilde{m} \sim 10^{-6} M_{\text{Pl}}$. Выразив \tilde{m} через исходные параметры, получаем связь последних между собой

$$\frac{m_\chi}{M_{\text{Pl}}} \frac{\sqrt{\lambda}}{\alpha} \sim 10^{-6}. \quad (6.35)$$

Поскольку естественно полагать, что $m_\chi \ll M_{\text{Pl}}$, неравенство (6.35), полученное из наблюдательных данных, не является обременительным и выполняется в широком интервале параметров.

Определим значение поля φ_U , при котором образовалась причинно связанная область, породившая видимую часть Вселенной. Известно, что число е-фолдов, необходимых для объяснения наблюдательных данных, есть $N_U \simeq 60$. Тогда, используя, как обычно, связь $N_U = \int_{\varphi_U}^{\varphi_{\text{end}}} H dt$, имеем

$$\begin{aligned} N_U &= \int_{\varphi_U}^{\varphi_c} \frac{H(\varphi)}{\dot{\varphi}} d\varphi + \int_{\varphi_c}^{\varphi_{\text{end}}} \frac{H(\varphi)}{\dot{\varphi}} d\varphi = \\ &= \frac{2\pi\alpha^2}{M_{\text{Pl}}^2 m_\chi^2} (\varphi_U^4 - \varphi_c^4) + \frac{\pi}{M_{\text{Pl}}^2} (\varphi_c^2 - \varphi_{\text{end}}^2). \end{aligned} \quad (6.36)$$

Учитывая различную временную зависимость поля φ на первом и втором этапах инфляции и используя малость поля φ_{end} сразу по окончании инфляции по сравнению с начальным значением φ_U , получаем искомое выражение

$$\varphi_U \simeq \left(\frac{N_U}{2\pi} \right)^{1/4} \sqrt{\frac{M_{\text{Pl}} m_\chi}{\alpha}}. \quad (6.37)$$

Отметим, что видимая часть Вселенной в данном случае может образоваться при $\varphi < M_{\text{Pl}}$, при более низких энергиях по сравнению с квадратичной инфляцией. Это объясняется тем, что на первом этапе поле движется “сверхмедленно”, и Вселенная успевает расшириться до нужных размеров.

Все приведенные рассуждения справедливы, если поле χ имеет достаточно большую массу, так что во время инфляции его можно было рассматривать находящимся в минимуме своего потенциала. Как известно, поле начинает быстро скатываться к минимуму, если параметр Хаббла становится меньше его массы, $H < m$. Параметр Хаббла зависит от времени, поэтому проведем необходимые оценки в начальный момент возникновения видимой Вселенной ($\varphi = \varphi_U$) и в момент окончания первого этапа

инфляции ($\varphi = \varphi_c$). Простые выкладки приводят к следующему результату:

$$\begin{aligned} m_\chi > H(\varphi_U) &\quad \mapsto \quad \frac{\sqrt{\lambda}}{\alpha} \leq 0.1, \\ m_\chi > H(\varphi_c) &\quad \mapsto \quad m_\chi \leq M_{\text{Pl}} \frac{\sqrt{\lambda}}{\alpha^2}. \end{aligned} \quad (6.38)$$

Эти неравенства выполняются в широком диапазоне параметров, по крайней мере вблизи окончания первого периода инфляции. Квантовые поправки к массе m_χ поля χ ограничивают массу снизу, поэтому далее будем полагать, что $m_\chi \geq \varkappa$, или, что то же самое, $\alpha \leq 1$.

Таким образом, взаимодействие инфлатонного поля и создаваемых им квантовых флуктуаций массивного поля другого сорта позволяет объяснить, например, малость флуктуаций плотности энергии при достаточно мягких условиях на параметры потенциала.

6.4. Фазовые переходы первого рода

В ранней Вселенной, по-видимому, происходило несколько фазовых переходов. К таковым можно отнести переход на масштабе Великого объединения при $\sim 10^{17}$ ГэВ и на масштабе нарушения электрослабого взаимодействия при ~ 100 ГэВ. До сих пор нами рассматривался класс моделей, базой которых служит хаотическая инфляция. Существует также и инфляционный сценарий, основанный на фазовом переходе первого рода, с которого начинается эволюция Вселенной [171]. В теории поля фазовые переходы первого рода происходят посредством распада метастабильного состояния (ложного вакуума) на сферически симметричные области, занятые истинным вакуумом, рис.6.2. Эти области начинают расширяться, занимая все больший объем. Вычисление вероятности такого распада представляет собой нетривиальную задачу.

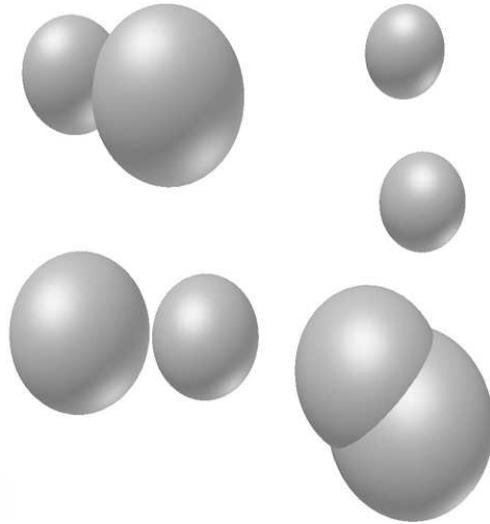


Рис. 6.2. Переход ложного вакуума в истинный происходит путем подбарьерных переходов с образованием пузырей истинного вакуума и их последующим расширением.

Эта проблема ранее неоднократно обсуждалась в литературе, причем в различных разделах физики, таких как теория поля, космология, физика твердого тела и физика фазовых переходов. Теоретические аспекты этого явления с квантово-механической точки зрения изучаются достаточно давно [22]. Существенное продвижение в исследованиях в квантовой теории поля начались в семидесятых годах, см. [120], в них заложены основы инстантонного подхода к вычислению распада вакуума. Этот же метод применим и при расчете вероятности квантово-механического туннелирования. Основная идея заключается в том, что контур интегрирования деформируется так, чтобы основной вклад в интеграл, представляющий собой амплитуду перехода, давала одна траектория — “инстантон” $x_{\text{inst}}(t_E)$, где t_E — время в евклидовом пространстве. Полученная таким образом амплитуда подбарьерного перехода содержит необходимый подавляющий фактор

$\exp -S[x_{\text{inst}}]$.

Дальнейшее продвижение в теории подбарьерных переходов выявило существование многочисленных дополнительных факторов, влияющих на динамику перехода. Эти факторы могут серьезно изменить физическую картину перехода и повлиять на конкретные выводы той или иной космологической модели.

Одним из основных факторов является температура, при которой происходит подбарьерный переход. В космологических фазовых переходах температурные поправки к эффективному потенциалу сильно искажают его форму, так что при высоких температурах фазовый переход может просто отсутствовать [210]. С понижением температуры появляются дополнительные минимумы потенциала, и фазовый переход становится возможным, хотя параметры эффективного потенциала на момент перехода отличаются от тех, которые наблюдаются в настоящее время.

Другим основополагающим фактором является окружающая среда. Взаимодействие с ней приводит обычно к уменьшению вероятности туннелирования. При изучении фазового перехода за счет распада ложного вакуума необходимо учитывать диссипацию энергии во время самого фазового перехода, при движении стенки, отделяющей истинный вакуум от ложного. Это особенно важно при изучении электрослабого перехода на ранней стадии развития Вселенной [215], когда плотность окружающей среды высока.

Третьим фактором является взаимодействие частицы или классического поля с виртуальными частицами другого сорта. Модель “тяжелого нуклона” также подтверждает вывод о том, что облако скалярных частиц, взаимодействующих с нуклоном, влияет на перенормировку как волновой функции, так и его массы.

В этом разделе предлагается метод вычисления квантовых поправок к вероятности распада вакуума в теории поля [20, 21]. Изложен способ суммирования низколежащих уровней, который позволяет корректно определить вклад от квантовых поправок.

Ниже показано, что поправки, будучи малыми по сравнению с основным вкладом, тем не менее приводят к заметному изменению величины основного эффекта — вероятности подбарьерного перехода. Изучается влияние облака виртуальных частиц на классическое движение скалярного поля, а также на вероятность распада ложного вакуума.

6.4.1. Квантовые поправки к вероятности распада вакуума

Здесь мы рассмотрим метод вычисления функциональных детерминантов, основанный на определении асимптотики функции Грина уравнения теплопроводности и исследовании дискретного спектра соответствующего уравнения Шредингера. Предлагаемый метод был использован для вычисления в однопетлевом приближении вероятности распада метастабильного вакуума в скалярной (3+1)-теории.

Рассмотрим вещественное скалярное поле с лагранжианом

$$L = \frac{1}{2}(\partial\Phi)^2 - U(\Phi), \quad (6.39)$$

где $U(\Phi)$ — полином четвертой степени, обладающий двумя невырожденными минимумами Φ_{\pm} . Хотя классически оба состояния стабильны, за счет квантовых поправок вакуумное состояние, построенное на поле Φ_{+} , становится нестабильным и со временем распадается. Состояние Φ_{+} принято называть метастабильным или ложным вакуумом. В теории поля анализ амплитуды распада метастабильного вакуумного состояния обычно проводится в рамках квазиклассического приближения с использованием функционального интегрирования. Однако при этом возникают сложности с вычислением функционального детерминанта, появляющегося при интегрировании по малым отклонениям от классической траектории. Так, если квазиклассические ведущие экспоненциальные факторы были вычислены

достаточно давно (см., например, [217]), то проблема вычисления функциональных детерминантов существует до сих пор.

В данном разделе предлагается метод вычисления функциональных детерминантов эллиптических операторов, основанный на нахождении асимптотики функции Грина уравнения теплопроводности для этого оператора и исследовании дискретного спектра соответствующего уравнения Шредингера в пространстве Минковского. Предлагаемый метод используется далее для вычисления вероятности распада метастабильного вакуума в скалярной (3+1)-теории.

Рассмотрим оператор D , заданный на евклидовом многообразии размерности d :

$$D(x) = -\Delta + u(x), \quad (6.40)$$

где Δ — оператор Лапласа и $u(x)$ — вещественная гладкая функция. Ядро оператора D можно представить через собственные функции $\Psi_n(y)$ и собственные значения λ_n оператора D следующим образом:

$$G(x, y, t) = \sum_n \Psi_n(x) \Psi_n(y) \exp(-\lambda_n t), \quad (6.41)$$

причем в этом случае $G(x, y, t)$ является функцией Грина уравнения теплопроводности

$$\left[\frac{\partial}{\partial t} + D(x) \right] G(x, y, t) = 0, \quad G(x, y, 0) = \delta(x - y), \quad t > 0. \quad (6.42)$$

Тогда, интегрируя по x и используя обратное преобразование Меллина, регуляризованный детерминант оператора D [164] можно представить в следующем виде:

$$\frac{\text{Det } D}{\text{Det } D_0} = \exp \left\{ - \lim_{d \rightarrow \infty} \frac{d}{ds} \frac{M^{2s}}{\Gamma(s)} \int_0^\infty dt t^{s-1} \cdot \int dx [G(x, x, t) - G_0(x, x, t)] \right\}, \quad (6.43)$$

где все выражение нормировано на детерминант оператора

$$D_0 = -\Delta + \mu_+^2, \quad u(x \rightarrow \pm\infty) \rightarrow \mu_+^2, \quad (6.44)$$

M — параметр с размерностью массы. Формулу (6.43) можно записать в несколько ином виде, используя определение (6.41) ядра оператора:

$$\frac{\text{Det } D}{\text{Det } D_0} = \exp \left\{ - \lim_{s \rightarrow 0} \frac{d}{ds} \frac{M^{2s}}{\Gamma(s)} \int_0^\infty dt t^{s-1} \sum_n \left[e^{-\lambda_n t} - e^{-\lambda_n^{(0)} t} \right] \right\}. \quad (6.45)$$

Выражения (6.43), (6.45) представляют собой две эквивалентные формы нормированного и регуляризованного через обобщенную ζ -функцию детерминанта оператора D . Таким образом, для вычисления функционального детерминанта нам необходимо найти функцию $G(x, y, t)$, удовлетворяющую уравнению (6.42), либо знать уровни дискретного спектра оператора D и фазовые сдвиги в континууме. В общем случае это неразрешимая задача. Однако ниже мы покажем, что приближенно функциональный детерминант может быть вычислен на основе известного асимптотического поведения функции Грина $G(x, y, t)$ при $t \rightarrow 0$ и знания уровней дискретного спектра.

Решение уравнения (6.42) может быть представлено в виде абсолютно сходящегося ряда

$$G(x, y, t) = \exp(-\mu_+^2 t) \left\{ Z(x, t; y, 0) + \int_0^t \int_L dy_1 d\tau_1 C(y_1) \cdot Z(y_1, \tau_1; y, 0) + \int_0^t \int_{L_1} \int_0^{\tau_1} \int_{L_2} dy_2 d\tau_2 dy_1 d\tau_1 Z(x, t; y_1, \tau_1) C(y_1) \cdot Z(y_1, \tau_1; y_2, \tau_2) C(y_2) Z(y_2, \tau_2; y, 0) + \dots \right\},$$

$$Z(x, t; y, \tau) = [4\pi(t - \tau)]^{-d/2} \exp \left[-\frac{(x - y)^2}{4(t - \tau)} \right],$$

$$C = C(x) = \mu_+^2 - u(x). \quad (6.46)$$

Разлагая функцию $C(x)$ в ряд Тейлора по степеням $(y - y_n)$ и выполняя соответствующие интегрирования в (6.46), получаем асимптотическое выражение для $G(x, y, t)$ при $t \rightarrow 0$:

$$G(x, x, t) = \frac{\exp(-\mu_+^2 t)}{(4\pi t)^{d/2}} \sum_{k=0}^{\infty} a_k \frac{t^k}{k!}, \quad (6.47)$$

где шесть первых коэффициентов a_k имеют вид

$$\begin{aligned} a_0 &= 1, & a_1 &= C, & a_2 &= C^2, & a_3 &= C^3 - \frac{1}{2}(\nabla C)^2, \\ a_4 &= C^4 + C^2 \Delta C + (1/5)(\Delta C)^2, \\ a_5 &= C^5 + (5/3)C^3 \Delta C + (1/3)C(\Delta C)^2 + \\ &\quad + (1/3)C^2 \Delta \Delta C + (1/14)\Delta C \Delta \Delta C. \end{aligned} \quad (6.48)$$

Хотя общего правила для построения коэффициентов a_k получить не удастся, однако, как было показано в работе [246], эти коэффициенты являются обобщением инвариантов одномерного уравнения Кортевега — де Фриза, и для нахождения этих коэффициентов была получена рекуррентная формула.

В случае и $u(x) = \mu_+^2$ для функции $G_0(x, x, t)$ получаем

$$G_0(x, x, t) = \frac{\exp(-\mu_+^2 t)}{(4\pi t)^{d/2}}. \quad (6.49)$$

Если бы разность функций $\int G_0(x, x, t)dx$ и $\int G(x, x, t)dx$ была быстро убывающей функцией переменной t , то функциональный детерминант (6.43) можно было бы приближенно вычислить, обрезая интеграл по t на некоторой величине δ . Однако если спектр оператора D содержит дискретные низколежащие уровни ($\lambda_n \sim 0$), то такой способ становится неэффективным, так как разность $\int G_0(x, x, t)dx - \int G(x, x, t)dx$ убывает медленно и область больших t дает существенный вклад, в результате чего при вычислениях уже недостаточно ограничиваться только несколькими первыми коэффициентами a_k . Указанной трудности можно избежать, разбив область интегрирования по t на две

области: $0 \div \delta$ и $\delta \div \infty$, причем в первой области использовать для функционального детерминанта формулу (6.43), а во второй — формулу (6.45):

$$\begin{aligned} \frac{\text{Det } D}{\text{Det } D_0} = \exp \left\{ - \lim_{s \rightarrow 0} \frac{d}{ds} \frac{M^{2s}}{\Gamma(s)} \int_0^\delta dt t^{s-1} \cdot \right. \\ \cdot \int dx [G(x, x, t) - G_0(x, x, t)] - \\ - \int_\delta^\infty dt t^{s-1} \sum_{n=0}^{n=n_0} [e^{-\lambda_n t} - e^{-\lambda_n^{(0)} t}] - \\ \left. - \int_\delta^\infty dt t^{s-1} \sum_{n=n_0+1}^{n=\infty} [e^{-\lambda_n t} - e^{-\lambda_n^{(0)} t}] \right\}, \quad (6.50) \end{aligned}$$

где $n = n_0 + 1$ соответствует началу континуума ($\lambda_{n_0} = \mu_+^2$). Если последнее слагаемое в выражении (6.50) является быстро убывающей функцией δ , то при больших δ этим слагаемым можно пренебречь.

Вычислим вероятность распада метастабильного вакуума в (3+1)-скалярной теории с лагранжианом вида (6.39), где

$$U(\Phi) = \frac{\lambda}{8}(\Phi^2 - a^2)^2 + \frac{\varepsilon}{2a}(\Phi - a). \quad (6.51)$$

Масса квантов поля Φ , соответствующих ложному вакууму Φ_+ , как обычно, определяется следующим образом:

$$\mu_+^2 = U''(\Phi_+) \simeq \mu^2 \left(1 - \frac{3\varepsilon}{2\lambda a^4} \right) + O(\varepsilon^2), \quad \mu^2 = \lambda a^2. \quad (6.52)$$

Сферически-симметричное решение Φ (баунс) уравнения движения

$$\Delta \Phi = U'(\Phi) \quad (6.53)$$

для потенциала (6.51) с граничным условием $\Phi(r \rightarrow \infty) \rightarrow \Phi_+$ соответствует минимуму действия. Решение уравнения (6.53) с

соответствующим граничным условием в первом приближении по параметру ε имеет вид

$$\Phi(r) = a \left[\operatorname{th} \left(\frac{\mu}{2}(r - R) \right) - \frac{\varepsilon}{2\lambda a^4} \right] + O(\varepsilon^2), \quad (6.54)$$

где $R = 2\lambda a^4/\varepsilon\mu$ — радиус пузыря. Использованное нами приближение $\varepsilon\lambda \ll \mu^4$ ($R\mu \gg 1$) соответствует так называемому приближению тонкой стенки, так как решение Φ напоминает большой пузырь истинного вакуума, отделенного от ложного вакуума в окружающем пространстве тонкой стенкой с толщиной $d = 2/\mu$. Классическое действие на экстремальной траектории (баунсе) в n -мерном пространстве есть

$$\begin{aligned} B \equiv S(\Phi) - S(\Phi_+) &= \int d^n x \left[\frac{1}{2}(\partial\Phi)^2 + U(\Phi) - U(\Phi_+) \right] = \\ &= \frac{4}{3} \frac{\pi^{n/2}}{n\Gamma(n/2)} \frac{R^{n-1}\mu^3}{\lambda}. \end{aligned} \quad (6.55)$$

После выполнения функционального интегрирования вблизи экстремальной траектории выражение для вероятности распада метастабильного вакуума с учетом однопетлевых квантовых поправок может быть представлено в виде

$$\Gamma = V \left(\frac{B}{2\pi} \right)^2 \left| \frac{\operatorname{Det}'[-\Delta + U''(\Phi)]}{\operatorname{Det}'[-\Delta + U''(\Phi_+)]} \right|^{-1/2} e^{-B}, \quad (6.56)$$

где V — нормировочный трехмерный объем, Det' означает, что из детерминанта выделены четыре нулевые моды, каждая из которых дала фактор $\sqrt{B/2\pi}$. Кроме того, следует отметить, что если мы используем приближение тонкой стенки, то искомым детерминант содержит только одну отрицательную моду.

Исследуем теперь спектр собственных значений уравнения Шредингера для оператора $D = -\Delta + U''(\Phi)$ в четырехмерном пространстве. Так как оператор D ротационно инвариантен, то угловая зависимость его собственных функций должна быть такой же, как у четырехмерных скалярных сферических

гармоник. Тогда, представляя собственные функции Ψ в виде $\Psi = f_{nl}(r)H_l(r)/r^{l+3/2}$, где $H_l(r)$ удовлетворяет уравнению Лапласа $\Delta H_l(r) = 0$, мы получаем уравнение для радиальных собственных функций

$$\left[\frac{d^2}{dr^2} + \frac{l(l+2) + 3/4}{r^2} + U''(\Phi) \right] f_{nl}(r) = \lambda_{nl} f_{nl}(r), \quad (6.57)$$

причем кратность вырождения собственных значений λ_{nl} равна $(l+1)^2$. Вводя новую переменную $z = \mu(r-R)/2$ и разлагая центробежный потенциал в ряд вблизи точки $r = R$ (это оправдано при исследовании дискретного спектра, так как связанные состояния локализованы в стенке пузыря, т.е. при $r \sim R$), получаем

$$\left[\frac{d^2}{dz^2} + \frac{4\lambda_{nl}}{\mu^2} - 4 - \frac{4l(l+2) + 3}{(R\mu)^2} + 6 \operatorname{ch}^{-2} z + v(z) \right] f_{nl}(r) = 0, \quad (6.58)$$

где величину $v(z)$ рассматриваем как возмущение:

$$v(z) = \frac{12}{R\mu} \operatorname{th} z + \frac{4z}{(R\mu)^3} [4l(l+2) + 3] + \dots$$

Согласно уравнению (6.57), непрерывный спектр начинается с $\lambda_{nl} = \mu_+^2$. Кроме того, существует и дискретный спектр оператора D . Для невозмущенной задачи ($v(z) = 0$) связанные состояния соответствуют следующим собственным значениям:

$$\lambda_{0l}^{(0)} = [l(l+2) + 3/4]/R^2, \quad \lambda_{1l}^{(0)} = \lambda_{0l}^{(0)} + 3\mu^2/4.$$

Так как ведущие слагаемые в $v(z)$ нечетны по z , возмущение $v(z)$ будет давать вклад в λ_{nl} только начиная с членов $\sim 1/(R\mu)^2$, причем для низколежащих уровней λ_{0l} , $l \ll R\mu$ эта поправка может быть легко определена исходя из того, что значение $l = 1$ соответствует нулевым модам, и тогда

$$\lambda_{0l} = [l(l+2) + 3/4]/R^2 + O(l^2/(R\mu)^4). \quad (6.59)$$

Для уровня $n = 1$ с необходимой нам точностью $\lambda_{1l} = \lambda_{1l}^{(0)}$, т.е.

$$\lambda_{1l} = \mu^2 \left[\frac{3}{4} + \frac{l(l+2)}{(R\mu)^2} \right] + O\left(\frac{1}{(R\mu)^2}\right). \quad (6.60)$$

Что касается оператора D_0 , то он имеет только непрерывный спектр, начинающийся с $\lambda_{nl}^0 = \mu_+^2$.

Запишем теперь выражение (6.56) несколько иначе, выделив из $\text{Det}' D$ отрицательную моду, а из $\text{Det} D_0$ пять первых мод. Тогда

$$\Gamma = V \left(\frac{B}{2\pi} \right)^2 \sqrt{\frac{\lambda_{00}^{(0)} \left(\lambda_{0l}^{(0)} \right)^4}{|\lambda_{00}|}} \left| \frac{\text{Det} D^{(5)}}{\text{Det} D_0^{(5)}} \right|^{-1/2} e^{-B}, \quad (6.61)$$

где $\text{Det} D^{(5)}$ означает $\text{Det} D$ без пяти первых мод. Для вычисления функционального детерминанта, содержащегося в формуле (6.61), воспользуемся полученным нами ранее соотношением (6.50):

$$\begin{aligned} \ln \left(\frac{\text{Det} D^{(5)}}{\text{Det} D_0^{(5)}} \right) &= A_1 + A_2 + A_3, \\ A_1 &= - \lim_{s \rightarrow 0} \frac{d}{ds} \frac{M^{2s}}{\Gamma(s)} \int_0^\delta dt t^{s-1} \left\{ \int dx [G(x, x, t) - G_0(x, x, t)] - \right. \\ &\quad \left. - \sum_{l=0}^1 (l+1)^2 [e^{-\lambda_{0l}t} - e^{-\mu_+^2 t}] \right\}, \\ A_2 &= - \int_\delta^\infty \frac{dt}{t} \left\{ \sum_{l=2}^{l_0} (l+1)^2 e^{-\lambda_{0l}t} + \sum_{l=0}^{l_1} (l+1)^2 e^{-\lambda_{1l}t} - \right. \\ &\quad \left. - \frac{1}{6} [-30 + (l_0+1)(l_0+2)(2l_0+3) + \right. \\ &\quad \left. + (l_1+1)(l_1+2)(2l_1+3)] e^{-\mu_+^2 t} \right\}, \quad (6.62) \end{aligned}$$

где мы положили $\lambda_{nl}^{(0)} \simeq \mu_+^2$ при конечных l и n , а l_0 и l_1 соответствуют началу континуума $\lambda_{0l_0} = \lambda_{0l_1} = \mu_+^2$.

Воспользовавшись соотношениями (6.47)–(6.49) и подставляя в них $C = \mu_+^2 - U''(\Phi)$, находим

$$\begin{aligned} & \int dx [G(x, x, t) - G_0(x, x, t) - \sum_{l=0}^1 (l+1)^2 [e^{-\lambda_0 l t} - e^{-\mu_+^2 t}] = \\ & = \frac{(R\mu)^3}{16\mu^2 t} \left[9 - 3\mu^2 t + \frac{7}{10}(\mu^2 t)^2 - \frac{3}{35}(\mu^2 t)^3 - \frac{1}{280}(\mu^2 t)^4 + \dots \right] + \\ & + O(R\mu). \end{aligned} \quad (6.63)$$

Выполняя интегрирование по t в выражении (6.62), получаем (с использованием рекуррентных формул [11, 246] при $\delta\mu^2 \ll 1$)

$$\begin{aligned} A_1 = -\frac{(R\mu)^3}{16} & \left\{ -3 \left[\gamma + \ln \mu^2 \delta - \ln \left(\frac{\mu^2}{M^2} \right) \right] - \frac{9}{\mu^2 \delta} + \frac{7}{10} \mu^2 \delta - \right. \\ & \left. - \frac{3}{70} (\mu^2 \delta)^2 - \frac{1}{840} (\mu^2 \delta)^3 + \dots \right\} + O(R\mu), \end{aligned} \quad (6.64)$$

где γ — постоянная Эйлера. Суммирование в области (δ, ∞) выполняется достаточно просто, после чего при произвольном $\delta > 0$

$$\begin{aligned} A_2 = -(R\mu)^3 & \left[-\frac{3}{8} (\gamma + \ln \mu^2 \delta) - \frac{3}{8} E_1(\mu^2 \delta) + \frac{\pi}{8\sqrt{3}} - \right. \\ & - \sum_{n=1}^{\infty} \frac{(-\mu^2 \delta)^n}{(2n+3)nn!} - \frac{1}{24} \sum_{n=1}^{\infty} \frac{(-1)^n}{nn!} \left(\frac{3}{4} \mu^2 \delta \right)^n \times \\ & \times \left(1 + \frac{n}{5} + \frac{n(n-1)}{3 \cdot 7 \cdot 2!} + \frac{n(n-1)(n-2)}{3^2 \cdot 9 \cdot 3!} + \dots \right) \left. \right] + \\ & + \frac{9}{8} (R\mu)^2 e^{-\mu^2 \delta} + O(R\mu), \end{aligned} \quad (6.65)$$

где $E_k(z) = \int_1^{\infty} dx e^{-zx}/x^k$ — интегральная показательная функция.

Подставим выражения (6.64), (6.65) в формулу (6.62); тогда при $\delta\mu^2 \ll 1$ получим

$$\ln \left(\frac{\text{Det } D^{(5)}}{\text{Det } D_0^{(5)}} \right) = -\frac{9}{16} (R\mu)^3 \left[\frac{1}{3} \ln \left(\frac{\mu^2}{M^2} \right) + \frac{2\pi}{9\sqrt{3}} - \frac{2}{3} E_1(\mu^2 \delta) - \right.$$

$$\begin{aligned}
& -\frac{1}{\mu^2\delta} - \gamma - \ln \mu^2\delta - \sum_{n=1}^{\infty} b_n(\mu^2\delta)^n \left] + \frac{9}{8} e^{-\mu^2\delta} (R\mu)^2 + \right. \\
& \left. + A_c + O(R\mu), \tag{6.66}
\end{aligned}$$

где $b_1 = -1/(1 \cdot 2!)$, $b_2 = -1/(2 \cdot 3!)$, $b_3 = -1/(3 \cdot 4!)$ (вычисление следующих коэффициентов b_n принципиальных трудностей не вызывает, хотя и достаточно утомительно). Таким образом, несколько первых членов ряда $\sum_{n=1}^{\infty} b_n(\mu^2\delta)^n$ представляют собой разложение интегральной показательной функции $E_2(z)$ в ряд:

$$E_2(z) = 1 + z \left[-1 + \gamma + \ln z + \sum_{n=1}^{\infty} \frac{(-1)^n}{n(n+1)!} z^n \right].$$

Поэтому в дальнейшем мы будем предполагать [19], что

$$\sum_{n=1}^{\infty} b_n(\mu^2\delta)^n = 1 - \gamma - \ln \mu^2\delta + [E_2(\mu^2\delta) - 1] / \mu^2\delta,$$

и тогда

$$\begin{aligned}
\ln \left(\frac{\text{Det } D^{(5)}}{\text{Det } D_0^{(5)}} \right) &= -\frac{9}{16} (R\mu)^3 \left[\frac{1}{3} \ln \left(\frac{\mu^2}{M^2} \right) + \frac{2\pi}{9\sqrt{3}} - 1 - \right. \\
&\left. - \frac{2}{3} E_1(\mu^2\delta) - \frac{E_2(\mu^2\delta)}{\mu^2\delta} \right] + \frac{9}{8} e^{-\mu^2\delta} (R\mu)^2 + A_c + O(R\mu). \tag{6.67}
\end{aligned}$$

Хотя мы работаем с точностью до $O(R\mu)$, слагаемые $\sim \ln(R\mu)$ также могут быть вычислены без более точного решения уравнения движения (6.53) и более точного определения собственных значений λ_{0l} и λ_{1l} . Для этого необходимо заметить, что логарифмические факторы могут возникнуть в формуле (6.62) при интегрировании по области (δ, ∞) и суммировании по уровням λ_{0l} , так как

$$\int_{\delta}^{\infty} \frac{dt}{t} \sum_{l=2}^{l_0} \exp(-\lambda_{0l}t) = \sum_{l=2}^{l_0} (l+1)^2 E_1(\lambda_{0l}\delta),$$

причем, как легко видеть, вклад в слагаемые $\sim \ln R$ дают только низколежащие уровни ($2 \leq l < \Delta$, $\Delta \gg 1$). После несложных вычислений легко получить, что логарифмическая поправка к формуле (6.67) составляет $10 \ln(R\mu)$.

Продолжим аналитически выражение (6.67) на область $\mu\delta > 1$. Поскольку величина функционального детерминанта (6.67) не должна зависеть от δ , получаем

$$\ln \left(\frac{\text{Det } D^{(5)}}{\text{Det } D_0^{(5)}} \right) = -\frac{9}{16}(R\mu)^3 \left[\frac{1}{3} \ln \left(\frac{\mu^2}{M^2} \right) + \frac{2\pi}{9\sqrt{3}} - 1 \right] + 10 \ln(R\mu) + O(R\mu), \quad (6.68)$$

где мы использовали то, что $E_1(\mu^2\delta)$, $E_2(\mu^2\delta)$ и A_c являются убывающими функциями при $\delta \rightarrow \infty$.

Для того чтобы получить окончательный ответ для вероятности распада метастабильного вакуума, нам необходимо найти квантовую поправку к действию

$$\begin{aligned} \delta B &= \int d^n x \left[\frac{1}{2} (Z_3 - 1) (\partial\Phi)^2 + \frac{\Delta\lambda}{8} (\Phi^2 - a^2)^2 - \frac{\lambda a}{2} \Delta a (\Phi^2 - a^2) \right] = \\ &= 2B \left[(Z_3 - 1) + \frac{\Delta\lambda}{\lambda_R} + \frac{9}{2} \frac{\Delta a}{a_R} \right], \end{aligned} \quad (6.69)$$

где Z_3 — константа перенормировки волновой функции, λ_R и a_R — перенормированные параметры, $\Delta\lambda$ и Δa — соответствующие контрчлены. Используя то, что коэффициент a_3 в формуле (6.47) для функции Грина ведет к перенормировке волновой функции, находим

$$Z_3 = 1 - \frac{3\lambda_R}{64\pi^2} \left(1 + \frac{\varepsilon}{2\lambda_R a_R^4} \right) + O(\varepsilon^2). \quad (6.70)$$

Контрчлены же $\Delta\lambda$ и Δa могут быть зафиксированы условиями

$$U''_{\text{eff}}(\Phi_+) = (\mu_+^2)_R, \quad U'''_{\text{eff}}(\Phi_+) = 3\lambda_R, \quad (6.71)$$

где $U_{\text{eff}}(\Phi_c)$ — эффективный потенциал:

$$U_{\text{eff}}(\Phi) = U(\Phi) + \ln \det[-\partial^2 + U''(\Phi)] / (2 \int dx); \quad (6.72)$$

Φ_c — постоянное поле произвольной величины. Детерминант в выражении (6.72) вычисляется элементарно по формулам (6.43), (6.49), и эффективный потенциал принимает вид

$$U_{\text{eff}}(\Phi) = U(\Phi) + \frac{1}{64\pi^2}(U'')^2 \left(\ln \frac{U''}{M^2} - \frac{3}{2} \right). \quad (6.73)$$

Тогда для δB имеем

$$\delta B = B \frac{3\lambda_R}{32\pi^2} \left[3 \ln \left(\frac{\mu^2}{M^2} \right) - 109 + \frac{260}{R\mu} \right]. \quad (6.74)$$

Подставляя (6.55), (6.68), (6.74) в выражение (6.61) для вероятности распада метастабильного вакуума, окончательно получаем:

$$\Gamma \simeq V \varepsilon^2 R^4 \exp \left\{ -B + \frac{1}{16} \left(50 + \frac{\pi}{\sqrt{3}} \right) (R\mu)^3 - \frac{130}{16} (R\mu)^2 + O(R\mu) \right\}. \quad (6.75)$$

Отметим, что полученный нами предэкспоненциальный фактор совпадает с результатом работы [155].

Полученные квантовые поправки к действию B несомненно меньше самого действия, которое пропорционально $(R\mu)^3/\lambda$. Тем не менее, в рассматриваемом пределе $R\mu \gg 1$ эти поправки приводят к значительному эффекту для вероятности распада вакуума и должны учитываться при изучении динамики фазовых переходов. В работе [21] получено более точное выражение для вероятности распада

$$\frac{\Gamma}{V} \cong \frac{1}{9} \left(\frac{R}{\lambda} \right)^2 \exp \left\{ -B + \frac{1}{16} [(50 + \pi/\sqrt{3})(R\mu)^3 - 130(R\mu)^2 + (68\pi^2 - 239)R - (385 + 153\pi^2 + 4\zeta(3)/\pi^2)] \right\}. \quad (6.76)$$

6.4.2. Подавление распада вакуума виртуальными частицами

Рассмотрим фазовый переход первого рода, когда скалярное поле, совершающее фазовый переход, окружено облаком виртуальных частиц другого сорта. Как будет показано ниже, такое облако уменьшает вероятность подбарьерного перехода даже при нулевой температуре, причем подавление может стать достаточно сильным, чтобы полностью изменить картину перехода. Физически это достаточно очевидно, поскольку облаку виртуальных частиц требуется время на перестройку своего состояния.

Рассмотрим систему с действием вида (6.19) в евклидовом пространстве. Действуя в полной аналогии с предыдущим, получим эффективное евклидово действие для основного поля φ :

$$S_{\text{eff}}^E = \int d^4x \left[-\frac{1}{2} \varphi_{,\mu} \varphi^{,\mu} - V(\varphi) \right] + \frac{\varkappa^2}{2} \int d^4x \int d^4x' u(\varphi(x)) G_E(x, x') u(\varphi(x')). \quad (6.77)$$

(Здесь и далее гравитационные эффекты не учитываются). Для дальнейшего удобно преобразовать это выражение к виду

$$S_E = \int d^4x \left[\frac{1}{2} (\partial\varphi)^2 + V(\varphi) - \frac{\alpha^2}{2} u(\varphi)^2 \right] + \frac{\alpha^2}{2} \int d^4x \frac{\partial u(\varphi(x))}{\partial x_\mu} \frac{\partial u(\varphi(x))}{\partial x_\mu}. \quad (6.78)$$

Третье слагаемое в этом выражении должно быть включено в перенормировку потенциала,

$$V_{\text{ren}}(\varphi) = V(\varphi) - \frac{\alpha^2}{2} u(\varphi(x))^2, \quad \alpha := \varkappa/m_\chi,$$

а последнее слагаемое может оказывать существенное влияние на поведение системы. Именно это нелокальное слагаемое и отражает влияние облака виртуальных частиц сорта χ на динамику

поля φ . Здесь учтено также, что для функции Грина $G_E(x - x')$ выполняется равенство

$$\int dx G_E(x - x') = 1/m_\chi^2. \quad (6.79)$$

Пусть потенциал V_{ren} имеет два минимума и начальное значение поля φ_F отвечает метастабильному состоянию. В этом случае распад метастабильного состояния происходит путем образования и расширения пузырьков истинного вакуума φ_T . Этот процесс описывается $O(4)$ -инвариантным решением $\varphi_B(r)$ классического уравнения поля φ с граничными условиями $\varphi_B(0) = \varphi_T$, $\varphi_B(\infty) = \varphi_F$. Плотность вероятности распада вакуума была впервые получена в статье [120]:

$$\Gamma/V = \left(\frac{S_E(\varphi_B)}{2\pi} \right)^2 \left| \frac{\text{Det}' \hat{D}(\varphi_B)}{\text{Det} \hat{D}(\varphi_F)} \right|^{-1/2} e^{-S_E(\varphi_B)}, \quad (6.80)$$

где ядро K оператора $\hat{D}(\varphi)$ есть

$$K(x, y) \equiv \frac{\delta^2 S_E(\varphi)}{\delta\varphi(x) \delta\varphi(y)}. \quad (6.81)$$

Вычисление нормировочного детерминанта в знаменателе обычно не представляет трудностей. Поскольку $\varphi_F = \text{const}$, все сводится к вычислению хорошо известного детерминанта типа $\text{Det}(-\partial^2 + \text{const})$. В рассматриваемом же случае эффективное действие нелокально, и вычисление нормировочного детерминанта представляет проблему. Если ввести определения

$$\begin{aligned} \Omega^2 &\equiv V''(\varphi_F) + \alpha^2 \left(\frac{du}{d\varphi_F} \right)^2, \\ M^4 &\equiv \varkappa^2 \left(\frac{du}{d\varphi_F} \right)^2, \end{aligned} \quad (6.82)$$

то проблема сводится к вычислению детерминанта оператора $\hat{D}(\varphi_F)$, ядро которого имеет вид

$$K_F(x, y) = \{\delta(x - y)(-\partial_x^2 + \Omega^2) - M^4 G_E(x - y)\}. \quad (6.83)$$

Это ядро не диагонально, но ввиду его специфики приводится к диагональному виду, после чего легко вычисляется (см. [202]). Введем оператор $\hat{G}_E^{-1} \equiv (-\partial_x^2 + m_\chi^2)$. Тогда

$$\begin{aligned} \text{Det } \hat{D}(\varphi_F) &= \text{Det } G_E^{-1} \hat{D}(\varphi_F) / \text{Det } G_E^{-1} = \\ &= \text{Det}[(-\partial_x^2 + \Omega^2)(-\partial_x^2 + m_\chi^2) - M^4] / \text{Det } G_E^{-1} = \\ &= \text{Det}(-\partial_x^2 + \omega_1^2) \cdot \text{Det}(-\partial_x^2 + \omega_2^2) / \text{Det } G_E^{-1}, \end{aligned} \quad (6.84)$$

где введены параметры

$$\omega_{1,2}^2 = \frac{1}{2} \left[\Omega^2 + m_\chi^2 - \sqrt{(\Omega^2 - m_\chi^2)^2 + 4M^4} \right]. \quad (6.85)$$

Задача вычисления нормировочного множителя в предэкспоненте выражения (6.80) свелась к вычислению известных детерминантов от операторов осцилляторного типа.

На самом деле квантовые флуктуации при высоких энергиях приводят к серьезным изменениям в эффективном действии, так что уже главный, экспоненциальный множитель оказывает существенное влияние на скорость распада вакуума. Покажем это на примере потенциала с двумя минимумами — локальным и абсолютным

$$V_{\text{ren}} = \frac{\lambda}{8}(\varphi^2 - a^2)^2 + \frac{\varepsilon}{2a}(\varphi - a). \quad (6.86)$$

Инстантонное решение евклидовского уравнения движения можно параметризовать следующим образом:

$$\varphi(x) = \varphi_{\text{inst}}(r) = A \text{th} \left(\frac{M}{2}(r - R) \right) - B, \quad (6.87)$$

где $r^2 \equiv \sum_{\alpha=1}^4 x_\alpha^2$, а параметры R и M — произвольные величины, определяемые посредством минимизации действия. Параметры A и B находятся из граничных условий

$$\varphi_{\text{inst}}(r \rightarrow \infty) = \varphi_+, \quad \varphi_{\text{inst}}(r \rightarrow 0) = \varphi_-. \quad (6.88)$$

Здесь φ_+ отвечает правому, локальному минимуму потенциала (6.86) (ложный вакуум), а φ_- — левому, абсолютному минимуму (истинный вакуум).

Действие (6.78) записано в пренебрежении высшими производными по полю φ . Это приближение справедливо при условии $\partial\varphi_B/m_\chi\varphi_B \ll 1$ (m_χ — масса χ -частиц, образующих облако). На инстантонной траектории производная $\partial\varphi_B$ пропорциональна массе кванта основного поля m_φ . Таким образом, критерием справедливости приближения является условие

$$m_\varphi/m_\chi \ll 1. \quad (6.89)$$

На самом деле область перехода от ложного вакуума к истинному значительно уширяется под влиянием запаздывающих эффектов, обусловленных облаком виртуальных χ -частиц. Поэтому вышеприведенное условие заведомо выполняется и даже является чересчур жестким.

Для проведения численного анализа запишем классическое уравнение, решением которого является $O(4)$ -симметричная инстантонная траектория

$$\partial_r^2\varphi + \frac{3}{r}\partial_r\varphi - V'_{\text{ren}}(\varphi) + \frac{\alpha^2}{m_\chi^2}u'(\varphi) \left[\frac{3}{r}u'(\varphi)\partial_r\varphi + u''(\varphi)(\partial_r\varphi)^2 + u'(\varphi)\partial_r^2\varphi \right] = 0. \quad (6.90)$$

Решение этого уравнения ищется в виде (6.87) с неизвестными параметрами M и R . Параметры A и B определяются граничными условиями.

Результаты численных расчетов показаны на рис. 6.3. Ось абсцисс выбрана совпадающей с известным результатом [120, 121], полученным без учета влияния облака виртуальных частиц. Из графика видно, что действие может возрастать на порядки и, следовательно, вероятность фазового перехода оказывается экспоненциально подавленной по сравнению с обычными предсказаниями.

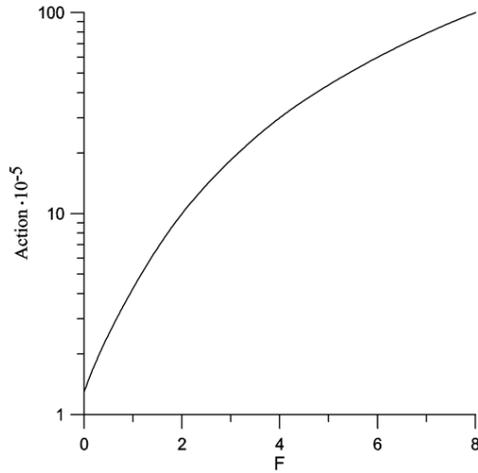


Рис. 6.3. Зависимость евклидова эффективного действия от параметра $F = \varkappa^2/m_\chi^4$. Значения параметров: $a = 1$; $\lambda = 0.1$; $\varepsilon = 0.01$ в единицах $m_\chi = 1$.

Для более конкретной оценки предположим, что действие на инстантонной траектории без учета взаимодействия с дополнительным полем есть S_0 , а с его учетом — ρS_0 . Очевидно, с учетом этого взаимодействия время жизни метастабильного состояния возрастает в $q \approx \exp[(\rho - 1)S_0]$ раз. Используя график 6.3, выберем промежуточное значение $\rho = 10$. Вероятность подбарьерного перехода обычно вычисляется в квазиклассическом приближении, когда $S_0 \gg 1$. Выберем минимально возможное разумное значение действия $S_0 = 10$, что позволяет получить оценку на величину $q = e^{90} \sim 10^{40}$. Такое резкое возрастание времени жизни метастабильного состояния способно изменить выводы, получаемые на основе конкретной модели. Обычно действие много больше выбранного значения, и подавление оказывается еще гораздо более сильным.

Очевидно, расчет вероятности распада ложного вакуума представляет собой непростую задачу, требующую учета квантовых

поправок первого порядка. Последние могут в корне изменить оценку времени жизни метастабильного состояния и повлиять, например, на скорость образования нашей Вселенной, если это произошло за счет фазового перехода первого рода, или на скорость электрослабого перехода.

Глава 7

Барионная асимметрия Вселенной

Известно, что, несмотря на симметрию между частицами и античастицами, в нашей Вселенной преобладают именно барионы. Это один из важнейших ингредиентов современной космологической модели. В какой момент произошло столь значительное нарушение симметрии? Каков механизм бариосинтеза? Этой проблеме посвящено множество обзоров, см. например [35, 131]), где изучаются самые разные варианты. Здесь рассмотрен механизм, предсказывающий возможное существование крупномасштабных пространственных областей антиматерии. Принципиальная возможность образования областей антиматерии обсуждалась и ранее [117].

Поскольку острова (области) антиматерии пока экспериментально не обнаружены, число антибарионов в целом по Вселенной должно быть много меньше числа барионов. По той же причине острова антиматерии не могут быть слишком большими — иначе мы наблюдали бы фотоны от аннигиляции по краям областей, заполненных антибарионами.

С другой стороны, если размер областей недостаточно велик, та же самая аннигиляция на границе приводила бы к их

быстрому испарению. Именно это и имеет место в большинстве моделей бариосинтеза. Обычно процессы спонтанного нарушения CP-симметрии проявляются в ходе фазовых переходов первого и второго родов, имевших место в ранней Вселенной. В этом случае любые области антиматерии оказываются слишком малыми, чтобы сохраниться до настоящего времени [131]. Как было показано в [177], граница, отделяющая область антиматерии от барионного окружения, сдвигается за счет аннигиляции не более чем на 0.5 пк к моменту окончания радиационно-доминированной эпохи. Чтобы сохраниться до настоящего времени, области антиматерии должны быть достаточно большими. Это означает, что будущие области антиматерии должны начать свое формирование на инфляционной стадии, поскольку основным свойством этой стадии является возможность эффективно "растягивать" пространственные размеры.

7.1. Механизм спонтанного бариогенезиса

Существует большое количество механизмов, приводящих к барионной асимметрии Вселенной, несмотря на симметрию законов. Мы будем базироваться на одном из них [131, 132], где инфляционный период играет существенную роль. Именно, предположим, что, кроме инфлатона, существует комплексное скалярное поле χ , обладающее барионным зарядом.

Предположим, что лагранжиан поля χ , взаимодействующего с тяжелыми кварками Q и лептонами L , имеет вид

$$L = -\partial_\mu \chi^* \partial^\mu \chi - V(\chi) + i\bar{Q}\gamma^\mu \partial_\mu Q + i\bar{L}\gamma^\mu \partial_\mu L - m_Q \bar{Q}Q - m_L \bar{L}L + (g\chi \bar{Q}L + \text{h.c.}). \quad (7.1)$$

Предполагается, что тяжелые кварки Q и лептоны L связаны с обычными кварками и лептонами полей материи. Поля χ и Q обладают барионным числом, но не лептонным. Поле L , напротив, обладает лишь лептонным зарядом. Следовательно, слагаемое в лагранжиане (7.1), описывающее их взаимодействие,

приводит к несохранению лептонного числа [132]. Сохранение барионного числа есть следствие $U(1)$ симметрии лагранжиана относительно преобразований

$$\chi \rightarrow \exp(i\beta)\chi, \quad Q \rightarrow \exp(i\beta)Q, \quad L \rightarrow L. \quad (7.2)$$

Выберем потенциал в первом приближении в виде

$$V(\chi) = -m_\chi^2 \chi^* \chi + \lambda_\chi (\chi^* \chi)^2 + V_0, \quad (7.3)$$

где константа V_0 добавлена для того, чтобы минимум потенциала равнялся нулю. Потенциал (7.3) — “сомбреро” — представляет собой первое, основное приближение. В дальнейшем нам понадобятся квантовые поправки к потенциалу.

Поле χ может быть представлено в виде

$$\chi(\vartheta) = \frac{f}{\sqrt{2}} \exp\left(\frac{i\vartheta}{f}\right). \quad (7.4)$$

Нарушение $U(1)$ симметрии означает, что радиальная компонента поля χ принимает фиксированное значение

$$f = m_\chi / \sqrt{\lambda_\chi},$$

и переменная ϑ в уравнении (7.4) теперь имеет смысл безмассового скалярного поля, поскольку $V(\vartheta) = \text{const}$. Минимум потенциала расположен на окружности радиуса f . Далее работаем с безразмерной переменной $\theta = \vartheta/f$.

В результате нарушения симметрии фаза β поля χ приобретает случайное значение θ , и эффективный лагранжиан имеет вид

$$L = -\frac{f^2}{2} \partial_\mu \theta \partial^\mu \theta + i\bar{Q}\gamma^\mu \partial_\mu Q + i\bar{L}\gamma^\mu \partial_\mu L - m_Q \bar{Q}Q - m_L \bar{L}L + \left(\frac{g}{\sqrt{2}} f \bar{Q}L + \text{h.c.}\right) + \partial_\mu \theta \bar{Q}\gamma^\mu Q. \quad (7.5)$$

Квантовые поправки приводят к изменению формы потенциала, так что его симметрия (7.2) нарушается и появляется небольшой наклон “сомбреро”

$$V(\theta) = \Lambda^4 (1 - \cos \theta); \quad \Lambda \ll f. \quad (7.6)$$

Этот потенциал имеет счетное множество минимумов при $\theta = 2\pi N$, $N = 0, 1, 2, \dots$. Отметим, что потенциал (7.6) является удобной аппроксимацией более сложного выражения, см., например, [39].

Для бариогенезиса важным параметром является наклон потенциала (7.6) вблизи минимума, который фиксирует массу поля θ :

$$m_\theta^2 = \frac{\Lambda^4}{f^2}. \quad (7.7)$$

Поскольку наклон потенциала обусловлен квантовыми поправками, то естественно предположить малость параметра Λ . Это означает, что вклад в плотность энергии от поля θ мал по сравнению с полной плотностью энергии на инфляционной стадии. По этой же причине поле θ является эффективно безмассовым, так как

$$m_\theta \ll H, \quad (7.8)$$

где H — параметр Хаббла во время инфляции. По окончании инфляции, когда условие (7.8) нарушается, поле θ осциллирует около минимума потенциала (7.6). При этом энергия $\rho_\theta \simeq \theta_i^2 m_\theta^2 f^2$ поля θ переходит в энергию рождающихся барионов и антибарионов [118, 132]. Знак барионного заряда в конечном состоянии зависит от начального значения фазы на момент нарушения симметрии.

Оценим число барионов и антибарионов, рождающихся при осцилляциях фазы θ с произвольной начальной фазой θ_i . Выражение для концентрации образовавшихся барионов $n_B(\bar{B})$ в пределе малой фазы θ_i имеет вид [132]

$$n(Q, L) = \frac{1}{V} \sum_{s_Q, s_L} \int \frac{d^3 p}{(2\pi)^3 2p^0} \frac{d^3 q}{(2\pi)^3 2q^0} |\langle Q(p, s_Q) L(q, s_L) | 0 \rangle|^2. \quad (7.9)$$

После стандартных выкладок с использованием последнего слагаемого в (7.1) в качестве лагранжиана взаимодействия получаем плотность рожденных барионов

$$n_{B(\bar{B})} = \frac{g^2}{\pi^2} \int_{m_Q+m_L}^{\infty} \omega d\omega \left| \int_{-\infty}^{\infty} dt \chi(t) e^{\pm 2i\omega t} \right|^2, \quad (7.10)$$

что справедливо в случае $\chi(t \rightarrow -\infty) = \chi(t \rightarrow +\infty) = 0$.

Для более общего случая $\chi(t \rightarrow -\infty) \neq 0$, $\chi(t \rightarrow +\infty) = 0$ ответ можно получить, проинтегрировав по частям выражение (7.10), что приводит к следующей формуле:

$$N_{B(\bar{B})} = \frac{g^2}{4\pi^2} \Omega_{\theta_i} \int d\omega \left| \int_{-\infty}^{\infty} d\tau \dot{\chi}(\tau) e^{\pm 2i\omega\tau} \right|^2, \quad (7.11)$$

где Ω_{θ_i} есть объем, содержащий начальную фазу θ_i . При получении формулы (7.11) поверхностный член равен нулю при $t = \infty$ благодаря фейнмановским условиям на излучение поля χ при $t = -\infty$.

Приближенно численный результат можно получить, полагая зависимость фазы от времени в виде

$$\theta(t) \approx \theta_i(1 - m_\theta t) \quad (7.12)$$

в течение начального периода осцилляций. Подставляя (7.12) и (7.4) в (7.11), после несложных вычислений получим:

$$N_{B(\bar{B})} \approx \frac{g^2 f^2 m_\theta}{8\pi^2} \Omega_{\theta_i} k(\theta_i), \quad k(\theta_i) = \theta_i^2 \int_{\mp \frac{\theta_i}{2}}^{\infty} d\omega \frac{\sin^2 \omega}{\omega^2}, \quad (7.13)$$

где знак на нижнем пределе интеграла соответствует барионному либо антибарионному избытку.

Для пространственно-однородного поля $\chi = (f\sqrt{2}) e^{i\theta}$ имеем следующую формулу для образовавшегося барионного заряда:

$$Q = i[\chi^* d\chi/dt - (d\chi^*/dt)\chi] = -f d\theta/dt. \quad (7.14)$$

Очевидно, барионный заряд в некоторой области $Q > 0$, если $\theta > 0$ в процессе классического движения к нулю фазы θ . Следовательно, движение по часовой стрелке приводит к избытку барионов, а движение против часовой стрелки приводит к избытку антибарионов.

Во время рехитинга плотность энергии инфлатона преобразуется в плотность энергии релятивистских частиц. При этом предполагается, что распад инфлатона на легкие частицы должен быть довольно быстрым по сравнению с периодом колебаний вокруг минимума (напомним, что из-за наклона потенциала минимум находится в точке $\theta = 0$). Таким образом, имеем $\Gamma_{\text{tot}} \gg m_\theta$. Осцилляции поля θ начинаются, когда трением в среде можно пренебречь, т.е. при $H \approx m_\theta$. Изменение фазы со временем приводит к рождению барионов или антибарионов согласно обсуждению, приведенному выше. Плотность энтропии после термализации есть

$$s = \frac{2\pi^2}{45} g_* T^3, \quad (7.15)$$

где g_* — полное эффективное число безмассовых степеней свободы. Предполагается, что температура превосходит шкалу нарушения электрослабой симметрии. При этой температуре все степени свободы Стандартной Модели находятся в равновесии и $g_* \simeq 106.75$. Температура связана со скоростью расширения следующим образом:

$$T = \sqrt{\frac{M_{\text{Pl}} H}{1.66 g_*^{1/2}}} \approx \frac{\sqrt{M_{\text{Pl}} m_\theta}}{g_*^{1/4}}. \quad (7.16)$$

В последнем выражении (7.16) предполагается, что релаксационные процессы начинаются при выполнении условия $H \approx m_\theta$. Используя формулы (7.13), (7.15) и (7.16), получаем различные выражения для концентрации барионов и антибарионов:

$$\frac{n_{B(\bar{B})}}{s} = \frac{45 g_*^2}{16 \pi^4 g_*^{1/4}} \left(\frac{f}{M_{\text{Pl}}} \right)^{3/2} \frac{f}{\Lambda} k(\theta_i). \quad (7.17)$$

Функция $k(\theta_i)$ включает зависимость амплитуды и знака барионной асимметрии от начальной фазы, различной в разных пространственных областях. Поведение этой функции легко поддается численному анализу с использованием выражения (7.13).

Выражение (7.17) позволяет получить наблюдаемую барионную асимметрию в среднем по Вселенной $n_B/s \approx 3 \cdot 10^{-10}$. В данной модели полагалось, что $f \geq H \simeq 10^{-6} M_{\text{Pl}}$. Естественной величиной константы связи является значение $g \leq 10^{-2}$, что приводит к наблюдаемой барионной асимметрии при разумном соотношении $f/\Lambda \geq 10^5$, [39].

7.2. Крупномасштабные флуктуации барионного заряда

Интересная возможность появляется, если учесть эффект флуктуаций фазы во время инфляции. Для этого подробнее рассмотрим движение фазы вдоль долины $|\chi| = f/\sqrt{2}$, представляющей собой окружность (см. рис. 7.1, 7.2). Пусть фаза $\theta = 0$ соответствует северному полюсу, а $\theta = \pi$ — южному. Минимум потенциала находится в северном полюсе. Используя формулу (7.14), нетрудно показать, что преимущественное производство антибарионов осуществляется при движении поля к северному полюсу против часовой стрелки, что происходит, когда Вселенная изначально образуется со значением поля ϑ именно в этой области (область АВ на рис. 7.1). Производство барионов имеет место в области В, когда поле движется по часовой стрелке.

Описанная выше картина справедлива “в среднем”, когда не учитываются квантовые флуктуации фазы, дающие новый эффект. Рассмотрим его подробнее (см. [188]). Прежде всего отметим, см. гл. 6, что флуктуации (почти) безмассового поля, каковым является фаза θ , имеют гауссово распределение [210, 214,

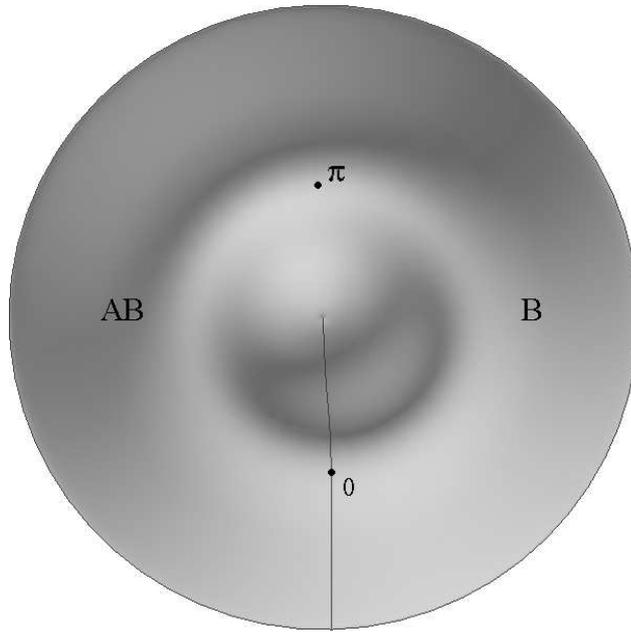


Рис. 7.1. Потенциал “сомбреро”, вид сверху.

282, 283]

$$P(\theta_l, l) = \frac{1}{\sqrt{2\pi}\sigma_l} \exp \left\{ -\frac{(\theta_{60} - \theta_l)^2}{2\sigma_l^2} \right\}. \quad (7.18)$$

Как уже говорилось, Вселенная будет барионно-доминированной в среднем, если начальное значение фазы $\theta = \theta_U$ находится в интервале $[0, \pi]$. Значение θ_U является стартовым для броуновского движения фазы вдоль круговой долины во время инфляции. Средняя флуктуация фазы есть $\delta\theta = H/(2\pi f)$ за каждый е-фолд. Поскольку характерный пространственный размер флуктуации $\delta\theta$, возникающей за один е-фолд, равен H^{-1} , начальный домен, содержащий фазу θ_U , оказывается разделенным на $e^3 \approx 20$ отдельных, причинно не связанных доменов размером H^{-1} . Каждый домен содержит фазу с амплитудой $\theta_{59} \approx \theta_U \pm \delta\theta$.

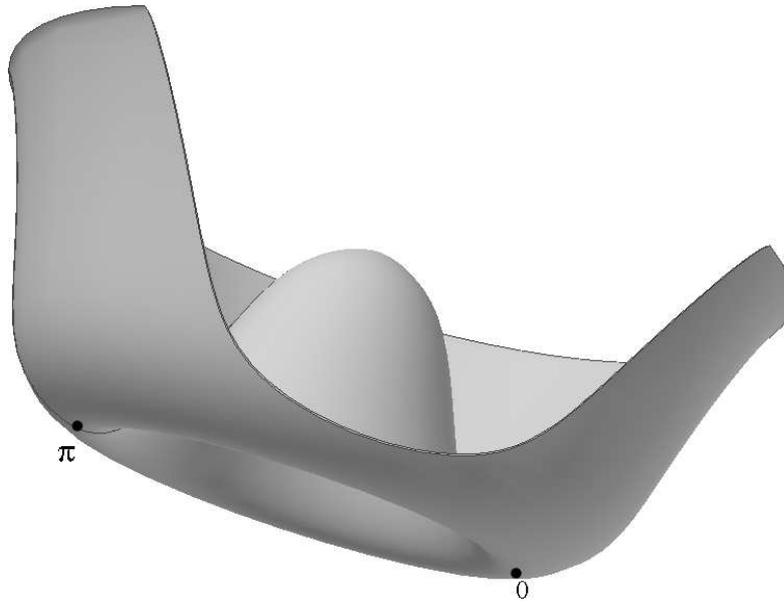


Рис. 7.2. Потенциал “сомbrero”, вид сбоку.

В половине доменов фаза приблизилась к π (южный полюс), в другой половине — к нулю (северный полюс). На следующем е-фолде каждый домен размером H^{-1} точно так же разделится на примерно 20 причинно несвязанных областей. В некоторых из них фаза продвинется еще ближе к Северному полюсу. Нетрудно сообразить, что при некотором е-фолде отдельные пространственные области пересекут нулевую отметку и по окончании инфляции будут находиться в области "В", см. рис. 7.1. Такие области окажутся заполненными преимущественно антибарионами.

Распределение областей по размерам рассмотрено в [188]. Предположим, что к моменту времени t , соответствующему е-фолду N_t до окончания инфляции, объем $V(\bar{\theta}, N_t)$ содержал фазу $\bar{\theta}$. Тогда объем, заполненный фазой $\bar{\theta}$, при е-фолде

$N_{t+\Delta t} = N_t - \Delta N$ может быть найден посредством следующей итерационной процедуры:

$$V(\bar{\theta}, N_{t+\Delta t}) = e^3 V(\bar{\theta}, N_t) + (V_U(N_t) - e^3 V(\bar{\theta}, N_t)) P(\bar{\theta}, N_{t+\Delta t}) \sigma_{N_{t+\Delta t}}, \quad (7.19)$$

где $V_U(N_t) \approx e^{N_t} H^{-1}$ — объем Вселенной при e -фолде N_t . Используя выражение (7.19), можно вычислить пространственное распределение доменов, заполненных наперед заданной фазой поля θ .

Чтобы быть реалистичной, модель бариосинтеза, основанная на данном механизме, должна нетривиальным образом подавлять крупномасштабные флуктуации барионного заряда. В противном случае такие флуктуации давали бы вклад в спектр реликтового излучения и были бы наблюдаемы. Действительно, согласно данному механизму, флуктуации температуры реликтового излучения пропорциональны плотности энергии материи $\delta\rho/\rho \simeq \delta\theta\Omega_B/\Omega_{\text{tot}}$. Здесь $\delta\theta = H/(2\pi f)$, см. выше, Ω_B — относительная плотность энергии барионов, а Ω_{tot} — полная относительная плотность энергии. Следовательно, имеем ограничение $\delta\theta \leq 10^{-3}$, по крайней мере на больших масштабах. Но в этом случае за 60 e -фолдов в течение инфляции фаза продвинется к нулю не более чем на угол $\Delta\theta \sim 60 \cdot 10^{-3} \sim 0.1$ радиана, и переход через нулевое значение маловероятен.

Для того чтобы подавить крупномасштабные флуктуации, и только их, введем взаимодействие поля χ с инфлатоном φ , с некоторой константой связи g . Тогда потенциал примет вид

$$V(\varphi, \chi) = \lambda(|\chi|^2 - f^2/2)^2 - g|\chi|^2(\varphi - cM_{\text{Pl}})^2, \quad (7.20)$$

где λ , g и c есть параметры потенциала. Этот потенциал также имеет форму “сомбреро” с минимумом при

$$|\chi| \equiv f_{\text{eff}}(\varphi) = \sqrt{f^2 + \frac{g}{\lambda}(\varphi - cM_{\text{Pl}})^2}. \quad (7.21)$$

Очевидно, что теперь положение минимума не есть константа. Напротив, оно сильно зависит от классического поля φ , меняющегося в ходе инфляции

$$\varphi(t) = \varphi_U - \frac{m_\varphi M_{\text{Pl}}}{2\sqrt{3}\pi} t \quad (7.22)$$

для квадратичного потенциала $U(\varphi) = m_\varphi^2 \varphi^2 / 2$. Напомним, что амплитуда флуктуаций фазы θ поля χ обратно пропорционально величине $f_{\text{eff}}(\varphi)$. Нетрудно получить зависимость величины инфлатона от числа е-фолдов N :

$$\varphi = \varphi_N = \varphi_U - \frac{M_{\text{Pl}}}{2\sqrt{3}\pi} N. \quad (7.23)$$

Здесь учтено, что $m_\varphi \cong H$ и $N = Ht$. Таким образом, через N е-фолдов после образования Вселенной эффективный масштаб f_{eff} (7.21) выглядит следующим образом:

$$f_{\text{eff}}(N) = f \sqrt{1 + \frac{g}{\lambda} \frac{M_{\text{Pl}}^2}{f^2} \left[\left(\frac{\varphi_U}{M_{\text{pl}}} - c \right) - \frac{N}{2\sqrt{3}\pi} \right]^2}. \quad (7.24)$$

Обозначая $(\varphi_U/M_{\text{pl}} - c) \equiv N_f/(2\sqrt{3}\pi)$, приходим к окончательному выражению вида

$$f_{\text{eff}}(N) = f \sqrt{1 + \frac{g}{12\pi\lambda} \frac{M_{\text{Pl}}^2}{f^2} (N_f - N)^2}. \quad (7.25)$$

Это выражение обладает важным свойством. Действительно, имеется изначально большой параметр $M_{\text{Pl}}^2/f^2 \sim 10^{10}$, и функция $f_{\text{eff}}(N)$ имеет резкий минимум при $f_{\text{eff}}(N_f) = f$ при разумном соотношении между параметрами. Отсюда сразу следует, что амплитуда флуктуации фазы θ

$$\langle \delta\theta \rangle = \frac{H}{2\pi f_{\text{eff}}(N)} \quad (7.26)$$

резко возрастает в окрестности е-фолда $N = N_f$.

Оценим возможный интервал параметров g и λ . Первое неравенство следует из того, что флуктуации должны подавляться достаточно сильно, т.е. $f_{\text{eff}} \gg f$, и выражения (7.21)

$$\langle \chi \rangle \sim \sqrt{g/\lambda} M_{\text{Pl}} \gg f. \quad (7.27)$$

Второе неравенство связано с тем, что введенное в потенциал дополнительное слагаемое перенормирует также и массу инфлатона,

$$m_{\text{ren}}^2 = m^2 - g \langle \chi^2 \rangle.$$

Требование малости перенормировки приводит к неравенству $m^2 \gg g \langle \chi^2 \rangle \sim g^2 M_{\text{Pl}}^2 / \lambda$. Таким образом, второе неравенство имеет вид

$$m \gg g M_{\text{Pl}} / \sqrt{\lambda}. \quad (7.28)$$

Выбрав численные значения параметров

$$f = 10^{-5} M_{\text{Pl}}, \quad m = 10^{-6} M_{\text{Pl}}, \quad (7.29)$$

и подставляя их в выражения (7.27), (7.28), приходим к ограничениям на параметры g, λ :

$$\sqrt{g/\lambda} \gg 10^{-5}, \quad g/\sqrt{\lambda} \ll 10^{-6}. \quad (7.30)$$

Эта система неравенств не очень обременительна. Например, если $\lambda \sim 1$, то $10^{-10} \ll g \ll 10^{-6}$. Малость параметров связана с выбором малых параметров в (7.29), обычным для инфляционных моделей. При таких ограничениях на параметры флуктуации фазы имеют резкий пик при некотором значении e -фолда, обозначенного нами как N_f .

7.3. Обсуждение

В этой главе рассмотрен конкретный пример инфляционной модели с неоднородным бариогенезисом — острова антиматерии,

достаточно большие пространственно, появляются естественным образом. Их число, размер и плотность антиматерии внутри них регулируются несколькими параметрами модели.

Отметим еще одну проблему данного сценария нарушения барионной симметрии, как и многих других, в которых используются фазовые переходы. Это образование топологических дефектов. В нашем случае — это возможность появления крупномасштабных доменных стенок [278, 296]. В данном случае области антиматерии оказываются окруженными замкнутыми доменными стенками.

Плотность поверхностной энергии стенки есть

$$\Delta \approx 8f\Lambda^2. \quad (7.31)$$

После завершения инфляционной стадии стенки сжимаются под действием внутреннего натяжения и осциллируют, стремясь к минимуму энергии. Процесс сопровождается излучением квантов поля и гравитационных волн.

Кроме модели бариосинтеза, рассмотренной выше, существует и множество других возможностей. В качестве примера рассмотрим кратко модель [41], [131] с потенциалом вида

$$U(\varphi) = m^2|\varphi|^2 + \frac{1}{2}\lambda_1|\varphi|^4 + \frac{1}{4}\lambda_2(\varphi^4 + \varphi^{*4}).$$

Модель интересна тем, что барионный заряд поля φ не сохраняется благодаря наличию последнего слагаемого в потенциале,

$$\partial_\mu j_B^\mu \equiv i\partial_\mu(\varphi^* \cdot \partial^\mu \varphi - \partial^\mu \varphi^* \cdot \varphi) = i\lambda_2(\varphi^{*4} - \varphi^4).$$

Здесь для простоты предполагается $\lambda_1 = -\lambda_2 \equiv \lambda > 0$. После окончания инфляции поле движется, подчиняясь классическим уравнениям движения

$$\begin{aligned} \ddot{\varphi}_1 + 3H\dot{\varphi}_1 + (m^2 + 4\lambda\varphi_2^2)\varphi_1 &= 0, \\ \ddot{\varphi}_2 + 3H\dot{\varphi}_2 + (m^2 + 4\lambda\varphi_1^2)\varphi_2 &= 0. \end{aligned}$$

Около минимума потенциала $\varphi_1 = \varphi_2 = 0$ поле вращается с почти постоянным моментом, что приводит к появлению ненулевого барионного заряда. Направление вращения определяется начальными условиями на инфляционной стадии. Если начальные условия немного изменить, то после окончания инфляции поле будет вращаться в другую сторону. Но именно на инфляционной стадии начальные условия в различных пространственных областях отличаются друг от друга за счет квантовых флуктуаций. Поэтому в одних областях будут преобладать барионы, а в других — антибарионы. После окончания инфляции происходит взаимоуничтожение областей за счет аннигиляции на границах. Таким образом, суммарный барионный заряд Вселенной определялся начальными условиями.

Глава 8

Крупномасштабная структура

8.1. Космический микроволновый фон

В 1964 году Арно Пензиас и Роберт Уилсон решили провести исследования в области радиоастрономии и спутниковой связи. Для тестирования антенны была выбрана длина волны 7,35 см. Вскоре стало понятно, что антенна постоянно фиксирует дополнительный необъяснимый посторонний шум, от которого никак не удавалось избавиться. Изотропия шума и его постоянство во времени означали, что его источник находится за пределами Солнечной системы. Даже если бы причина крылась в нашей Галактике, то интенсивность излучения изменялась бы из-за вращения Земли вокруг своей оси и вокруг Солнца, изменяющего направление антенны на те или иные участки космоса. Следовательно, шум был внегалактического происхождения. Измеренная на длине волны 7,35 см, интенсивность этого радиосигнала оказалась равной интенсивности излучения абсолютно черного тела с температурой около 3 К. Так было открыто реликтовое излучение.

Что же такое реликтовое излучение? Согласно теории Большого взрыва, Вселенная возникла приблизительно 14 миллиардов лет назад в результате грандиозного “взрыва”, создавшего

пространство и время, всю материю и энергию, которая существует в нашей Вселенной. До возраста приблизительно 300 тысяч лет молодая Вселенная состояла из горячей плазмы элементарных частиц и, конечно, фотонов. Общее расширение Вселенной постепенно охлаждало эту среду, и когда температура упала до значения нескольких тысяч градусов, появились нейтральные атомы. После этой эпохи первоначальное излучение — фотоны — распространялось свободно, так как оно почти не взаимодействует с образовавшимися нейтральными атомами. В ходе космического расширения длина волны этого излучения постоянно увеличивалась. Именно оно и являлось тем самым шумом, впервые обнаруженным Пензиасом и Уилсоном.

Реликтовое излучение имеет спектр, соответствующий температуре $T = 2.73$ К. Исследования реликтового излучения под разными углами выявили удивительную малость флуктуаций температуры — порядка 10^{-5} в относительных единицах. Тем не менее, в этих флуктуациях заложена ценная информация о ранней Вселенной.

Разберемся на качественном уровне с эффектами, приводящими к флуктуациям температуры $\Delta T(\theta, \varphi)$. Для удобства принято разлагать возмущения по сферическим гармоникам

$$\frac{\Delta T(\theta, \varphi)}{T} = \sum_{l,m} a_{lm} Y_{lm}(\theta, \varphi) \quad (8.1)$$

и изучать график величины $l(l+1)C_l$, где $C_l = \langle |a_{lm}|^2 \rangle$.

График содержит важнейшую информацию. Например, сейчас до нас долетают фотоны, которые находились около горизонта l_{rec} на момент рекомбинации. Именно после рекомбинации Вселенная стала прозрачной. (Напомним, что горизонт — это расстояние, на которое улетает световой сигнал в расширяющейся Вселенной от некоторого момента к данному моменту времени). Пока Вселенная была непрозрачной, движение массивных частиц также было затруднено трением за счет взаимодействия с фотонами.

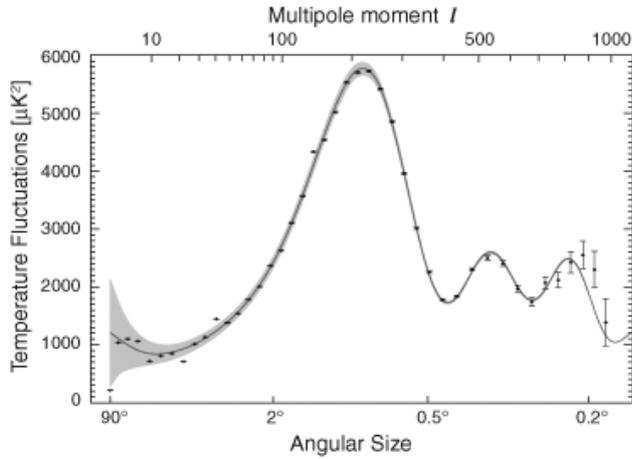


Рис. 8.1. Флуктуации температуры реликтового излучения. См. http://map.gsfc.nasa.gov/media/080999/080999_PowerSpectrum.pdf.

С другой стороны, в тот же период времени имелись флуктуации плотности частиц и гравитационного потенциала. Под действием гравитационного поля флуктуации плотности могли бы увеличиваться, но этого не происходит, поскольку, как говорилось выше, движение заряженных частиц сильно тормозится. В период рекомбинации происходят два события: а) трение исчезает, частицы начинают двигаться к “центру” флуктуации плотности. При этом они ускоряются, что в результате увеличивает температуру области; б) фотоны от этих подогретых областей движутся свободно.

Как влияют размеры области флуктуации плотности на дальнейшие события?

1. Области, сравнимые с размером горизонта на момент рекомбинации $L_{\text{rec}} = 2H_{\text{rec}}^{-1}$, являются причинно связанными областями наибольшего размера на тот момент времени. Именно они обеспечивают наибольшую амплитуду колебаний и, следовательно, наибольший контраст температуры. Фотоны от таких подо-

гретых областей именно сейчас долетают до нас, и на графике мы должны видеть пик температуры. Оценим размер нагретых областей L_0 в настоящий момент:

$$L_0 = \frac{1}{a(t)} L_{\text{rec}} = \frac{T}{T_0} L_{\text{rec}} \approx 300 \text{ Мпк.}$$

Поскольку $H_{\text{rec}}^{-1} \approx \frac{3}{2} t_{\text{rec}}$, $T_{\text{rec}} \approx 3000 \text{ К}$.

Угол, под которым мы видим такую область, есть

$$\theta = \frac{L_0}{L_U} \approx \frac{300}{9000} = 0.03 \text{ рад} \approx 2^\circ.$$

Следовательно, должен наблюдаться пик в области примерно 2° . Этот акустический, или доплеровский, пик отчетливо виден на графике.

2. Области размером много больше размера горизонта L_{rec} в период рекомбинации эволюционируют незначительно. Их средняя температура слабо зависит от размера — на графике отчетливо видно плато (Sachs-Wolfe plateau) при больших углах.

3. Область много меньше l_{rec} . После рекомбинации эти области успели не только сжаться, но и расшириться из-за внутреннего давления. Поэтому они слегка охладились, и мы должны видеть минимум на графике.

8.2. Квантовые флуктуации во время инфляции

8.2.1. Предварительное рассмотрение

Как известно, классическая эволюция систем слегка искажается квантовыми флуктуациями. В пространстве Минковского их роль незначительна во многих случаях, поскольку квантовые поправки пропорциональны малому параметру — постоянной Планка \hbar . Кроме того, согласно принципу неопределенности Гейзенберга, чем больше флуктуация, тем меньше время ее существования. Ситуация кардинально меняется на инфляционной

стадии, которая приближенно описывается в терминах пространства де Ситтера. Основное свойство инфляции состоит в том, что любая пространственная неоднородность растягивается вместе с расширением пространства. Флуктуации поля также представляют собой неоднородности, и их поведение совсем не похоже на поведение флуктуаций в пространстве Минковского. В пространстве де Ситтера их размер быстро растет, превышая размер горизонта. Этот процесс похож на рождение пары частиц в сильном поле. После рождения виртуальные частицы разлетаются, а не аннигилируют, как в пространстве Минковского. Дополнительная энергия создается за счет работы внешнего поля. В нашем случае — гравитационного.

Гипотетическое поле, благодаря которому существует инфляционная стадия, обычно предполагается скалярным. Уравнения Эйнштейна, в которых источником гравитационного поля служит это скалярное поле φ , выглядит следующим образом:

$$\ddot{\varphi} + 3H\dot{\varphi} - H^{-2} e^{-2Ht} \Delta\varphi + V'(\varphi) = 0. \quad (8.2)$$

Уравнение (8.2) следует из уравнения (5.38) и (5.4) после несложных вычислений. Здесь пренебрежено слабым изменением параметра Хаббла $H = \dot{a}/a$ на инфляционной стадии, так что масштабный фактор $a(t) = H^{-1} e^{Ht}$.

Далее, разделим поле на “классическую” — Φ — и “квантовую” — q — части:

$$\varphi(\mathbf{x}, t) = \Phi(\mathbf{x}, t) + q(\mathbf{x}, t). \quad (8.3)$$

Слагаемое $q(\mathbf{x}, t) \equiv \delta\phi(\mathbf{x}, t)$, представляющее собой малые флуктуации, необходимо проквантовать, для чего, как обычно, используем разложение Фурье

$$\delta\varphi(x, t) = \int \frac{d^3k}{(2\pi)^3} \left[a_k f_k(t) e^{ikx} + \text{h.c.} \right],$$

где a_k — оператор уничтожения моды k . Отметим, что в сопутствующей системе координат, используемой нами в данном анализе, пространственная координата x безразмерна, равно как и

волновой вектор k . Из основного уравнения (8.2), полагая $H \simeq \text{const}$, получаем

$$\ddot{f}_k + 3H\dot{f}_k + \frac{k^2}{(H e^{Ht})^2} f_k + V''(\Phi) f_k = 0. \quad (8.4)$$

Во время инфляции поле меняется медленно, (это уравнение справедливо при $V''(\Phi) = \text{const}$) поэтому пренебрежем последним слагаемым, что позволяет найти подходящее решение этого уравнения

$$f_k(t) = \frac{H}{\sqrt{2k^3}} (i + k e^{-Ht}) e^{ik e^{-Ht}}. \quad (8.5)$$

Поведение решения крайне любопытно. При больших волновых числах оно быстро осциллирует, а при малых волновых числах выходит на константу. Граничным волновым числом, разделяющим два режима, является $k_H = a(t)H(t) = e^{Ht}$. Волновое число обратно пропорционально характерному размеру области, в данном случае флуктуации, возникшей в некий момент на инфляционной стадии.

Вырисовывается следующая картина. Квантовые эффекты порождают флуктуацию поля φ , пространственный размер которой, l , естественно, меньше размера горизонта, поскольку флуктуация может возникнуть только в причинно связанной области. Условие $l < 1$ означает $k > 1$ (напомним, что используются безразмерные координаты, и соответствующие им импульсы также безразмерны).

Предположим, что в начальный момент возникла мелкомасштабная полевая конфигурация размера $l_0 \ll 1$ и характерным импульсом $k_0 \gg 1$. Функция f_k , описывающая эту конфигурацию, (8.5), быстро осциллирует. Граничный импульс $k_H(t) = e^{Ht}$ растет экспоненциально быстро, и всегда наступит такой момент времени t_H , что характерный импульс системы окажется сравнимым с граничным, $k_0 \sim k_H(t_H)$. Начиная с этого момента времени, функция f_k перестает осциллировать и выходит на

константу. Очевидно, что значение последней распределено случайным образом. Закон распределения рассмотрен в следующем разделе.

До сих пор мы работали в с сопутствующими координатами. Посмотрим на эти процессы с точки зрения физических расстояний $R_{\text{phys}} = a(t)r = H^{-1} e^{Ht}r$ и соответствующих физических импульсов $P_{\text{phys}} = p/a(t) = H e^{-Ht}k$. Граничный физический импульс, при котором амплитуда некоторой флуктуации “замерзает”, есть $P_{\text{phys}}^H = H e^{-Ht}k_H = H$. При этом характерный размер $L_{\text{phys}} = 1/P_{\text{phys}}^H = 1/H$. Итак, возникшая за счет квантовых эффектов флуктуация на инфляционной стадии развивается следующим образом: величина поля быстро флуктуирует, а ее пространственный размер экспоненциально растет до тех пор, пока не станет равным обратному значению параметра Хаббла.

8.2.2. Подробное рассмотрение

Вернемся к разложению (8.3). Классическая часть Φ ассоциируется с плавным, медленным изменением поля. Часто встречается следующее разложение, основанное на фурье-преобразовании:

$$\begin{aligned}\varphi(\mathbf{x}, t) &= \int \frac{d^3\mathbf{k}}{(2\pi)^{3/2}} \left[a_{\mathbf{k}}(t) e^{-i(\mathbf{k}\mathbf{x})} + a_{\mathbf{k}}^\dagger(t) e^{i(\mathbf{k}\mathbf{x})} \right]; \\ \Phi(\mathbf{x}, t) &= \int_{|\mathbf{k}| < k^*} \frac{d^3\mathbf{k}}{(2\pi)^{3/2}} \left[a_{\mathbf{k}}(t) e^{-i(\mathbf{k}\mathbf{x})} + a_{\mathbf{k}}^\dagger(t) e^{i(\mathbf{k}\mathbf{x})} \right], \\ q(\mathbf{x}, t) &= \int_{|\mathbf{k}| > k^*} \frac{d^3\mathbf{k}}{(2\pi)^{3/2}} \left[a_{\mathbf{k}}(t) e^{-i(\mathbf{k}\mathbf{x})} + a_{\mathbf{k}}^\dagger(t) e^{i(\mathbf{k}\mathbf{x})} \right],\end{aligned}\quad (8.6)$$

где граничный импульс k^* отделяет “быструю” подсистему от “медленной”, классической. Этот выбор неоднозначен и зависит от рассматриваемой проблемы. В пространстве Минковского наиболее удобен выбор базисных функций в виде $a_{\mathbf{k}}(t) \sim e^{ik_0 t}$, являющихся решениями уравнения д’Аламбера

$$\square g(\mathbf{x}, t) = 0.$$

В пространстве де Ситтера также удобно выбрать в качестве базисных функций решения этого уравнения. После преобразования Фурье

$$\tilde{g}_{\mathbf{p}}(t) = \int d^3x e^{i(\mathbf{p}\mathbf{x})} g(\mathbf{x}, t)$$

уравнение приобретает вид

$$\frac{\partial^2 \tilde{g}_{\mathbf{p}}(t)}{\partial t^2} + 3H \frac{\partial \tilde{g}_{\mathbf{p}}(t)}{\partial t} + (H\mathbf{p})^2 e^{-2Ht} \tilde{g}_{\mathbf{p}}(t) = 0, \quad (8.7)$$

где H — параметр Хаббла. Отметим, что импульс \mathbf{p} безразмерен, как и сопутствующая координата \mathbf{x} . Часто используется импульс $\mathbf{P} \equiv H\mathbf{p}$, имеющий правильную размерность. Тогда набор решений уравнения (8.7) выражается через функции Ханкеля

$$H_{3/2}^{(2)}(y) = \left[H_{3/2}^{(1)}(y) \right]^* = -\sqrt{\frac{2}{\pi y}} e^{-iy} \left(1 + \frac{1}{iy} \right).$$

Решение может быть представлено в виде

$$\tilde{g}_{\mathbf{p}}(t) = \frac{\sqrt{\pi}}{2} H \eta^{3/2} \left[c_1(p) H_{3/2}^{(1)}(\eta P) + c_2(p) H_{3/2}^{(2)}(\eta P) \right].$$

Здесь введено так называемое конформное время

$$\eta \equiv -H^{-1} e^{-Ht}. \quad (8.8)$$

Эта величина упрощает уравнения при анализе процессов в пространстве де Ситтера. Две константы определяются из дополнительных условий. Точнее, естественно предположить, что на расстояниях, много меньших размера горизонта, т.е. при $p \rightarrow \infty$, решение должно совпадать с решением в пространстве Минковского. Это возможно при $c_1 = 0$, $c_2 = -1$. Таким образом, набор ортонормированных решений имеет вид [210]

$$\tilde{g}_{\mathbf{p}}(t) = \frac{iH}{P^{3/2}\sqrt{2}} \left(1 + \frac{P}{iH} e^{-Ht} \right) \exp \left(\frac{iP}{H} e^{-Ht} \right). \quad (8.9)$$

Анализ поведения функции $\tilde{g}_{\mathbf{p}}(t)$ указывает на наличие порогового значения импульса

$$P^*(t) \equiv H e^{Ht}. \quad (8.10)$$

В каждый момент времени t решение осциллирует при $P \gg P^*$ и стремится к константе при $P \ll P^*$:

$$\tilde{g}_{\mathbf{p}}(t) \simeq \frac{iH}{P^{3/2}\sqrt{2}}, \quad P \ll P^*. \quad (8.11)$$

Чтобы прояснить физический смысл полученного результата, перейдем к физическим расстояниям, $\mathbf{R}_{\text{phys}} = a(t)\mathbf{r}$ и импульсам. Очевидно, физический импульс \mathbf{P}_{phys} связан с координатным \mathbf{p} следующим образом: $\mathbf{P}_{\text{phys}} = \mathbf{p}/a(t) = \mathbf{P}/(a(t)H)$. На инфляционной стадии масштабный фактор связан с параметром Хаббла, $a(t) \simeq H^{-1}e^{Ht}$. Следовательно, пороговое значение физического импульса P_{phys}^* есть

$$P_{\text{phys}}^* \equiv \frac{P^*}{a(t)H} = H.$$

Характерный размер флуктуации есть $L_{\text{phys}} \sim P_{\text{phys}}^{-1}$. Если этот размер больше размера горизонта H^{-1} , то амплитуда флуктуации стремится к константе.

Проследим за временной зависимостью произвольной флуктуации, определяемой выражением (8.9) и, следовательно, имеющей фиксированный импульс P . Очевидно, что амплитуда флуктуации стремится со временем к константе (8.11). С другой стороны, физический размер L_{phys} флуктуации растет экспоненциально быстро:

$$L_{\text{phys}} \sim P_{\text{phys}}^{-1} = \frac{a(t)H}{P} = \frac{1}{P} e^{Ht}. \quad (8.12)$$

Такое поведение характерно именно для пространства де Ситтера. В пространстве Минковского время жизни квантовой флуктуации — порядка $1/\Delta E$ (ΔE есть энергия флуктуации). В про-

странстве де Ситтера квантовая флуктуация, родившись, увеличивает свой размер экспоненциально быстро согласно выражению (8.12). В то же время ее амплитуда определяется выражением (8.11).

Теперь можно вернуться к проблеме разделения скалярного поля φ на классическую и квантовую части. Квантовая часть — см. (8.6) — может быть записана в виде

$$q(\mathbf{x}, t) \equiv \int \frac{d^3\mathbf{p}}{(2\pi)^{3/2}} W(P, t) \left[\hat{a}_{\mathbf{p}} \tilde{g}_{\mathbf{p}}(t) e^{-i(\mathbf{p}\mathbf{x})} + \hat{a}_{\mathbf{p}}^\dagger \tilde{g}_{\mathbf{p}}^*(t) e^{i(\mathbf{p}\mathbf{x})} \right]. \quad (8.13)$$

Здесь введены операторы рождения и уничтожения $\hat{a}_{\mathbf{p}}^\dagger$, $\hat{a}_{\mathbf{p}}$, согласно стандартному способу квантования полей. Вместо грубого обрезания момента условием $P > P^*$, используем функцию $W(P, t)$ со свойствами $W(P \rightarrow 0, t) \rightarrow 0$, $W(P \rightarrow \infty, t) \rightarrow 1$. Это может быть, например, функция вида

$$W(P, t) = \theta(P - \varepsilon H e^{Ht}), \quad \varepsilon \ll 1. \quad (8.14)$$

Как будет видно позднее, конечный результат не зависит от малого, а в остальном произвольного параметра ε . Подставляя выражения (8.3) и (8.13) в уравнение (8.2), получим

$$\begin{aligned} \frac{\partial \Phi}{\partial t} - \frac{1}{3H} \left[e^{-2Ht} \Delta \Phi - \frac{\partial V(\Phi)}{\partial \Phi} \right] &= y(\mathbf{x}, t), \\ y(\mathbf{x}, t) &\equiv \left(\frac{1}{3H} \frac{\partial^2}{\partial t^2} - \frac{\partial}{\partial t} + \frac{1}{3H} e^{-2Ht} \Delta \right) q(\mathbf{x}, t). \end{aligned} \quad (8.15)$$

В этом уравнении сделано приближение — учитывая медленное движение поля во время инфляции, опущена вторая производная и пренебрежено высшими степенями функции $y(\mathbf{x}, t)$. Уравнение (8.15) описывает динамику поля Φ под действием случайной “силы” y . Последняя предполагается малой, так что мы можем искать решение уравнения в виде [256]

$$\Phi = \Phi_{\text{det}} + \phi. \quad (8.16)$$

Основная часть классического поля Φ_{det} подчиняется уравнению

$$\frac{\partial \Phi_{\text{det}}}{\partial t} - \frac{1}{3H} \left[e^{-2Ht} \Delta \Phi_{\text{det}} - \frac{\partial V(\Phi_{\text{det}})}{\partial \Phi_{\text{det}}} \right] = 0, \quad (8.17)$$

тогда как случайная часть ϕ полностью зависит от квантовых флуктуаций согласно линейному уравнению

$$\frac{\partial \phi}{\partial t} - \frac{1}{3H} \left[e^{-2Ht} \Delta \phi - V''(\Phi_{\text{det}}) \phi \right] = y(\mathbf{x}, t). \quad (8.18)$$

(Рассматривается предел $\Phi_{\text{det}} \gg \phi$, что оправдано, если случайная “сила” $y(\mathbf{x}, t)$ мала). Проводя вычисления случайной “силы”, появившейся благодаря квантовым флуктуациям, получим (см. детали в [256])

$$\begin{aligned} y(\mathbf{x}, t) &\equiv \left(\frac{1}{3H} \frac{\partial^2}{\partial t^2} - \frac{\partial}{\partial t} + \frac{1}{3H} e^{-2Ht} \Delta \right) q(\mathbf{x}, t) = \\ &= \left(\frac{1}{3H} \frac{\partial^2}{\partial t^2} - \frac{\partial}{\partial t} + \frac{1}{3H} e^{-2Ht} \Delta \right) \int \frac{d^3 \mathbf{p}}{(2\pi)^{3/2}} \theta(P - \varepsilon H e^{Ht}) \times \\ &\quad \times \left[\hat{a}_{\mathbf{p}} \tilde{g}_{\mathbf{p}}(t) e^{-i(\mathbf{p}\mathbf{x})} + \hat{a}_{\mathbf{p}}^\dagger \tilde{g}_{\mathbf{p}}^*(t) e^{i(\mathbf{p}\mathbf{x})} \right] \simeq \\ &\simeq \frac{H^3 \varepsilon}{\sqrt{2}} e^{Ht} \int \frac{d^3 p}{(2\pi p)^{3/2}} \delta(P - \varepsilon H e^{Ht}) \left[\hat{a}_{\mathbf{p}} e^{-i(\mathbf{p}\mathbf{x})} - \hat{a}_{\mathbf{p}}^\dagger e^{i(\mathbf{p}\mathbf{x})} \right]. \end{aligned} \quad (8.19)$$

Здесь использовано уравнение (8.7), сильно упрощающее выкладки, и опущена вторая производная, пропорциональная малому параметру ε^2 . Также имеем $P = \sqrt{\mathbf{P}^2} = \sqrt{(\mathbf{p}H)^2}$. Кроме того, в последней строчке принято во внимание приближение (8.11). Справедливость этого предположения вытекает из малости импульса, входящего в аргумент δ -функции.

Рассмотрим коррелятор

$$D(\mathbf{x}, t, \mathbf{x}', t') \equiv \langle 0 | y(\mathbf{x}, t), y(\mathbf{x}', t') | 0 \rangle.$$

Используя выражение (8.19), полученное выше, и свойства операторов рождения и уничтожения a^\dagger, a , можно получить аналитическое выражение для этой величины

$$D(\mathbf{x}, t, \mathbf{x}', t') = \frac{H^3}{4\pi^2} \delta(t - t') \frac{\sin \varepsilon |\mathbf{x} - \mathbf{x}'| e^{Ht}}{\varepsilon |\mathbf{x} - \mathbf{x}'| e^{Ht}},$$

$$\varepsilon \ll 1. \quad (8.20)$$

Случай однородного распределения

Согласно выражению (8.20), коррелятор $D(\mathbf{x}, t, \mathbf{x}', t')$ оказывается резкой функцией расстояния $|\mathbf{x} - \mathbf{x}'|$. По этим же причинам можно пренебречь пространственными производными в формуле (8.18), что сильно упрощает уравнения. Следовательно, есть смысл рассматривать однородное распределение $\Phi = \Phi(t)$, подчиняющееся простому уравнению

$$\frac{\partial \Phi}{\partial t} + \frac{1}{3H} \frac{\partial V(\Phi)}{\partial \Phi} = 0, \quad (8.21)$$

$$\frac{\partial \phi}{\partial t} + \frac{m^2}{3H} \phi = y(t). \quad (8.22)$$

Здесь введено обозначение

$$m^2 \equiv V''(\Phi_{\text{det}}).$$

Для квадратичного потенциала

$$V(\phi) = V_0 + \frac{1}{2} m^2 \phi^2$$

эта величина является константой, а в общем случае произвольного потенциала она почти постоянна во время инфляции. Коррелятор (8.20) случайной функции $y(t)$ в пределе $\varepsilon \ll 1$ может быть аппроксимирован следующим образом:

$$\langle y(t_1) y(t_2) \rangle = D(\mathbf{x}, t, \mathbf{x}, t') = \frac{H}{4\pi^2} \delta(t_1 - t_2).$$

Появление δ -функции в правой части уравнения означает гауссов закон распределения случайной функции $y(t)$ с энергией

$$W(y) = \text{const} \cdot \exp \left[-\frac{1}{2\sigma^2} \int y(t)^2 dt \right], \quad \sigma = \frac{H^{3/2}}{2\pi}.$$

Благодаря линейной связи функций ϕ и $y(t)$, см. (8.22), распределения их вероятностей пропорциональны друг другу. Это означает, что вероятность обнаружить значение функции $\phi(t)$ внутри некоторого малого интервала равно (см. [146])

$$dP(\phi) = \text{const} \cdot \mathcal{D}\phi \exp \left[-\frac{1}{2\sigma^2} \int \left[\frac{\partial\phi}{\partial t} + \frac{m^2}{3H}\phi \right]^2 dt \right].$$

Мера имеет вид $\mathcal{D}\phi \equiv \prod_{i=1}^N d\phi(t_i)$, $N \rightarrow \infty$.

Теперь нетрудно получить вероятность найти значение поля ϕ_2 в момент t_2 при условии, что задана величина ϕ_1 в момент t_1 . Очевидно, необходимо проинтегрировать по всем значениям поля в интервале (t_1, t_2) с исключением граничных точек $\phi_1 \equiv \phi(t_1)$, $\phi_2 \equiv \phi(t_2)$, что приводит к выражению

$$\begin{aligned} dP(\phi_2, t_2; \phi_1, t_1) &= \\ &= \text{const} \cdot d\phi_2 \int_{\phi_1}^{\phi_2} \mathcal{D}\phi \exp \left[-\frac{1}{2\sigma^2} \int_{t_1}^{t_2} \left[\frac{\partial\phi}{\partial t} + \frac{m^2}{3H}\phi \right]^2 dt \right]. \end{aligned} \quad (8.23)$$

Неизвестная константа находится из условия нормировки

$$\int_{-\infty}^{\infty} d\phi_2 P(\phi_2, t_2; \phi_1, t_1) = 1.$$

Функциональный интеграл (8.23) может быть вычислен точно [146] методом перевала, подразумевающим поиск экстремума действия

$$\ddot{\phi} - \mu^2 \phi = 0, \quad \mu \equiv \frac{m^2}{3H}$$

с граничными условиями

$$\phi(t_1) = \phi_1, \quad \phi(t_2) = \phi_2.$$

Решение этого уравнения представляется в виде

$$\begin{aligned} \phi(t) &= A e^{\mu t} + B e^{-\mu t}. \\ A &= \frac{\phi_2 - \phi_1 e^{-\mu T}}{2sh(\mu T)}, \quad B = \frac{-\phi_2 + \phi_1 e^{\mu T}}{2sh(\mu T)}; \quad T = t_2 - t_1. \end{aligned}$$

Подставляя его в интеграл в экспоненте в выражении (8.23), находим требуемую вероятность

$$\begin{aligned} dP(\phi_2, t_1 + T; \phi_1, t_1) &= d\phi_2 \cdot \sqrt{\frac{r}{\pi}} \exp[-r(\phi_2 - \phi_1 e^{-\mu T})^2], \\ r &\equiv \frac{\mu}{\sigma^2} \frac{1}{1 - e^{-2\mu T}}, \\ \mu &= \frac{m^2}{3H} \simeq \text{const}, \quad \sigma = \frac{H^{3/2}}{2\pi} \simeq \text{const}. \end{aligned} \quad (8.24)$$

В пределе безмассового поля приходим к более простой формуле

$$dP(\phi_2, t_1 + T; \phi_1, t_1) = d\phi_2 \sqrt{\frac{2\pi}{H^3 T}} \exp\left[-\frac{2\pi^2}{H^3 T} (\phi_2 - \phi_1)^2\right]. \quad (8.25)$$

Эта формула широко используется в различных инфляционных сценариях, где движение поля должно быть медленным и, следовательно, второй производной потенциала, а значит, и массой можно пренебречь.

Картина эволюции поля выглядит следующим образом. Поле предполагается состоящим из двух компонент — классической (детерминистской) и квантовой, см. выражение (8.16). Классическая часть Φ_{det} подчиняется уравнениям движения (8.21) со случайной “силой” y . Как показано выше, влияние этой силы включено в случайную компоненту ϕ и имеет плотность вероятности (8.24). Нетрудно вычислить среднее $\langle \phi^2 \rangle$ для оценки от-

клонения от классического значения Φ_{det} со временем. Используя выражение для вероятности (8.24), получаем

$$\begin{aligned}\langle \phi(t)^2 \rangle &= \int_{-\infty}^{\infty} \phi^2 dP(\phi, t_1 + t; \phi_1, t_1) = \\ &= \frac{1}{2r} = \frac{\sigma^2}{2\mu} (1 - e^{2\mu t}).\end{aligned}\quad (8.26)$$

Эта формула может быть значительно упрощена в случае безмассового поля. Как уже говорилось, поле можно считать приближенно безмассовым во время инфляции, когда выполняется условие $m \ll H$. Разлагая экспоненту в уравнении (8.26), получаем

$$\sqrt{\langle \phi(t)^2 \rangle} = \sigma = \frac{H}{2\pi} \sqrt{Ht}.\quad (8.27)$$

В терминах e -фолдов, $N \equiv Ht$, и мы получаем формулу, которая в дальнейшем будет широко использоваться:

$$\sqrt{\langle \phi(t)^2 \rangle} = \sigma = \frac{H}{2\pi} \sqrt{N}.\quad (8.28)$$

В частности, можно сделать вывод, что флуктуации с амплитудой $\sim H/2\pi$ формируются на временном интервале $t \sim H^{-1}$ ($N = 1$). Выражение (8.10) указывает на момент времени, когда флуктуация перестает меняться, так что ее пространственный размер становится фиксированным в тот же временной период $t \sim H^{-1}$.

Пространственный размер флуктуации также может быть оценен. Заметим для этого, что коррелятор (8.20) не мал при $t \sim H^{-1}$, если сопутствующее расстояние $|\mathbf{x} - \mathbf{x}'| \lesssim 1$. Это означает, что существенными являются флуктуации с сопутствующим размером порядка единицы. Их размеры в физических координатах растут обычным образом:

$$L_{\text{fluct}} = a(t)|\mathbf{x} - \mathbf{x}'| \approx H^{-1} e^{Ht}\quad (8.29)$$

и равны H^{-1} в момент $t \sim H^{-1}$. Это замечание очень важно для дальнейшего. Кстати, результаты (8.28) и (8.29) могут быть легко воспроизведены даже в пространстве Минковского. Действительно, выбрав лагранжиан для безмассового скалярного поля, проведем оценки по порядку величины для действия

$$S = \int d^4x \sqrt{-g} \frac{1}{2} g^{\mu\nu} \partial_\mu \varphi \partial_\nu \varphi.$$

В результате получим

$$S \sim \frac{(\Delta\varphi)^2}{H^{-2}} H^{-4}.$$

Здесь $\Delta\varphi$ обозначает амплитуду флуктуации, возникшей на время $t \sim H^{-1}$. Также принято во внимание, что размер флуктуации порядка H^{-1} , что верно для безмассового поля, распространяющегося со скоростью света. Вероятность такой флуктуации порядка единицы при условии, что действие $S \sim 1$. Соответственно, получаем оценку

$$\Delta\varphi \sim H,$$

которая по порядку величины совпадает с более корректно полученным результатом (8.28).

Существует и другой путь исследования поведения флуктуаций, основанный на том, что уравнение (8.22) есть не что иное, как уравнение Ланжевена. Тогда сразу можно сделать утверждение, что плотность вероятности удовлетворяет уравнению Фоккера — Планка

$$\frac{\partial P}{\partial t} = \frac{H^3}{8\pi^2} \frac{\partial^2 P}{\partial \phi^2} + \frac{m^2}{3H^2} \frac{\partial^2(\phi P)}{\partial \phi^2}. \quad (8.30)$$

Выражение (8.24) для вероятности, полученное ранее, является решением этого уравнения ($dP = P d\phi$ в наших обозначениях).

8.2.3. Классическая эволюция квантовых флуктуаций

Один из важнейших результатов предыдущего раздела состоит в следующем. На инфляционной стадии флуктуации плотности, созданные квантовыми флуктуациями, быстро увеличивают свой пространственный размер. В некоторый момент он становится больше размера горизонта — флуктуация пересекает горизонт первый раз. К окончанию инфляции, масштаб флуктуации может существенно превосходить горизонт. Амплитуда флуктуации меняется до тех пор, пока масштаб флуктуации не пересечет горизонт второй раз, уже после окончания инфляции. Оказывается, в промежутке между двумя пересечениями горизонта имеет место простое соотношение

$$\frac{\delta\rho}{p+\rho} \simeq \text{const}, \quad (8.31)$$

которое понадобится нам в дальнейшем. Прежде всего докажем это соотношение, используя уравнение (5.10) в виде

$$\frac{d\rho}{p+\rho} = -\frac{3da}{a}.$$

Предположим, что имеется взаимно-однозначное соответствие между плотностью и давлением материи $\rho \leftrightarrow p$. Следовательно, давление есть некоторая функция плотности $p(\rho)$ и можно формально проинтегрировать предыдущее выражение,

$$\int_{\rho(t_0)}^{\rho(t)} \frac{d\rho}{p(\rho)+\rho} = -3 \int_{a(t_0)}^{a(t)} \frac{da}{a}, \quad (8.32)$$

где t_0 есть момент времени на инфляционной стадии, а t — произвольный момент, удовлетворяющий соотношению $t - t_0 \gg 1/H_e$, H_e — параметр Хаббла в конце инфляции. Если квантовая флуктуация создает возмущение плотности энергии $\delta\rho(t_0)$, уравнение (8.32) должно быть записано в виде

$$\int_{\rho(t_0)+\delta\rho(t_0)}^{\rho(t)+\delta\rho(t)} \frac{d\rho}{p(\rho)+\rho} = -3 \int_{a(t_0)+\delta a(t_0)}^{a(t)+\delta a(t)} \frac{da}{a}, \quad (8.33)$$

справедливым для пространственной области, занятой флуктуацией.

Пусть теперь t_0 и t означают моменты первого и второго пересечения горизонта данной флуктуацией. Разложим обе части уравнения (8.33) на сумму интегралов:

$$\begin{aligned} \int_{\rho(t_0)+\delta\rho(t_0)}^{\rho(t_0)} \frac{d\rho}{p(\rho)+\rho} + \int_{\rho(t)}^{\rho(t)+\delta\rho(t)} \frac{d\rho}{p(\rho)+\rho} = \\ = -3 \int_{a(t_0)+\delta a(t_0)}^{a(t_0)} \frac{da}{a} - 3 \int_{a(t)}^{a(t)+\delta a(t)} \frac{da}{a}. \end{aligned}$$

Здесь уравнение (8.32), справедливое для объема, включающего в себя объем, занятый флуктуацией, используется для сокращения основных слагаемых в обеих частях равенства. Оставшиеся интегралы легко оценить, принимая во внимание малость флуктуаций:

$$\left(\frac{\delta\rho}{p+\rho} + 3 \frac{da}{a} \right)_{t_0} \simeq \left(\frac{\delta\rho}{p+\rho} + 3 \frac{da}{a} \right)_t.$$

Масштабный фактор a растет экспоненциально быстро начиная с момента вхождения в инфляционную стадию, так что мы можем пренебречь вторыми слагаемыми в обеих частях уравнения, чтобы получить требуемое равенство (8.31).

Формула (8.31), переписанная в виде

$$\left(\frac{\delta\rho}{p+\rho} \right)_{t_f} \simeq \left(\frac{\delta\rho}{p+\rho} \right)_{t_{in}}, \quad (8.34)$$

может быть использована для получения спектра крупномасштабных флуктуаций. Здесь $t_{in} = t_0$ — время появления флуктуации на инфляционной стадии и пересечения горизонта первый раз, t_f — время второго пересечения горизонта на стадии Фридмана — Робертсона — Уокера. Левая часть этого равенства может быть выражена в терминах скалярного поля (инфлатона) $\varphi = \varphi(t_{in})$. Правая часть равенства особенно проста на стадии

доминантности материи, поскольку давление равно нулю, $p = 0$. Запишем выражения для $\delta\rho$ и $p + \rho$ на инфляционной стадии:

$$\delta\rho = V'(\varphi)\delta\varphi, \quad p + \rho = \dot{\varphi}^2.$$

Теперь величина $\delta\rho/\rho|_{t_f}$, которую и надо оценить, также выражается через инфлатонное поле:

$$\left(\frac{\delta\rho}{\rho}\right)_{t_f} \simeq \left(\frac{V'(\varphi)\delta\varphi}{\dot{\varphi}^2}\right)_{t_{\text{in}}}.$$

В приближении медленного скатывания имеем $\delta\varphi \approx H(\varphi)/2\pi$. Используя уравнение инфлатонного поля, получаем

$$\left(\frac{\delta\rho}{\rho}\right)_{t_f} = \left(\frac{9H(\varphi)^3}{2\pi V'(\varphi)}\right)_{t_{\text{in}}}. \quad (8.35)$$

На инфляционной стадии параметр Хаббла H выражается через потенциал $V(\varphi)$, и мы приходим к формуле для вычисления амплитуды флуктуации плотности

$$\left(\frac{\delta\rho}{\rho}\right)_{t_f} = \frac{9}{5} \left(\frac{4}{3}\right)^{3/2} \frac{m\varphi^2}{M_{\text{Pl}}^3}. \quad (8.36)$$

Флуктуации характеризуются значениями поля φ на момент первого пересечения горизонта. В данную формулу включен для полноты фактор $2/5$, который возникает при учете доминантности материи при втором пересечении горизонта [210]. Для дальнейшего введенный фактор особого значения не имеет.

Другим важным параметром флуктуации является ее характерный размер, l . Он зависит от моментов t_{in} и t_f первого и второго пересечения горизонта соответственно. Оценка размера флуктуации на момент времени t , такой, что $t_{\text{in}} < t < t_f$, может быть получена с использованием нормировки на размер видимой части Вселенной, L_U :

$$l = L_U \exp(N - N_U), \quad (8.37)$$

где N есть число e -фолдов, при котором образуется масштаб l данной флуктуации на инфляционной стадии, L_U означает размер нашей Вселенной, $L_U \approx 10^4 \text{ Мпк} \approx 10^{28} \text{ см}$.

С другой стороны, число e -фолдов выражается через величину инфлатона φ согласно выражению (6.5)

$$N \approx \frac{2\pi}{M_{\text{Pl}}^2} \varphi^2,$$

где потенциал выбран в виде $V(\varphi) = (m/2)\varphi^2$ и предполагается выполненным неравенство $\varphi \equiv \varphi_{\text{in}} \gg \varphi_{\text{f}}$. Подставляя эту формулу и формулу (8.37) в выражение (8.35), приходим к связи между амплитудой флуктуации $(\delta\rho/\rho)_{t_{\text{f}}}$ и ее размером l [205]

$$\begin{aligned} \left(\frac{\delta\rho}{\rho}\right)_{t_{\text{f}}} &= \left(\frac{\delta\rho}{\rho}\right)_{t_U} \left[1 + \frac{1}{N_U} \ln\left(\frac{l}{L_U}\right)\right] \simeq \\ &\simeq \left(\frac{\delta\rho}{\rho}\right)_{t_U} \left(1 + \ln\frac{l}{L_U}\right)^{1/N_U} \simeq \left(\frac{\delta\rho}{\rho}\right)_{t_U} \left(\frac{l}{L_U}\right)^{1/N_U}. \end{aligned} \quad (8.38)$$

Здесь принято во внимание, что второе слагаемое в скобках много меньше единицы на больших масштабах, таких, что $l/L_U \geq 0.01$. Следовательно, амплитуда крупномасштабных флуктуаций очень слабо зависит от их пространственного размера. В этом случае говорят о почти плоском спектре флуктуаций. Проведенное разложение в ряд по $l \approx L_U$ приводит к степенному спектру со спектральным индексом $n \equiv 1 - 2/N_U < 1$. Отметим, что по наблюдаемой величине спектрального индекса можно судить о перспективности конкретной модели инфляции.

8.3. Развитие возмущений плотности

8.3.1. Возмущения плотности в пространстве Минковского

Проследим за развитием возмущений плотности энергии. Будем считать, что субстанция — носитель энергии может рассматри-

ваться как идеальная жидкость с тензором энергии-импульса

$$T_{\mu\nu} = pg_{\mu\nu} + (p + \rho)U_\mu U_\nu.$$

Здесь U_μ — компоненты скорости элементарного объема, занятого жидкостью. В сопутствующей системе отсчета

$$U^t = 1, \quad U^i = 0.$$

Гидродинамические уравнения движения имеют вид

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \nabla(\rho \bar{v}) &= 0, \\ \frac{\partial \bar{v}}{\partial t} + (\bar{v} \nabla) \bar{v} + \frac{1}{\rho} \nabla p + \nabla \varphi &= 0, \\ \Delta \varphi &= 4\pi G \rho. \end{aligned} \tag{8.39}$$

Решение этого уравнения в нулевом приближении элементарно:

$$\bar{v}_0 = 0, \quad \rho_0 = \text{const}, \quad p_0 = \text{const}.$$

Вообще говоря, это лишь приближенное решение, удовлетворительное при слабом влиянии потенциала гравитационного поля. Рассмотрим поведение малых флуктуаций, отметив их индексом 1. Система (8.39) упрощается:

$$\begin{aligned} \frac{\partial \rho_1}{\partial t} + \nabla(\rho_0 \bar{v}_1) &= 0, \\ \frac{\partial \bar{v}_1}{\partial t} + \frac{v_s^2}{\rho_0} \nabla \rho_1 + \nabla \varphi_1 &= 0, \\ \Delta \varphi_1 &= 4\pi G \rho_1. \end{aligned} \tag{8.40}$$

Скорость звука $v_s^2 \equiv (\partial p / \partial \rho)_{\text{adiabat}} \simeq \frac{p_1}{\rho_1}$.

Из уравнений (8.40) нетрудно получить волновое уравнение

$$\frac{\partial^2 \rho_1}{\partial t^2} - v_s^2 \nabla^2 \rho_1 = 4\pi G \rho_0 \rho_1,$$

которое нужно исследовать на устойчивость. Для этого ищем решение в виде плоских волн

$$\rho_1(\bar{r}, t) = A\rho_0 e^{-i\bar{k}\bar{r} + i\omega t}.$$

Из выражения для частоты

$$\omega = \pm \sqrt{k^2 v_s^2 - 4\pi G \rho_0}$$

видно, что при малых частотах, таких, что

$$k < k_J \equiv \sqrt{4\pi G \rho_0 / v_s^2},$$

решение неустойчиво.

Теория гравитационной неустойчивости впервые обсуждалась в статье Джеймса Джинса “Стабильность сферической туманности” (1902 г.), в которой количественно рассматривалась гравитационная нестабильность газового облака. Встречающийся в данном разделе индекс “J” связан с фамилией ученого.

Растущие моды разрушают нулевое приближение в виде однородного распределения плотности энергии. Под действием гравитации происходит уплотнение структур. Происходит экспоненциальный рост флуктуаций, который заканчивается рождением звезд и звездных скоплений. Размер нестабильной области по порядку величины есть

$$l \equiv l_J > 2\pi/k_J.$$

Можно также оценить джинсовскую массу вещества M_J , выше которой структура оказывается неустойчивой:

$$M_J \equiv \frac{4\pi}{3} \left(\frac{\pi}{k_J} \right)^3 \rho_0 = \frac{\pi^{5/2}}{6} \frac{v_s^2}{G^{3/2} \rho_0^{1/2}}.$$

8.3.2. Возмущения плотности в расширяющейся Вселенной

В предыдущем разделе не учитывалось космологическое расширение. Поэтому вывод справедлив в ограниченном числе ситуаций, например, при конденсации внутригалактического газа.

В однородной Вселенной расширение учитывается в виде зависимости от времени масштабного фактора $a(t)$. При этом уравнения (8.39) сохраняют свой вид, а их решение меняется кардинально. В качестве нулевого приближения выберем решение в виде

$$\rho_0(t) = a^{-3}(t)\rho_0(t_0), \quad \vec{v}_0(t) = H(t)\vec{r}, \quad \nabla\varphi_0 = \frac{4\pi G}{3}\rho_0(t)\vec{r}.$$

Вводим малые возмущения, отмечая их индексом 1. Действуя аналогично предыдущему, получим

$$\ddot{\delta}_k + 2H\dot{\delta}_k + \left(\frac{v_s^2 k^2}{a^2(t)} - 4\pi G\rho_0(t) \right) \delta_k = 0,$$

где δ_k есть фурье-образ относительной плотности $\delta\rho/\rho \simeq \rho_1/\rho_0$. Скорость звука v_s пропорциональна скорости частиц, следовательно, $v_s \sim 1/a(t)$. Нестабильный режим наступает при больших длинах волн, т.е. при

$$k < k_J \equiv a(t)\sqrt{4\pi G\rho_0(t)/v_s^2(t)}.$$

Рассмотрим рост длинноволновых флуктуаций, таких, что $k \ll k_J$. Вспоминая, что на стадии доминантности материи выполняются соотношения

$$H(t) = \frac{\dot{a}}{a} = \frac{2}{3t}, \quad \rho_0(t) = \frac{1}{6\pi Gt^2},$$

получим из исходного уравнения:

$$\ddot{\delta}_k + \frac{4}{3t}\dot{\delta}_k - \frac{2}{3t}\delta_k = 0,$$

которое приводит к растущим решениям

$$\delta_k(t) = \delta_k(t_{\text{in}}) \left(\frac{t}{t_{\text{in}}} \right)^{2/3}.$$

Образовавшиеся длинноволновые флуктуации растут, но гораздо медленнее, чем в стационарной Вселенной. В дальнейшем из этих флуктуаций плотности образуются галактики и скопления галактик.

8.4. Массивные первичные черные дыры

Процесс образования галактик включает в себя рождение квантовых флуктуаций плотности энергии и их пространственный рост на инфляционной стадии, рост плотности после захода флуктуации под горизонт и последующим образованием звезд. Тем не менее, не все так очевидно в данной картине, где темным местом являются черные дыры (ЧД). В стандартной картине звезды постепенно оседают к центру и сливаются, образуя массивную черную дыру, масса которой растет со временем. Однако возможность формирования галактик со сверхмассивными черными дырами в центре становится все более интригующей и неясной, особенно в связи с открытием далеких квазаров с красным смещением $z > 6$. Максимальное красное смещение $z = 6.41$ принадлежит квазару, светимость которого соответствует аккреции вещества на ЧД с массой $3 \cdot 10^9 M_\odot$. Подобное раннее образование черных дыр с массой $\sim 10^9 M_\odot$ может оказаться серьезной проблемой для стандартных моделей образования ЧД [9]. Таким образом, сценарии с первичными массивными ЧД [61, 109, 267, 314] привлекают внимание как возможная альтернатива стандартному сценарию. Подобные первичные черные дыры (ПЧД) могут быть центрами конденсации барионов и темной материи [9, 129] с последующим образованием протогалактик. Имеются две возможности: формирование изначально массивных ПЧД с последующим ростом до масс порядка $\sim 10^9 M_\odot$

благодаря аккреции материи, либо первичное рождение ПЧД небольших масс с их последующим слиянием в сверхмассивную ЧД в процессе иерархического скучивания протогалактик, содержащих ПЧД.

Существует несколько механизмов формирования массивных ПЧД. Рассмотрим один из них, основанный на том факте, что современные модели инфляции широко используют потенциалы сложной формы с несколькими минимумами. Фактически он уже использовался при обсуждении гибридной инфляции в разделе 6.2.1, [265]. Ниже мы покажем, что самые разные потенциалы приводят к формированию массивных ПЧД. Более того, оказывается, что не так просто избежать их перепроизводства в ранней Вселенной. Фактически рассмотрение любой инфляционной модели, использующей потенциал с двумя или более минимумами, должно учитывать этот механизм.

На рис. 8.2 изображен гипотетический потенциал инфлатонного поля. Пространственная область, схематично нарисованная слева, заполнена полем, величина которого отмечена черной точкой на кривой потенциала. Во время инфляции эта область, как и все остальные, быстро увеличивается в размерах. Кроме того, поле внутри области меняется со временем случайным образом за счет квантовых флуктуаций. На среднем рисунке показано, что поле внутри подобласти, закрашенной черным, может случайно оказаться по другую сторону “холма” — максимума потенциала. Затем инфляция увеличит многократно размер этой маленькой области. Когда инфляция закончится, поле начнет быстро скатываться к минимуму потенциала, как шарик с горки. Но в то время как поле в основной области скатывается к своему минимуму, точка “а” на нижнем рисунке, поле внутри подобласти движется к другому минимуму, “b”.

Следующий рисунок, рис. 8.3 (Г), показывает возможное расположение в пространстве обеих областей после того как поле скатилось в соответствующие минимумы. Расширение пространства во время инфляции многократно увеличило размеры обе-

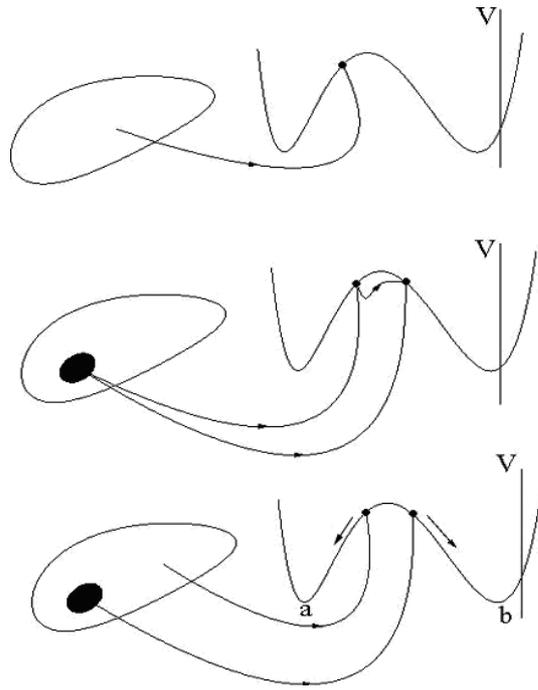


Рис. 8.2. Рождение полевой стенки в результате квантовой флуктуации вблизи максимума потенциала.

их областей. Поскольку основная пространственная область и ее подобласть увеличились в одинаковое число раз, последняя осталась внутри основной области. Теперь обширная область занята полем “а”, в то время как меньшая область содержит поле “b”. Ясно, что если мы пойдем (мысленно) из внутренней области наружу в любом направлении, т.е. из одного минимума в другой, то обязательно побываем в точке, где потенциал максимален. Значит, маленькая область окружена замкнутой поверхностью с повышенной плотностью энергии. Возникла так называемая замкнутая полевая стенка.

Исследуем вероятность образования замкнутых стенок опре-

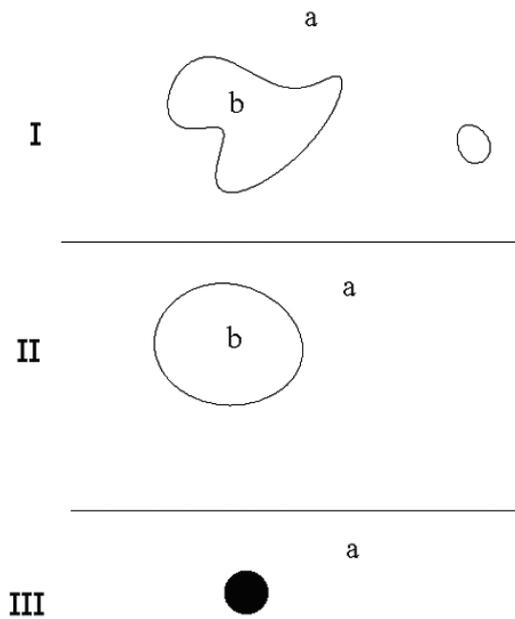


Рис. 8.3. Стенки больших размеров коллапсируют, образуя черные дыры. Стенки малых размеров исчезают, оставляя после себя колебания скалярного поля.

деленного размера. Для этого предположим, что наша Вселенная родилась из квантовой флуктуации вблизи максимума потенциала или его седловой точки. Это позволяет аппроксимировать потенциал, ограничиваясь квадратичным слагаемым:

$$V = V_0 - \frac{m^2}{2}\phi^2. \quad (8.41)$$

Полагаем без потери общности, что максимум находится при $\phi = 0$. Тогда плотность вероятности найти определенное значение поля ϕ имеет вид [188] (по сравнению с цитированной книгой учтено, что движение происходит около максимума потенциала,

а не минимума)

$$dP(\phi, T; \phi_{\text{in}}, 0) = d\phi \sqrt{\frac{a}{\pi(e^{2\mu T} - 1)}} \exp \left[-a \frac{(\phi - \phi_{\text{in}} e^{\mu T})^2}{e^{2\mu T} - 1} \right]. \quad (8.42)$$

Здесь $a = \mu/\sigma^2$, $\mu \equiv m^2/(3H)$ и $\sigma = H^{3/2}/(2\pi)$, где параметр Хаббла

$$H \simeq \sqrt{8\pi V_0/(3m_4)}.$$

Пусть начальное значение поля в момент рождения Вселенной удовлетворяет неравенству $\phi_{\text{in}} > 0$. Тогда, в точности по сценарию, описанному чуть выше, среднее значение поля увеличивается со временем, достигая некоего минимума $\phi = \phi_+ > 0$. Основная часть пространства будет содержать именно это значение поля. Тем не менее уже с самого начала квантовые флуктуации в некоторых пространственных областях заставляют поле “перепрыгивать” максимум потенциала. Поле в таких областях движется затем к другому минимуму $\phi_- < 0$. Следовательно, нам надо интересоваться флуктуациями, приводящими к $\phi \leq 0$. Последние сильно зависят от значения поля ϕ_{in} . На рис. 8.4 представлена вероятность

$$P_0(\phi_{\text{in}}, T) = \int_{\phi=-\infty}^{\phi=0} dP(\phi, T; \phi_{\text{in}}, 0) \quad (8.43)$$

найти поле $\phi \leq 0$ в произвольной точке пространства, что эквивалентно доле пространственного объема, занятого “другим” вакуумом. Очевидно, эта доля резко возрастает с приближением начального значения поля ϕ_{in} к максимуму потенциала.

Если области с $\phi = \phi_-$ составляют небольшую долю пространства, то они оказываются окруженными замкнутыми полевыми стенками.

Дальнейшая эволюция замкнутой стенки происходит следующим образом. Стенка, обладая избытком энергии, стремится в энергетически наиболее выгодное состояние. Она, как мыльный пузырь, приобретает сферическую форму и сжимается, см.

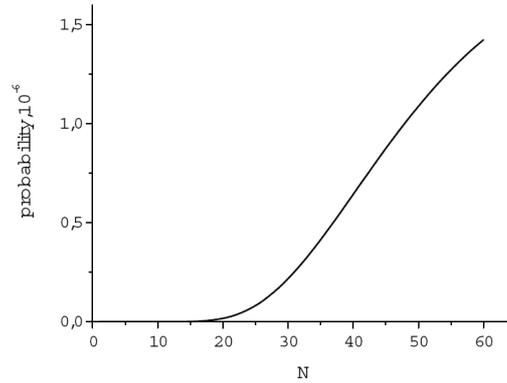


Рис. 8.4. Часть пространства, заполненного другим вакуумом. Время измеряется в e-фолдах. Начальное значение поля $\phi_{\text{in}} = \phi_{\text{max}} + 3H$, где ϕ^{max} — поле в минимуме потенциала. Масса инфлатона принята равной $m = 0.3H$, где H есть параметр Хаббла при $V = V_0$.

рис. 8.3 (II). (Если бы внутри мыльного пузыря не было воздуха, он делал бы то же самое.) Судьба маленькой стенки малоинтересна — она сжимается к центру, происходит столкновение разных областей стенки друг с другом, в результате чего вся ее энергия переходит в разлетающиеся волны поля, из которого состояла стенка. Совсем другая ситуация с большими стенками. Сжимаясь, они доставляют колоссальную энергию в небольшой объем пространства. Эта энергия удерживается гравитационным полем, и образуется компактный объект, рис. 8.3 (III). Более подробный анализ показывает, что такие объекты — черные дыры — образуются целыми семьями, кластерами, рис. 8.5. Как видно, никаких звезд для образования этих массивных первичных черных дыр не требуется! Наоборот, именно такой кластер черных дыр способен создать гравитационную яму и собрать там достаточное количество барионов, из которых и образуются затем звезды. Уже имеются сведения об одиноких сверхмассивных черных дырах, в окружающем пространстве которых отсутству-

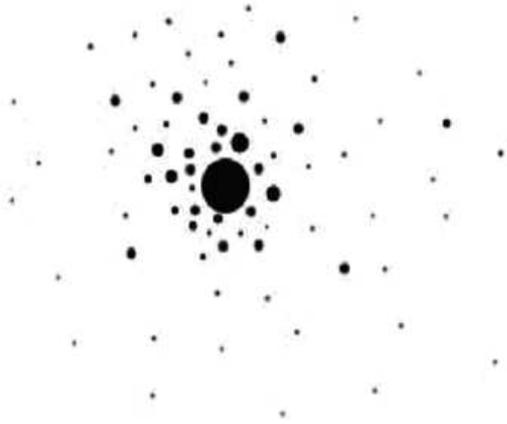


Рис. 8.5. Типичная картина кластера первичных черных дыр.

ет светящееся вещество. Это можно рассматривать как косвенное подтверждение механизма образования массивных черных дыр, рассмотренного выше.

Конкретная реализация этого механизма была развита в работах [129, 187, 188, 265, 267]. В этих статьях были исследованы свойства кластеров ПЧД, получены их массовые и пространственные распределения. В качестве исходного был принят лагранжиан скалярного поля с потенциалом типа “сомбреро на бекрень”. Оказалось, что свойства кластеров ПЧД сильно зависят от величины начальной фазы. Кроме того, эти свойства зависят от наклона потенциала Λ , масштаба нарушения симметрии f в начале инфляции, что приводит к существенным вариациям распределения по массам кластеров ПЧД. В статье [129] выбраны такие параметры модели, при которых образуются изначально большие кластеры массивных ПЧД с массами центральных ЧД $\sim 4 \cdot 10^7 M_{\odot}$, которые увеличиваются благодаря аккреции до значений $\sim 10^9 M_{\odot}$ и, следовательно, объясняют наблюдаемую

активность удаленных квазаров.

В данном сценарии имеется несколько стадий формирования ЧД и галактик:

- 1) Образование замкнутых стенок скалярного поля сразу после окончания инфляции и их коллапс в кластер ПЧД согласно [187, 267, 269]. Наиболее массивная ПЧД образуется в центре кластера после пересечения космологического горизонта.
- 2) Отделение центральной плотной области кластера от космологического расширения и его вириализация. Большое количество менее массивных ПЧД, окружающих центральную, сливаются с ней, увеличивая ее массу.
- 3) Отделение внешних областей кластера, в которых доминирует темная материя, от космологического расширения и рост протогалактики. Прекращение роста протогалактики благодаря взаимодействию с обычными флуктуациями темной материи.
- 4) Охлаждение газа и звездообразование, слияние протогалактик и формирование современных галактик.

Не исключено, что оба сценария образования галактик — стандартный и с ПЧД — реализуются одновременно. При этом, поскольку на начальном этапе галактики интенсивно сливаются, достаточно небольшой доли протогалактик с массивной черной дырой в центре, чтобы наблюдать современное обилие сверхмассивных черных дыр в центрах галактик.

Глава 9

Многомерие

9.1. Введение

Идея дополнительных измерений, берущая начало от пионерских работ Т. Калуцы [180] и О. Клейна [194], в последние десятилетия стала необходимой составляющей практически всех попыток объединения четырех физических взаимодействий. Наибольшую популярность снискали модели размерностей $D = 10$ (суперструны) и 11 (М-теория), однако имеется целый ряд моделей и более высоких размерностей. Бозонный сектор подобных теорий обычно содержит, помимо метрики, различные виды скалярных и векторных полей, а также поля антисимметричных форм; в то же время известно, что дополнительные (по отношению к обычным четырем измерениям) компоненты D -мерной метрики естественно порождают скалярные и векторные поля при редукции к четырехмерному пространству-времени (см., напр., [5] и приведенные там ссылки).

Простейший случай такой редукции, с появлением в результате эйнштейновской гравитации и максвелловского электромагнитного поля, и описан в работе [180]. Более сложные единые многомерные теории, обеспечивающие получение из многомерной гравитации бозонных полей, входящих в Стандартную мо-

дель физики частиц, описаны в книге [5] (там же см. ссылки на другие работы этого направления, включая многочисленные работы группы Ю.С. Владимирова). Отдельную проблему представляет геометризация фермионных полей.

Особый интерес представляют многомерные теории с гравитационным действием, содержащим нелинейные по кривизне члены. Их появление является прямым следствием квантовой теории поля в искривленном пространстве-времени [7, 68] и потому даже не должно рассматриваться как независимый постулат.

Согласно [94, 98], многомерная гравитация с лагранжианами, нелинейными по кривизне, приводит к множеству различных низкоэнергетических теорий благодаря произвольности исходных параметров эффективного действия и различным вариантам начальных данных. Под последними подразумеваются как метрика, так и топология дополнительных измерений, которые могут возникать в результате квантовых флуктуаций при высоких энергиях. Таким образом, даже при фиксированных параметрах исходного эффективного лагранжиана возможно появление многих вселенных с сильно различающимися свойствами. Даже при сравнительно простых предположениях относительно многомерной геометрии в результате редукции к четырем наблюдаемым измерениям возникают эффективные скалярные поля со сложным видом потенциалов, свойства которых во многом и определяют физические свойства той или иной вселенной.

Такая точка зрения близка к концепции хаотической инфляции, согласно которой постоянно рождается неограниченно много вселенных при флуктуациях скалярного (инфлатонного) поля. При достаточно сложном виде потенциала ситуация напоминает так называемую концепцию ландшафта в теории струн: число различных вакуумов в теории гетеротических струн достигает 10^{1500} [207]); число наиболее реалистичных деситтеровских вакуумов [178] также огромно, хотя и конечно. К предсказанию различных видов вселенных ведет и концепция случайного потенциала [264, 266]. Мы пытаемся сделать дальнейший шаг, объ-

ясняя происхождение таких потенциалов с позиций многомерной гравитации.

Получение конкретных результатов в многомерной нелинейной гравитации существенно облегчается при использовании приближения медленных изменений [94, 98], которое предполагает малость всех производных (в смысле, обсуждаемом ниже) и малость всех плотностей энергии по сравнению с планковскими масштабами. Такие предположения оказываются работающими даже при исследовании инфляционных моделей в космологии, характеризуемых масштабом энергии Большого объединения, не говоря уже о процессах при меньших энергиях и в условиях современной космологии. Этим методом можно рассматривать самые различные виды гравитационного действия в широких областях значений содержащихся в них параметров.

В концепции Калуцы — Клейна дополнительные измерения пространства остаются невидимыми в силу их чрезвычайной малости. В последние годы, однако, широко обсуждается альтернативная концепция — концепция мира на бране, которая рассматривает нашу Вселенную как выделенную трехмерную (или, с учетом временного измерения, четырехмерную) поверхность или слой, называемый браной, в многомерном пространстве, в котором дополнительные измерения имеют большие или даже бесконечные размеры. Они остаются невидимыми в силу того, что поля Стандартной модели (а следовательно, и наблюдатели) предполагаются сосредоточенными на бране, а гравитация (и, как правило, только она) распространяется в окружающем объеме. История подобных моделей прослеживается с начала 80-х годов - с работ Акамы [44], Рубакова и Шапошникова [262] и др. Недавний всплеск интереса к подобным моделям связан в основном с достижениями теории струн и М-теории, в частности, с известной 11-мерной моделью Хожавы и Виттена [172, 311], в которой одно из дополнительных измерений обладает значительно большими размерами, чем остальные. В рамках этого подхода был предложен естественный механизм решения проблемы

иерархии в физике частиц (работы Рэндалл и Сундрума [252] и др.), при возможности сохранения ньютоновского поведения слабого поля тяготения на бране в согласии с современными экспериментами. В моделях мира на бране оказалось возможным существенное продвижение в нерешенных фундаментальных проблемах как физики частиц, так и космологии.

В данной главе мы опишем несколько классов моделей Вселенной, возникающих из нелинейной многомерной гравитации, без привлечения каких-либо полей, кроме метрического, в исходном действии. Среди них космологические модели, способные описывать как инфляцию в ранней Вселенной, так и современный этап ускоренного расширения Вселенной, а также модели мира на бране смешанного типа — с бесконечным пятым измерением и произвольным числом компактных калуца-клейновских измерений.

9.2. Компактные дополнительные пространства

Мысль о том, что число измерений в нашем мире больше четырех возникла в начале прошлого века. Идея оказалась одновременно и простой, и глубокой. Простой — потому что вообразить ситуацию довольно просто, по крайней мере на двух- либо трехмерных аналогах. Так, лист бумаги воспринимается нами (и используется) как двумерный объект, но для микробов он однозначно трехмерен. Очевидно, что благодаря малому размеру, скажем, пятого измерения, оно способно ускользать от нашего восприятия. Глубокой же идея оказалась благодаря многочисленным и нетривиальным следствиям, из нее вытекающим. Для получения значимых результатов требуется владение соответствующим математическим аппаратом. Поэтому далее кратко излагаются необходимые сведения.

Рассмотрим D -мерное дифференцируемое многообразие \mathbb{M}_D ,

которое разложимо в прямое произведение

$$\mathbb{M}_D = \mathbb{M}_4 \times \mathbb{M}_d. \quad (9.1)$$

Пространство дополнительных измерений \mathbb{M}_d предполагается компактным и имеющим объем $V_d < \infty$. Вводим координаты многообразия \mathbb{M}_D : $\{X\} = (\{x\}, \{y\})$, где набор x^μ , $\mu = 1, 2, 3, 4$ описывает четырехмерное пространство \mathbb{M}_4 , а набор y_a , $a = 1, 2, \dots, d$ — дополнительное пространство \mathbb{M}_d . Интервал записывается с помощью метрического тензора обычным образом:

$$ds^2 = G_{MN} dX^M dX^N, \quad (9.2)$$

где индексы M, N пробегает значения $1, 2, \dots, d + 4$. Объем дополнительного пространства выражается через положительно-определенную внутреннюю метрику γ_{ab} :

$$V_d = V(\mathbb{M}_d) = \int d^d y \sqrt{\gamma} < \infty, \quad (9.3)$$

где $\gamma = \det(\gamma_{ab})$. Последнее неравенство — необходимое условие компактности этого подпространства.

Скалярные поля и дополнительное пространство

Для демонстрации возможностей рассмотрим поведение скалярного поля Φ , определенного в полном пространстве \mathbb{M}_D , действие для которого имеет вид

$$S = \int_{\mathbb{M}_D} d^D X \sqrt{|G|} \left[\frac{1}{2} (\partial_M \Phi \partial^M \Phi) - U(\Phi) \right] \quad (9.4)$$

Варьируя действие по Φ при фиксированной метрике, приходим к уравнению движения

$$\square_D \Phi + dU/d\Phi = 0.$$

Здесь D -мерный оператор д'Аламбера определяется следующим образом:

$$\square_D \Phi \equiv \nabla_M \nabla^M \Phi = \frac{1}{\sqrt{|G|}} \partial_M \left(\sqrt{|G|} \partial^M \Phi \right), \quad (9.5)$$

где $G = \det(G_{MN})$. Если ограничиться малыми отклонениями от положения равновесия (минимума потенциала), то можно аппроксимировать потенциал квадратичным слагаемым $U(\Phi) \simeq M^2\Phi^2/2$. Тогда уравнение движения упрощается, принимая вид

$$(\square_D + M^2)\Phi = (\square_4 + \square_d + M^2)\Phi = 0,$$

где \square_4 и \square_d — операторы д'Аламбера в соответствующих подпространствах.

Введем полный ортонормированный набор функций $Y_n(y)$, таких, что

$$\square_d Y_n(y) = \lambda_n Y_n(y). \quad (9.6)$$

Поскольку \mathbb{M}_d — компактное пространство, собственные значения оператора д'Аламбера λ_n дискретны. Свойство ортонормированности записывается следующим образом:

$$\int_{\mathbb{M}_d} d^d y \sqrt{|\gamma|} Y_n(y) Y_m(y) = \delta_{nm} \quad (9.7)$$

Разлагая Φ в ряд по собственным функциям $Y_n(y)$,

$$\Phi(X) = \sum_n Y_n(y) \Phi_n(x), \quad (9.8)$$

и учитывая выражения (9.6) и (9.7), нетрудно выписать уравнение поля Φ в пространстве \mathbb{M}_4 :

$$(\square_4 + \lambda_n + M^2)\Phi_n(x) = 0. \quad (9.9)$$

Поля $\Phi_n(x)$ воспринимаются наблюдателем как скалярные поля с массами $M_n = M^2 + \lambda_n$. Ненаблюдаемость скалярных частиц обычно объясняется их большой массой.

Скалярные поля. d -мерный тор

Предположим, что дополнительное пространство \mathbb{M}_d имеет геометрию d -мерного тора, $\mathbb{M}_d = \mathbb{T}_d = \mathbb{S}^1 \times \mathbb{S}^1 \times \mathbb{S}^1 \times \dots \times \mathbb{S}^1$, $y^m \in [0, 2\pi R]$. Тогда явный вид собственных значений (9.6) хорошо известен:

$$\lambda_n \equiv \lambda_{\{n_1, n_2, \dots, n_d\}} = \sum_{j=1}^d \frac{n_j^2}{R^2}. \quad (9.10)$$

При этом собственные функции имеют вид

$$Y_{\{n_1, n_2, \dots, n_d\}}(y) = \frac{1}{\sqrt{V_T}} \exp \left[\frac{i \sum n_j y^j}{R} \right], \quad V_T = (2\pi R)^d. \quad (9.11)$$

Если радиус тора R , характеризующий размер дополнительного пространства, мал, то все возбуждения с $j \geq 1$ соответствуют сверхмассивным частицам.

9.2.1. Модель Калуцы — Клейна с одним дополнительным измерением

Замкнутое одномерное многообразие (замкнутая кривая) с точки зрения его внутренней геометрии представляет собой окружность \mathbb{S}^1 . Поэтому структура (9.1) в случае $d = 1$ имеет вид

$$\mathbb{M}_5 = \mathbb{M}_4 \otimes \mathbb{S}^1. \quad (9.12)$$

Удобно представить метрику полного пространства \mathbb{M}_5 в виде матрицы

$$G_{AB} = \begin{pmatrix} G_{\mu\nu} & G_{\mu 5} \\ G_{5\mu} & G_{55} \end{pmatrix}, \quad (9.13)$$

где G_{55} , $G_{5\mu}$, $G_{\mu\nu}$ — скаляр, вектор и тензор в четырехмерном пространстве \mathbb{M}_4 .

Как извлечь калибровочную симметрию из дополнительных измерений

Чтобы понять идею, введем ограничение $G_{55} = \text{const} = -1$. В этом случае преобразования вида

$$\begin{aligned} x'^{\mu} &= x^{\mu}, \\ y' &= y + \varepsilon(x), \quad \varepsilon(x) \ll 1 \end{aligned} \quad (9.14)$$

вливают на метрический тензор следующим образом:

$$G'_{\mu\nu} = \frac{\partial x^A}{\partial x'^{\mu}} \frac{\partial x^B}{\partial x'^{\nu}} G_{AB}, \quad G'_{\mu 5} = \frac{\partial x^A}{\partial x'^{\mu}} G_{A5},$$

а коэффициент G_{55} остается неизменным; в нашем случае бесконечно малых сдвигов

$$\begin{aligned} G'_{\mu\nu} &= G_{\mu\nu} - G_{\nu 5} \partial_{\mu} \varepsilon - G_{\mu 5} \partial_{\nu} \varepsilon + O(\varepsilon^2), \\ G'_{\mu 5} &= G_{\mu 5} - G_{55} \partial_{\mu} \varepsilon = G_{\mu 5} + \partial_{\mu} \varepsilon. \end{aligned} \quad (9.15)$$

Сама идея дополнительных измерений возникла из желания описать единым образом гравитацию и электромагнитное (калибровочное) поле. В качестве компонент последнего могут выступать недиагональные элементы метрического тензора $G_{\mu 5}$, поскольку, как это видно из (9.15), их преобразования очень напоминают калибровочные. С выбором метрического тензора четырехмерного пространства ситуация сложнее. Выбор тензора $G_{\mu\nu}$ в качестве метрического тензора в основном подпространстве \mathbb{M}_4 не представляется удачным, поскольку он зависит от выбора координаты в дополнительном пространстве, см. (9.15). С другой стороны, нетрудно сообразить, что тензор

$$g_{\mu\nu} = G_{\mu\nu} + G_{\mu 5} G_{\nu 5}$$

инвариантен относительно инфинитезимальных преобразований (9.14). Именно его и удобно выбрать в качестве метрического тензора в основном подпространстве \mathbb{M}_4 . Введя обозначение $B_{\mu} = -G_{\mu 5}$, запишем метрический тензор в виде

$$G_{AB}(x) = \begin{pmatrix} g_{\mu\nu} - B_{\mu} B_{\nu} & -B_{\mu} \\ -B_{\nu} & -1 \end{pmatrix}. \quad (9.16)$$

Обратная к G_{AB} матрица имеет вид

$$G^{AB}(x) = \begin{pmatrix} g^{\mu\nu} & -B^\mu \\ -B^\nu & -1 + B_\lambda B^\lambda \end{pmatrix}. \quad (9.17)$$

Преобразование величины B_μ , при координатных преобразованиях (9.14)

$$B'_\mu = B_\mu - \partial_\mu \varepsilon, \quad (9.18)$$

особенно интересно: именно так преобразуется векторный потенциал электромагнитного поля при действии группы $U(1)$ на волновую функцию заряженной частицы. Важнейшим достоинством теорий типа Калуцы — Клейна является то, что калибровочные симметрии вытекают из симметрии компактного подпространства, а не постулируются дополнительно.

Инвариантность действия (9.4) относительно преобразований (9.14) оказывается причиной появления калибровочной инвариантности этого действия относительно преобразований (9.18).

Посмотрим, как влияет преобразование (9.14) на действительное скалярное поле $\Phi(x, y)$. Его можно представить рядом Фурье (9.8)

$$\Phi(x, y) = \sum_n \Phi_n(x) Y_n(y),$$

где $Y_n(y)$ — ортонормированный набор собственных функций оператора ∂_y^2 , действующего в дополнительном пространстве \mathbb{M}_1 . Так как в данном случае оно представляет собой окружность некоторого радиуса r , то

$$Y_n(y) = \frac{1}{\sqrt{2\pi r}} e^{-iny/r}.$$

Симметрия дополнительного пространства (окружности) проявляется в виде инвариантности функций $Y_n(y)$ относительно сдвигов $y \rightarrow y + 2\pi r$. Скалярное поле преобразуется при замене $y \rightarrow y + \varepsilon$ следующим образом:

$$\Phi'(x, y) = \Phi(x, y + \varepsilon).$$

Сравнивая разложения Фурье обеих частей равенства, получаем, что

$$\Phi'_n(x) = \Phi_n(x)e^{i(n/r)\varepsilon}.$$

Именно так преобразуется заряженное поле при калибровочных преобразованиях. Значит, Фурье-компонента $\Phi_n(x)$ может восприниматься как скалярное поле с зарядом $q_n \equiv n/r$.

Итак, мы видим, что $U(1)$ симметрия дополнительного пространства означает наличие калибровочного 4-вектора B_μ и заряженных частиц в пространстве M_4 . Очевидно, что лагранжиан в четырехмерии, который мы собираемся получить, будет инвариантен относительно калибровочных преобразований, поскольку они генерируются преобразованиями (9.14), относительно которых исходный лагранжиан заведомо инвариантен. Тем не менее, чтобы получить явный вид низкоэнергетического действия, необходимы дальнейшие рассуждения.

Пусть действие произвольного скалярного поля Φ имеет вид

$$S = \int d^5z \sqrt{G} \left(-\frac{1}{2} \Phi \square_G \Phi - \frac{1}{2} m^2 \Phi^2 \right).$$

Часто кинетическое слагаемое в действии записывается в виде

$$-\frac{1}{2} \frac{\partial \Phi}{\partial x^A} \frac{\partial \Phi}{\partial x_A}$$

Две записи отличаются на полную производную, которая не влияет на классические уравнения движения поля Φ . Метрику пятимерного пространства выберем в простейшем виде и аппроксимируем следующим образом, см. (9.16):

$$G_{AB}(x) = \begin{pmatrix} \eta_{\mu\nu} - B_\mu B_\nu & -B_\mu \\ -B_\nu & -1 \end{pmatrix}. \quad (9.19)$$

Здесь пренебрежено зависимостью калибровочного поля $B_\mu(x)$ от координаты y дополнительного пространства и динамикой метрического коэффициента G_{55} . Физически это означает, что,

по предположению, релаксационные процессы в пятом измерении уже закончились. Тогда оператор д'Аламбера представляется в виде

$$\square_G = \eta^{\mu\nu}(\partial_\mu - B_\mu \partial_y)(\partial_\nu - B_\nu \partial_y) + \square_y.$$

Подставляя это выражение в действие, получаем:

$$S = \frac{1}{2} \sum_n \int d^4x \sqrt{-g} \Phi_n^* (-\square_n + m_n^2) \Phi_n,$$

где

$$\begin{aligned} \square_n &= g^{\mu\nu}(\partial_\mu - iq_n A_\mu)(\partial_\nu - iq_n A_\nu), & A_\mu &\equiv \frac{-1}{\sqrt{2\kappa}} B_\mu, \\ q_n &= n \left(\sqrt{2\kappa}/R \right), & m_n^2 &= m^2 + n^2/R^2. \end{aligned}$$

На языке теории поля, полученное эффективное действие описывает набор скалярных полей с массами m_n^2 и зарядами q_n , взаимодействующих с калибровочным полем A_μ в пространстве Минковского.

Обобщение — $G_{55} \neq \text{const}$

До сих пор рассматривался простейший вариант 5-мерной теории, когда метрический коэффициент G_{55} не зависел от координат. Расширим пределы применимости вышеизложенного, полагая $G_{55} \neq \text{const}$.

Удобно ввести поле B_μ , обобщая ранее введенное следующим образом:

$$G_{\mu 5} = -\sqrt{|G_{55}|} B_\mu. \quad (9.20)$$

Как будет показано далее, 5-метрика может быть записана в виде

$$\begin{pmatrix} g_{\mu\nu} - e^{2\phi} B_\mu B_\nu & -e^\phi B_\mu \\ -e^\phi B_\mu & -e^{2\phi} \end{pmatrix}, \quad (9.21)$$

где введено скалярное поле ϕ :

$$G_{55} \equiv -e^{2\phi}. \quad (9.22)$$

Полученный результат позволяет выделить калибровочную симметрию в пятимерном действии Гильберта-Эйнштейна. Действительно, скаляр Риччи R_5 пятимерного пространства выражается через аналогичный скаляр R_4 четырехмерного пространства следующим образом:

$$R_5 = R_4 - 2e^{-\phi} \square e^\phi - \frac{1}{4} e^{2\phi} F_{\mu\nu} F^{\mu\nu}; \quad (9.23)$$

$$F_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu. \quad (9.24)$$

При этом пятимерное действие Эйнштейна-Гильберта после серии выкладок приобретает вид

$$\begin{aligned} S &= \frac{1}{16\pi G_5} \int_{\mathbb{M}_5} d^5 X \sqrt{|G|} (R_5 - 2\Lambda) = \\ &= \frac{1}{16\pi G_4} \int_{\mathbb{M}_4} d^4 x \sqrt{|g|} \left(R_4 - 2\Lambda' - \frac{3}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{4} e^{3\phi} F_{\mu\nu} F^{\mu\nu} \right). \end{aligned} \quad (9.25)$$

Вывод этой формулы опущен ввиду его длительности. В ходе выкладок было проведено конформное преобразование

$$g_{\mu\nu} \rightarrow \tilde{g}_{\mu\nu} = e^\phi g_{\mu\nu}$$

и введены обозначения $G_4 = G_5/(2\pi R)$, $\Lambda' = \frac{4}{3}\Lambda$.

Итак, исходя из предположения о существовании одного дополнительного измерения, мы получили четырехмерную гравитацию с безмассовым скалярным полем, взаимодействующим с абелевым калибровочным (электромагнитным) полем (теорию с таким взаимодействием часто называют дилатонной гравитацией, а поле ϕ дилатоном). Если $\phi = \text{const}$, то это обычная теория Эйнштейна-Максвелла.

Геометрический смысл $g_{\mu\nu}$

Дополнительное пространство в виде окружности обладает симметрией (группой изометрий) $U(1)$, преобразованиями которой являются повороты на произвольный угол. Метрический тензор $g_{\mu\nu}$ основного пространства также инвариантен относительно таких преобразований. Именно эта группа играет роль калибровочной группы для полученного из геометрии калибровочного поля B_μ , интерпретируемого как электромагнитное.

Благодаря простоте дополнительного пространства нам легко удалось сконструировать подходящую комбинацию из тензоров $G_{\mu\nu}$, $G_{\mu 5}$ и G_{55} . В более сложных случаях это процедура не столь очевидна. Поэтому необходимо более пристально разобраться в смысле введенных величин.

Поскольку метрический тензор G_{AB} содержит недиагональные элементы, пространство \mathbb{M}_4 с метрикой $G_{\mu\nu}$ не ортогонально пространству \mathbb{M}_1 . Найдем пространство \mathbb{U}_4 , ортогональное \mathbb{M}_1 . Малый вектор смещения, ортогональный к \mathbb{M}_1 , будем искать в виде

$$\vec{\delta} \equiv dx^\mu \vec{e}_\mu + dy \vec{e}_5, \quad (9.26)$$

где \vec{e}_μ, \vec{e}_5 — пятерка единичных векторов, причем \vec{e}_5 является касательным к \mathbb{M}_1 . Из условия ортогональности, записанном в виде

$$(dx^\mu \vec{e}_\mu + dy \vec{e}_5) \cdot \vec{e}_5 = 0,$$

находим смещение dy :

$$dy = -G_{55}^{-1} G_{5\mu} dx^\mu.$$

При этом модуль вектора смещения равен

$$\vec{\delta}^2 = \delta^M G_{MN} \delta^N = (G_{\mu\nu} - G_{55}^{-1} G_{\mu 5} G_{\nu 5}) dx^\mu dx^\nu \equiv g_{\mu\nu} dx^\mu dx^\nu. \quad (9.27)$$

Здесь использованы равенства $\delta^\mu = dx^\mu$, $\delta^5 = dy$, см. (9.26). Теперь очевиден смысл тензора $g_{\mu\nu}$ — он определяет интервал в пространстве, ортогональном к \mathbb{M}_1 .

Реперы

Введем реперный формализм, удобный для обобщения полученных результатов на компактные дополнительные пространства d измерений.

Известно, что вместо метрического тензора для описания гравитации можно ввести совокупность линейно-независимых реперных векторов e_A . Индекс A нумерует векторы в пространстве \mathbb{M}_5 , $A = 1, 2, \dots, 5$. Компоненты метрического тензора g_{AB} в базисе векторов e_A определяются через их скалярное произведение

$$g_{AB} = e_A \cdot e_B.$$

Базисные векторы можно подобрать так, чтобы метрические компоненты тензора, определенные с их помощью, имели канонический вид $\eta_{AB} = \text{diag}(1, -1, -1, \dots, -1)$. Репер из векторов $e_{\bar{A}}$, таких, что $e_{\bar{A}}e_{\bar{B}} = \eta_{AB}$, называется лоренцевым базисом. Один базис можно разложить по другому:

$$e_A = b_A^{\bar{A}}e_{\bar{A}} = b_A^{\bar{\mu}}e_{\bar{\mu}} + b_A^{\bar{5}}e_{\bar{5}}. \quad (9.28)$$

Суммирование подразумевается. Векторы $e_{\bar{5}}$ и $e_{\bar{5}}$ выбираются сонаправленными, $e_{\bar{5}} = b_{\bar{5}}^{\bar{5}}e_{\bar{5}}$. Тогда из выражения (9.28) сразу следует, что $b_{\bar{5}}^{\bar{\mu}} = 0$. Разложение (9.28) можно представить в матричном виде с матрицей

$$\begin{pmatrix} b_{\bar{\mu}}^{\bar{\mu}} & 0 \\ b_{\bar{\mu}}^{\bar{5}} & b_{\bar{5}}^{\bar{5}} \end{pmatrix},$$

действующей на столбцы $(e_{\bar{\mu}}, e_{\bar{5}})$, $e_{\bar{\mu}}, e_{\bar{5}}$. Обратная матрица

$$\begin{pmatrix} h_{\bar{\mu}}^{\bar{\mu}} & h_{\bar{\mu}}^{\bar{5}} \\ h_{\bar{5}}^{\bar{\mu}} & h_{\bar{5}}^{\bar{5}} \end{pmatrix}$$

обладает важным свойством

$$h_{\bar{5}}^{\bar{5}}b_{\bar{5}}^{\bar{5}} = 1.$$

Компоненты метрического тензора в пространстве \mathbb{M}_4 выражаются следующим образом:

$$\begin{aligned} G_{\mu\nu} &\equiv e_\mu \cdot e_\nu = b_\mu^{\bar{A}} b_\nu^{\bar{B}} e_{\bar{A}} e_{\bar{B}} = b_\mu^{\bar{\mu}} b_\nu^{\bar{\nu}} e_{\bar{\mu}} e_{\bar{\nu}} + b_\mu^{\bar{5}} b_\nu^{\bar{5}} e_{\bar{5}} e_{\bar{5}} = \\ &= b_\mu^{\bar{\mu}} b_\nu^{\bar{\nu}} \eta_{\bar{\mu}\bar{\nu}} + b_\mu^{\bar{5}} b_\nu^{\bar{5}} \eta_{\bar{5}\bar{5}} = g_{\mu\nu} - b_\mu^{\bar{5}} b_\nu^{\bar{5}}, \end{aligned} \quad (9.29)$$

где введен тензор $g_{\mu\nu} \equiv b_\mu^{\bar{\mu}} b_\nu^{\bar{\nu}} e_{\bar{\mu}} e_{\bar{\nu}} = b_\mu^{\bar{\mu}} b_\nu^{\bar{\nu}} \eta_{\bar{\mu}\bar{\nu}}$, очевидно, инвариантный относительно преобразований координат в дополнительном пространстве. Кроме того,

$$G_{\mu 5} \equiv e_\mu \cdot e_5 = b_\mu^{\bar{A}} b_5^{\bar{B}} e_{\bar{A}} e_{\bar{B}} = b_\mu^{\bar{5}} b_5^{\bar{5}} e_{\bar{5}} e_{\bar{5}} = -b_\mu^{\bar{5}} b_5^{\bar{5}},$$

$$G_{55} \equiv e_5 \cdot e_5 = b_5^{\bar{5}} b_5^{\bar{5}} e_{\bar{5}} e_{\bar{5}} = -b_5^{\bar{5}} b_5^{\bar{5}}.$$

Определив поле B_μ следующим образом:

$$B_\mu \equiv B_\mu^{\bar{5}} \equiv h_{\bar{5}}^{\bar{5}} b_\mu^{\bar{5}},$$

приходим к выражениям, полученным выше, (см. (9.21), (9.22)), где $b_5^{\bar{5}} = e^\phi$:

$$G_{\mu\nu} = g_{\mu\nu} - b_5^{\bar{5}} b_5^{\bar{5}} B_\mu B_\nu;$$

$$G_{\mu 5} = -b_5^{\bar{5}} B_\mu;$$

$$G_{55} = -b_5^{\bar{5}} b_5^{\bar{5}}.$$

9.2.2. Модель Калуцы-Клейна. Общий случай.

Не существует причин, по которым необходимо остановиться на одном дополнительном измерении. Как будет видно в дальнейшем, богатые возможности возникают при увеличении числа измерений. Поэтому рассмотрим риманово пространство с произвольным числом измерений, выделив, по понятным причинам, физическое четырехмерное подпространство

$$\mathbb{M}_D = \mathbb{M}_4 \otimes \mathbb{M}_d \quad (9.30)$$

В предыдущем разделе был рассмотрен метод расщепления метрического тензора с использованием понятия реперов. Этот

метод легко обобщается на произвольное число дополнительных измерений. Далее действуем полностью аналогично 5-мерному варианту.

Вводим базисные вектора: координатный базис e_A для разбиения $\mathbb{M}_4 \otimes \mathbb{M}_d$ и $e_{\bar{A}}$ - лоренцев базис того же пространства. Один базис можно разложить по другому

$$e_A = b_A^{\bar{A}} e_{\bar{A}} = b_A^{\bar{\mu}} e_{\bar{\mu}} + b_A^{\bar{a}} e_{\bar{a}}. \quad (9.31)$$

Суммирование подразумевается. Индексы μ, ν относятся к пространству \mathbb{M}_4 , а a, b - к пространству \mathbb{M}_d .

Разложение (9.31), в полной аналогии с предыдущим случаем, может быть представлено в матричном виде с матрицей перехода

$$\begin{pmatrix} b_{\mu}^{\bar{\mu}} & b_a^{\bar{\mu}} \\ b_{\mu}^{\bar{a}} & b_a^{\bar{a}} \end{pmatrix}.$$

В лоренцевом базисе скалярное произведение $e_{\bar{\mu}} \cdot e_{\bar{a}} = 0$ по определению. Следовательно, $b_a^{\bar{\mu}} = 0$, а матрица

$$\begin{pmatrix} h_{\bar{\mu}}^{\mu} & h_{\bar{\mu}}^a \\ h_{\bar{a}}^{\mu} & h_{\bar{a}}^a \end{pmatrix}$$

представляет собой общий вид обратной матрицы.

Компоненты метрического тензора G в пространстве \mathbb{M}_4 выражаем через коэффициенты разложения одного базиса по другому

$$\begin{aligned} G_{\mu\nu} &\equiv e_{\mu} \cdot e_{\nu} = b_{\mu}^{\bar{A}} b_{\nu}^{\bar{C}} e_{\bar{A}} e_{\bar{C}} = b_{\mu}^{\bar{\mu}} b_{\nu}^{\bar{\nu}} e_{\bar{\mu}} e_{\bar{\nu}} + b_{\mu}^{\bar{a}} b_{\nu}^{\bar{c}} e_{\bar{a}} e_{\bar{c}} = \\ &= b_{\mu}^{\bar{\mu}} b_{\nu}^{\bar{\nu}} \eta_{\bar{\mu}\bar{\nu}} + b_{\mu}^{\bar{a}} b_{\nu}^{\bar{c}} \eta_{\bar{a}\bar{c}} = g_{\mu\nu} - b_{\mu}^{\bar{a}} b_{\nu}^{\bar{a}}. \end{aligned} \quad (9.32)$$

Кроме того,

$$G_{\mu a} \equiv e_{\mu} \cdot e_a = b_{\mu}^{\bar{A}} b_a^{\bar{C}} e_{\bar{A}} e_{\bar{C}} = b_{\mu}^{\bar{a}} b_a^{\bar{c}} e_{\bar{a}} e_{\bar{c}} = -b_{\mu}^{\bar{a}} b_a^{\bar{a}} \quad (9.33)$$

и

$$G_{ac} \equiv e_a \cdot e_c = b_a^{\bar{a}} b_c^{\bar{c}} e_{\bar{a}} e_{\bar{c}} = -b_a^{\bar{a}} b_c^{\bar{a}}. \quad (9.34)$$

Здесь использовалось основное свойство лоренцева базиса, $e_{\bar{A}}e_{\bar{C}} = \eta_{\bar{A}\bar{C}}$. Определим поле B_μ следующим образом:

$$b_\mu^{\bar{a}} = b_a^{\bar{a}} B_\mu^a,$$

или в явном виде

$$B_\mu \equiv B_\mu^a \equiv h_a^{\bar{a}} b_\mu^{\bar{a}}.$$

Кроме того, полезно ввести обозначение

$$G_{ac} \equiv \phi_{ac},$$

напоминающее, что это выражение является скаляром относительно преобразований в \mathbb{M}_4 . Подставив это определение в (9.32), (9.33), (9.34), получим искомые формулы для расщепления метрического тензора многомерного пространства

$$G_{\mu\nu} = g_{\mu\nu} + \phi_{ac} B_\mu^a B_\nu^c, \quad (9.35)$$

$$G_{\mu a} = \phi_{ac} B_\mu^c. \quad (9.36)$$

Выпишем явный вид координатных преобразований этих величин. А именно, преобразование, аналогичное 5-мерному случаю

$$x' = x'(x), \quad y' = y'(x, y),$$

приводит к следующим трансформациям:

$$g'_{\mu\nu} = \frac{\partial x^\lambda}{\partial x'^\mu} \frac{\partial x^\rho}{\partial x'^\nu} g_{\lambda\rho}, \quad G'_{ac} = \frac{\partial y^d}{\partial y'^a} \frac{\partial y^f}{\partial y'^c} G_{df}, \quad (9.37)$$

$$B_\mu'^a = \frac{\partial x^\nu}{\partial x'^\mu} \frac{\partial y'^a}{\partial y^c} B_\nu^c + \frac{\partial y'^a}{\partial y^c} \frac{\partial y^c}{\partial x^\nu}. \quad (9.38)$$

Все эти формулы справедливы и без учета симметрии дополнительного пространства. Ранее, при исследовании 5-мерной модели, можно было не заботиться о форм-инвариантности метрики дополнительного пространства относительно преобразования (9.14). В силу простоты геометрии этого пространства, вид

метрики дополнительного пространства \mathbb{M}_1 не менялся при применении преобразований (9.14). Если же дополнительное пространство является многомерным, то произвольные преобразования координат, вообще говоря, влияют на его метрику. По этой причине в низкоэнергетическом пределе лагранжиан не будет иметь стандартного калибровочно-инвариантного вида, хотя недиагональные компоненты метрического тензора будут ассоциироваться с векторными калибровочными полями.

Для того чтобы получить калибровочно инвариантную теорию при низких энергиях, необходимо, чтобы дополнительное пространство обладало соответствующей симметрией. Тогда суженный класс преобразований метрики этого пространства, а именно, оставляющий метрику форм-инвариантной, порождает калибровочно-инвариантную теорию. Посмотрим, к чему приводит наличие такой симметрии.

Векторы Киллинга

Ранее было показано, что локальные преобразования только дополнительного пространства \mathbb{M}_d

$$y^a \rightarrow y'^a = y^a + \epsilon^a(x) \quad (9.39)$$

можно рассматривать как преобразование калибровочных полей в гравитационном поле, по крайней мере, в 5-мерном случае. Вообще говоря, при этом меняется вид метрики G_{AB} как основного пространства \mathbb{M}_4 , так и дополнительного \mathbb{M}_d . Если в основном пространстве удастся выделить инвариантную часть, $g_{\mu\nu}$, то преобразование метрического тензора дополнительного пространства является нежелательным атрибутом. Поэтому стоит обратить внимание на координатные преобразования, оставляющие метрику дополнительного пространства неизменной.

Преобразование координат $y \rightarrow y'$ называется изометрией, если оно оставляет метрику форм-инвариантной,

$$G_{ab}(y) \rightarrow G'_{ab}(y') = G_{ab}(y'). \quad (9.40)$$

Если такие преобразования координат образуют группу, то она называется группой изометрии $\mathcal{G}(\mathbb{M}_d)$ пространства \mathbb{M}_d . Нас будут интересовать бесконечно малые преобразования.

Пусть имеется набор векторов $k_i^m(y)$, $i = 1, 2, \dots, r$, малые смещения вдоль которых

$$y^m \rightarrow y'^m = y^m + k_i^m(y)\varepsilon^i(x), \varepsilon \ll 1, \quad (9.41)$$

оставляют метрику пространства \mathbb{M}_d неизменной. Векторы $k_i^m(y)$ называются векторами Киллинга и являются генераторами изометрии данного пространства. Можно показать, что эти вектора удовлетворяют коммутационным соотношениям

$$k_i^a \partial_a k_j^b - k_j^a \partial_a k_i^b = f_{ij}^l k_l^b. \quad (9.42)$$

Структурные константы f_{ij}^l определяются конкретной группой изометрии. Максимально возможное число векторов Киллинга связано с размерностью пространства следующим образом

$$\max A = \frac{d(d+1)}{2}. \quad (9.43)$$

Уравнения для определения векторов Киллинга могут быть получены из (9.40) и имеют вид

$$\nabla_a k_{ib} + \nabla_b k_{ia} = 0, \quad k_{ia} \equiv G_{ab} k_i^b(y) \quad (9.44)$$

(уравнения Киллинга).

Неабелева калибровочная группа как группа изометрии дополнительного пространства

Применим вышеприведенные формулы для исследования связи симметрии дополнительного пространства с калибровочной инвариантностью низкоэнергетической теории в четырех измерениях. Выбираем зависимость метрики от координат

$$g_{\mu\nu} = g_{\mu\nu}(x), \quad G_{mn} = G_{mn}(y). \quad (9.45)$$

Далее, разложим вектор B_μ^m по базису векторов Киллинга k_i^n :

$$B_\mu^m = k_i^m(y)B_\mu^i(x). \quad (9.46)$$

Это можно сделать, поскольку B_μ^m является вектором в пространстве \mathbb{M}_d , обладающем соответствующей изометрией.

Теперь можно показать, что группа изометрии в \mathbb{M}_d реализуется в \mathbb{M}_4 как неабелева калибровочная группа. Для этого рассмотрим инфинитезимальные преобразования

$$x'^\mu = x^\mu, \quad y'^\alpha = y^\alpha + k_i^\alpha(y)\varepsilon^i(x).$$

После некоторых выкладок, из преобразований (9.37) следует закон преобразования объекта B_μ^i

$$\delta B_\mu^i = -\partial_\mu \varepsilon^i + f_{jl}^i B_\mu^j \varepsilon^l.$$

Это как раз то, что называется калибровочными преобразованиями группы G . Метрика дополнительного пространства $G_{ab}(y)$ инвариантна относительно этих преобразований, поскольку k_α^i есть векторы Киллинга дополнительного пространства.

Гравитационное действие в D мерном пространстве есть

$$S_G = \frac{1}{2\kappa_D} \int d^D X \sqrt{G}(R_D - 2\Lambda). \quad (9.47)$$

Длительные вычисления приводят к следующему виду скаляра Риччи:

$$R_D = R_4 + R_d - \frac{1}{4} G_{ab} k_i^a k_j^b F_{\mu\nu}^i F^{j\mu\nu}, \quad (9.48)$$

$$F_{\mu\nu}^i = \partial_\mu B_\nu^i - \partial_\nu B_\mu^i - f_{jl}^i B_\mu^j B_\nu^l, \quad F^{j\mu\nu} = g^{\mu\lambda} g^{\nu\rho} F_{\lambda\rho}^j. \quad (9.49)$$

Подставляя это выражение в действие, получаем

$$S_G = \frac{1}{2\kappa} \int d^4 x d^D y \sqrt{-g} \sqrt{G_d} (R_4 + R_d - \frac{1}{4} G_{ab} k_i^a k_j^b F_{\mu\nu}^i F^{j\mu\nu} - 2\Lambda).$$

Теперь воспользуемся тем, что рассматривается случай, когда метрика дополнительного пространства $G_{ab}(y)$ не зависит

от x . Это позволяет нормировать векторы Киллинга следующим образом:

$$\frac{1}{2\kappa} \int d^d y \sqrt{G} G_{ab} k_i^a(y) k_j^b(y) = \delta_{ij}.$$

Таким образом, чисто гравитационное действие в D мерном пространстве описывает гравитацию и калибровочное поле в четырехмерном пространстве

$$S = \int d^4 x \sqrt{-g} \left[\frac{-1}{2\kappa} R_4 - \frac{1}{4} F_{\mu\nu}^i(x) F^{i\mu\nu}(x) - 2\Lambda' \right]. \quad (9.50)$$

9.3. Нелинейная гравитация в моделях с одним дополнительным пространством

9.3.1. $F(R)$ -теория. Основные уравнения

Рассмотрим $(D = d_0 + d_1)$ -мерное многообразие \mathbb{M} , имеющее простейшую геометрическую структуру прямого произведения $\mathbb{M} = \mathbb{M}_0 \times \mathbb{M}_1$, с метрикой

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu + e^{2\beta(x)} b_{ab} dx^a dx^b, \quad (9.51)$$

где компоненты метрики в дополнительных измерениях b_{ab} не зависят от x^μ , координат наблюдаемого пространства-времени. Для реалистических моделей $d_0 = 4$, однако в формулах (2.37)–(9.55) оставлено произвольное d_0 для общности.

D -мерный тензор Римана имеет следующие ненулевые компоненты:

$$\begin{aligned} R^{\mu\nu}{}_{\rho\sigma} &= \bar{R}^{\mu\nu}{}_{\rho\sigma}, \\ R^{\mu a}{}_{\nu a} &= \delta_b^a B_\nu^\mu, \quad B_\nu^\mu := e^{-\beta} \nabla_\nu (e^\beta \beta^\mu), \\ R^{ab}{}_{cd} &= e^{-2\beta} \bar{R}^{ab}{}_{cd} + \delta_{cd}^{ab} \beta_\mu \beta^\mu, \end{aligned} \quad (9.52)$$

где заглавные латинские индексы пробегают все D значений, чертой сверху отмечены величины, полученные отдельно из $g_{\mu\nu}$

и b_{ab} , $\beta_\mu \equiv \partial_\mu \beta$ и $\delta^{ab}_{cd} \equiv \delta^a_c \delta^b_d - \delta^a_d \delta^b_c$. Ненулевые компоненты тензора Риччи и скалярная кривизна имеют вид

$$\begin{aligned} R^\nu_\mu &= \bar{R}^\nu_\mu + d_1 B^\nu_\mu, \\ R^b_a &= e^{-2\beta} \bar{R}^b_a + \delta^b_a [\square \beta + d_1 (\partial\beta)^2], \\ R &= \bar{R}[g] + e^{-2\beta} \bar{R}[b] + 2d_1 \square \beta + d_1 (d_1 + 1) (\partial\beta)^2, \end{aligned} \quad (9.53)$$

где $(\partial\beta)^2 \equiv \beta_\mu \beta^\mu$, $\square = g^{\mu\nu} \nabla_\mu \nabla_\nu - d_0$ -мерный оператор д'Аламбера, $\bar{R}[g]$ и $\bar{R}[b]$ — скаляры Риччи, относящиеся, соответственно, к $g_{\mu\nu}$ и b_{ab} .

Предположим, что b_{ab} описывает компактное d_1 -мерное пространство ненулевой постоянной кривизны, т.е. сферу ($k = 1$) или компактное d_1 -мерное гиперболическое пространство¹⁴ ($k = -1$) с фиксированным радиусом кривизны r_0 , нормированным на D -мерный аналог m_D планковской массы, т.е., $r_0 = 1/m_D$ (здесь, как и всюду, используются естественные единицы, в которых скорость света c и постоянная Планка \hbar равны единице), имеем:

$$\begin{aligned} \bar{R}^{ab}_{cd} &= k m_D^2 \delta^{ab}_{cd}, \\ \bar{R}^b_a &= k m_D^2 (d_1 - 1) \delta^b_a, \\ \bar{R}[b] &= k m_D^2 d_1 (d_1 - 1) = R_b. \end{aligned} \quad (9.54)$$

Масштабный фактор $b(x) \equiv e^\beta$ в (9.51) остается безразмерным; R_b имеет смысл характерного масштаба кривизны дополнительных измерений.

Рассмотрим в описанной геометрии достаточно общую нелинейную по кривизне теорию гравитации с действием

$$S = \frac{1}{2} m_D^{D-2} \int \sqrt{Dg} d^D x [F(R) + L_m], \quad (9.55)$$

¹⁴Компактные гиперболические пространства постоянной кривизны на основе обычного пространства Лобачевского \mathbb{H}^d изометричны \mathbb{H}^d/Γ , где Γ — некоторая нетривиальная дискретная группа изометрий \mathbb{H}^d , см., напр., [10]. О возможных применениях таких (трехмерных) пространств в космологии см., напр., [230] и приведенные там ссылки.

где $F(R)$ произвольная гладкая функция, L_m — лагранжиан материи и ${}^D g = |\det(g_{MN})|$. Дополнительные координаты легко устраняются интегрированием, и действие представляется в четырехмерном виде, $d_0 = 4$:

$$S = \frac{1}{2} \mathcal{V}[d_1] m_D^2 \int \sqrt{{}^4 g} d^4 x e^{d_1 \beta} [F(R) + L_m], \quad (9.56)$$

где ${}^4 g = |\det(g_{\mu\nu})|$ и $\mathcal{V}[d_1]$ — объем компактного d_1 -мерного пространства единичной кривизны.

9.3.2. Приближение медленных изменений. Картина Эйнштейна

Формула (9.56) описывает четырехмерную гравитацию, нелинейную по кривизне и содержащую неминимальную связь между эффективным скалярным полем β и кривизной. Упростим ее следующим образом:

(а) Выразим все величины через четырехмерные переменные и $\beta(x)$; отметим, что теперь

$$\begin{aligned} R &= R_4 + \phi + f_1, \\ R_4 &= \bar{R}[g], \quad f_1 = 2d_1 \square \beta + d_1(d_1 + 1)(\partial\beta)^2, \end{aligned} \quad (9.57)$$

где мы ввели эффективное скалярное поле

$$\phi(x) = R_b e^{-2\beta(x)} = k d_1 (d_1 - 1) m_D^2 e^{-2\beta(x)}. \quad (9.58)$$

Поскольку $k = \pm 1$ соответственно для положительной и отрицательной кривизны в d_1 дополнительных измерениях, поле ϕ в этих случаях, по определению, имеет разные знаки.

(б) Предположим, что все величины меняются медленно, т.е. рассматриваем каждую производную ∂_μ (включая те, которые входят в определение \bar{R}) как выражения, содержащие малый параметр ε ; отбросим все величины порядков выше $O(\varepsilon^2)$ (see [135]).

(с) Осуществим конформное отображение, приводящее к картине Эйнштейна, в которой четырехмерная кривизна минимально связана со скалярным полем.

В разложении (9.57) оба члена f_1 и R_4 считаются в нашем подходе малыми, это фактически означает, что все величины, включая четырехмерную кривизну, малы по сравнению с D -мерной планковской шкалой. Лишь один член, ϕ , не мал, и можно использовать тейлоровское разложение функции $F(R) = F(\phi + R_4 + f_1)$:

$$F(R) = F(\phi + R_4 + f_1) \simeq F(\phi) + F'(\phi) \cdot (R_4 + f_1) + \dots, \quad (9.59)$$

где $F'(\phi) \equiv dF/d\phi$. Подставляя это в (9.56), получим с точностью до $O(\varepsilon^2)$

$$S = \frac{1}{2} \mathcal{V}[d_1] m_D^2 \int \sqrt{^4g} d^4x e^{d_1\beta} [F'(\phi)R_4 + F(\phi) + F'(\phi)f_1 + L_m], \quad (9.60)$$

где β связана с ϕ согласно (9.58). Выражение (9.60) типично для скалярно-тензорной теории (СТТ) теории гравитации в картине Йордана.

Для анализа динамики скалярного поля и, в частности, для поиска стационарных точек имеет смысл перейти к картине Эйнштейна, в которой поле ϕ взаимодействует с гравитацией минимальным образом. После конформного преобразования

$$g_{\mu\nu} \mapsto \tilde{g}_{\mu\nu} = |f(\phi)|g_{\mu\nu}, \quad f(\phi) = e^{d_1\beta} F'(\phi), \quad (9.61)$$

с соответствующим преобразованием скалярной кривизны

$$R_4 = |f| \left(\tilde{R}_4 + \frac{3}{f} \tilde{\square} f - \frac{3}{2f^2} (\tilde{\partial} f)^2 \right) \quad (9.62)$$

(тильдой отмечены величины, полученные с помощью $\tilde{g}_{\mu\nu}$), действие (9.60) принимает вид

$$S = \frac{1}{2} \mathcal{V}[d_1] m_D^2 \int \sqrt{\tilde{g}} d^4x \left\{ [\text{sign } F'(\phi)] [\tilde{R}_4 + K] - \right.$$

$$-V(\phi) + \frac{e^{-d_1\beta}}{F'(\phi)^2} L_m \} \quad (9.63)$$

с кинетическим (K) и потенциальным (V) членами

$$\begin{aligned} K &= \frac{3}{2} \frac{(\partial f)^2}{f^2} - 2d_1 \frac{f^{\cdot\mu} \beta_\mu}{f} + (\partial\beta)^2, \\ V(\phi) &= -e^{-d_1\beta} F(\phi)/F'(\phi)^2. \end{aligned} \quad (9.64)$$

Остается выразить все через один скаляр, скажем, ϕ . Можно записать действие (9.63) в виде

$$\begin{aligned} S &= \frac{\mathcal{V}[d_1]}{2} m_D^2 \int \sqrt{\tilde{g}} (\text{sign } F') L, \\ L &= \tilde{R}_4 + \frac{1}{2} K_{\text{Ein}}(\phi) (\partial\phi)^2 - V_{\text{Ein}}(\phi) + \tilde{L}_m, \\ \tilde{L}_m &= (\text{sign } F') \frac{e^{-d_1\beta}}{F'(\phi)^2} L_m; \end{aligned} \quad (9.65)$$

$$K_{\text{Ein}}(\phi) = \frac{1}{2\phi^2} \left[6\phi^2 \left(\frac{F''}{F'} \right)^2 - 2d_1 \phi \frac{F''}{F'} + \frac{1}{2} d_1 (d_1 + 2) \right], \quad (9.66)$$

$$V_{\text{Ein}}(\phi) = -(\text{sign } F') \left[\frac{|\phi| m_D^{-2}}{d_1 (d_1 - 1)} \right]^{d_1/2} \frac{F(\phi)}{F'(\phi)^2}. \quad (9.67)$$

Во всех формулах индексы поднимаются и опускаются с помощью метрического тензора $\tilde{g}_{\mu\nu}$; всюду $F = F(\phi)$ и $F' = dF/d\phi$.

Формулы (9.65)–(9.67) верны как для положительной ($\phi > 0$), так и для отрицательной ($\phi < 0$) кривизны в дополнительных измерениях. Минимумы потенциала V_{Ein} , определенного в картине Эйнштейна, определяют стационарные точки скалярного поля, и они остаются стационарными и после перехода в другую конформную калибровку (картину), включая исходную йордановскую. Именно картина Эйнштейна определяет поведение скалярного поля, т.к. его динамика около экстремума прямо диктуется видом потенциала $V(\phi)$ и только опосредованно метрикой.

Если ϕ находится в минимуме потенциала V_{Ein} , этот потенциал становится эффективной космологической постоянной. В применении к современной эпохе он может определять плотность темной энергии, приводящей к ускоренному расширению Вселенной. В других моделях он может вызывать инфляцию в ранней Вселенной.

Потенциал (9.67) для произвольной функции $F(R)$ достаточно сложен и может иметь нетривиальные экстремумы. Сложный вид оказался и у кинетического члена (9.66). Как известно, его свойства так же важны для динамики поля, как и форма потенциала. Например, как показано в [203], нули и сингулярности кинетического члена могут описывать стабильные состояния скалярного поля.

Как эффективный потенциал (9.67), так и кинетический член (9.66) сингулярны при значениях ϕ , где величина $f(\phi) \sim F'(R)$ [множитель при R_4 в действии (9.60)] равна нулю. В отличие от многих работ, в которых вводится ограничение $F'(R) > 0$, мы включаем рассмотрение и модели с $F'(R) < 0$. Как мы увидим, это открывает новые интересные возможности получения минимумов эффективного потенциала, на которых стабилизируются дополнительные измерения.

В формулах (9.65)–(9.67) мы фактически изменили знак лагранжиана в случае $F' < 0$; чтобы сохранить притягивающий характер гравитации для обычной материи, необходимо, чтобы лагранжиан материи фигурировал в действии с самого начала с необычным знаком. В результате будет необычный знак полного действия материи и гравитации, что, однако, безразлично для уравнений движения материи; сохранится и обычный вид эффективных уравнений движения в стационарных точках поля ϕ . Действие же поля ϕ будет необычным согласно (9.66) и (9.67). Следует отметить, что общий знак полного действия не влияет и на квантовые переходы. В самом деле, амплитуда перехода в технике интегрирования по путям выражается как $\int \exp(iS[q])Dq$, где $q(t)$ — какая-либо динамическая переменная. Вероятность

перехода

$$\int \int \exp(iS[q_1] - iS[q_2]) Dq_1 Dq_2$$

инвариантна относительно замены $i \rightarrow -i$ (вместе с перестановкой переменных интегрирования q_1 и q_2).

Здесь необходимо сделать следующее замечание. Известно конформное преобразование [57, 58, 136, 142, 184, 221, 222, 310], приводящее действие (9.55) к виду, характерному для теории Эйнштейна с минимально связанным скалярным полем χ (вместо дополнительной степени свободы, связанной с R), потенциал которого выражается через функцию $F(R)$. Это D -мерная картина Эйнштейна. Полученную теорию можно далее редуцировать к четырем измерениям и еще одним конформным преобразованием, теперь зависящим от $\beta(x)$, к четырехмерной картине Эйнштейна. Это делалось, в частности, в статье [156, 157] для квадратичной гравитации. Наш метод технически проще и более универсален, так как (см. далее) позволяет работать и с другими инвариантами, не только $F(R)$. В случаях, когда действует наше приближение, результаты их совпадают, как показано в [98].

9.3.3. Некоторые оценки

Для сопоставления моделей с реальностью необходимо выбрать физическую (наблюдательную) конформную картину. Предположим, что исходная картина Йордана, в которой сформулирована теория, совпадает с наблюдательной, в которой измерительные эталоны не меняются в пространстве и времени. Общее обсуждение проблемы выбора физической конформной картины см. в работе [97].

Картину Эйнштейна, таким образом, мы используем как удобный инструмент для поиска стационарных состояний.

В наблюдательной картине, в предполагаемом стационарном состоянии, гравитационное действие имеет приближенный вид

действия Гильберта-Эйнштейна

$$S_{\text{GR}} = \frac{1}{2} m_4^2 \int \sqrt{^4g} d^4x (R_4^* - 2\Lambda_{\text{eff}}), \quad (9.68)$$

где $m_4 = (8\pi G_N)^{-1/2}$ — четырехмерная планковская масса, G_N — эффективная ньютоновская постоянная и R_4^* — наблюдаемая четырехмерная кривизна (при сделанных предположениях, $R_4^* = R_4$). Чтобы это стационарное состояние согласовывалось с нашим подходом и наблюдательными ограничениями, оно должно удовлетворять следующим требованиям:

- (i) Допустимость классического описания пространства-времени, т.е. истинный размер дополнительных измерений должен быть больше истинной D -мерной планковской длины $1/m_D$:

$$m_D b_0 = e^{\beta_0} \gg 1, \quad \Leftrightarrow \quad \phi/m_D^2 \ll d_1(d_1 - 1). \quad (9.69)$$

- (ii) Должно работать приближение медленных изменений, что, при $\phi = \text{const}$, сводится к требованию $R_4 \ll m_D^2$.

- (iii) Наблюдаемый размер b^* дополнительных измерений должен быть много больше планковской длины $l_4 = 1/m_4 \approx 8 \cdot 10^{-33}$ см, но не больше $\sim 10^{-17}$ см, что соответствует энергиям порядка ТэВ.

- (iv) Должна предсказываться эффективная космологическая постоянная Λ_{eff} , согласующаяся с наблюдениями:

$$\Lambda_{\text{eff}}/m_4^2 \sim 10^{-120}. \quad (9.70)$$

Получение столь малых значений Λ — одна из известных проблем теоретической космологии, т.к. величина Λm_4^2 , обычно связываемая с плотностью энергии вакуума, должна быть на много порядков меньше характерных масштабов плотности энергии для известных взаимодействий: так, для гравитации это планковская плотность m_4^4 .

Итак, предполагаем, что наблюдения происходят в картине Йордана. Соответствующее действие (9.60) аппроксимируется (9.68) с эффективными константами

$$\frac{1}{8\pi G_N} = m_4^2 = \mathcal{V}[d_1] m_D^{D-2} b_0^{d_1} F'_0, \quad \Lambda_{\text{eff}} = -\frac{F(\phi_0)}{2F'_0}, \quad (9.71)$$

где $b_0 = e^{\beta(\phi_0)}/m_D = b_*$ — наблюдаемый размер дополнительных измерений и $F'_0 = F'(\phi_0)$.

Формула (9.71) приводит к соотношению между безразмерными величинами $b_0 m_D$ и $b_0 m_4$:

$$b_0^2 m_4^2 = \mathcal{V}[d_1] F'_0 (b_0 m_D)^{d_1+2}. \quad (9.72)$$

Множитель $\mathcal{V}[d_1]$ порядка; того же можно ожидать от безразмерной величины F'_0 . Согласно пункту (iii) (см. выше), $b_0^2 m_4^2 \lesssim 10^{30}$, и уравнение (9.72) дает

$$\begin{aligned} b_0 m_D &\lesssim 10^{30/(d_1+2)}, \\ m_4/m_D &\sim (b_0 m_D)^{d_1} \lesssim 10^{15d_1/(d_1+2)}. \end{aligned} \quad (9.73)$$

Таким образом, m_D не очень сильно отличается от m_4 . Это, в частности, означает, что наше приближение медленных изменений (пункт (ii)) работает почти во всех мыслимых ситуациях, так как R_4 — наблюдаемая кривизна. Можно, в частности, показать [98], что оно работает в моделях хаотической инфляции на масштабе энергий Большого объединения, когда $R_4/m_4^2 \sim 10^{-6}$. Тем более это приближение справедливо на современной стадии эволюции.

Вернемся к неравенствам (9.73) и оценим Λ_{eff} . Отношение $\Lambda_{\text{eff}}/m_4^2$ можно выразить следующим образом:

$$\begin{aligned} -\frac{\Lambda_{\text{eff}}}{m_4^2} &= \frac{F(\phi_0)}{2m_D^2 F'_0} \frac{m_D^2 b_0^2}{m_4^2 b_0^2} = \\ &= \frac{1}{2F'_0} \frac{F(\phi_0)}{m_D^2} (F'_0 \mathcal{V}[d_1])^{\frac{-2}{d_1+2}} (m_4 b_0)^{\frac{-2d_1}{d_1+2}}. \end{aligned} \quad (9.74)$$

Если пытаться сделать это отношение малым, последний множитель в (9.74) дает в лучшем случае 10^{-30} (при достаточно большом d_1). Остальные 90 порядков величины должны объясняться малостью безразмерного отношения $F(\phi_0)/m_D^2$ (фактически — исходной космологической постоянной) и/или большим значением F'_0 . Последний вариант, однако, неприемлем, так как тогда становится невозможным $b_0 m_D \gg 1$, см. (9.72).

Таким образом, проблема тонкой настройки остается актуальной, хотя и несколько смягчается по сравнению с исходной: 30 порядков из 120 набираются достаточно естественно.

9.3.4. Квадратичная гравитация с космологической постоянной

В дальнейшем рассматриваем чистую гравитацию ($L_m = 0$) и полагаем $m_D = 1$.

В качестве первого примера рассмотрим действие (9.55) с функцией

$$F(\phi) = \phi + c\phi^2 - 2\Lambda, \quad c, \Lambda = \text{const.} \quad (9.75)$$

Тогда (9.66) и (9.67) дают эффективный потенциал

$$V_{\text{Ein}}^{(2)}(\phi) = -\frac{\text{sign}(1 + 2c\phi)}{[d_1(d_1 - 1)]^{d_1/2}} |\phi|^{d_1/2} \frac{c\phi^2 + \phi - 2\Lambda}{(1 + 2c\phi)^2} \quad (9.76)$$

и коэффициент K_{Ein} кинетического члена

$$K_{\text{Ein}}^{(2)}(\phi) = \frac{1}{(1 + 2c\phi)^2 \phi^2} \cdot \left[c^2 \phi^2 (d_1^2 - 2d_1 + 12) + d_1^2 c \phi + \frac{1}{4} d_1 (d_1 + 2) \right]. \quad (9.77)$$

В работе [156, 157] рассматривался случай $F'(R) > 0$ и было показано, что искомый минимум потенциала возможен лишь при $V_{\text{Ein}} < 0$ и отрицательной кривизне, $k = -1$. В данном подходе это подтверждается, см. примеры на рис. 9.1. В космологическом контексте это приводит к пространству анти-де Ситтера.

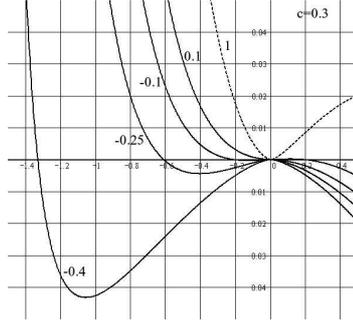


Рис. 9.1. Потенциал $V_{\text{Ein}}(\phi)$ для $d_1 = 3$, $c = 0.3$ в области $\phi > -5/3$, где $F'(R) > 0$, для некоторых значений Λ (показанных около кривых). Прерывистая линия ($\Lambda = 1$) имеет минимум при $\phi = 0$. Значения φ выражены в единицах $m_D = 1$.

Еще один минимум V_{Ein} при $F' > 0$, с $\Lambda > 0$, расположен в точке $\phi = 0$. Асимптотика $\phi \rightarrow 0$ соответствует растущему размеру дополнительных измерений: $b = e^\beta \sim 1/\sqrt{|\phi|} \rightarrow \infty$. В принципе всг же возможна модель, в которой этот процесс достаточно медленный и размер b к настоящему времени не достиг наблюдаемых значений. Проверка показывает, однако [98], что в такой модели эффективная гравитационная постоянная изменяется слишком быстро, почти в хаббловском темпе, следовательно, такая модель малоинтересна.

Мы также попытались найти минимумы потенциала в другой области, где $F' < 0$, в широком диапазоне значений параметров c и Λ . Примеры поведения потенциала при конкретном значении $c = 1.5$ показаны на рис.9.2. Потенциалы с $\Lambda < 0$ приводят к минимумам $V_{\text{min}} > 0$, так что $\Lambda_{\text{eff}} > 0$, так что возможны космологические модели с внешним пространством де Ситтера и стабильными дополнительными измерениями. Такие модели снова соответствуют $\phi < 0$, т.е. гиперболическим дополнительным измерениям. Тонкой настройкой констант c и Λ можно получить современные значения космологических параметров. Как и на

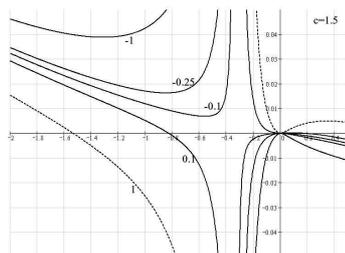


Рис. 9.2. То же, что на рис. 9.1, но с $c = 1.5$. Минимумы $V(\phi)$ расположены в области $\phi < -1/3$, где $F' < 0$.

рис.9.3.4, можно найти минимум при $\phi = 0$, свойства которого уже обсуждались.

Таким образом, система ведет себя существенно по-разному в различных областях значений ϕ и в зависимости от значений исходных параметров. Подтверждается, что чистая гравитация способна стабилизировать размер дополнительных измерений без включения других полей со специальным видом потенциалов. Положительные значения Λ_{eff} (плотности энергии вакуума) найдены в области, где $F' < 0$. Это, однако, лишь свойство квадратичной гравитации: как будет видно далее, в кубической гравитации можно получить $\Lambda_{\text{eff}} > 0$ в области, где $F' > 0$.

9.3.5. Кубическая гравитация

Рассмотрим действие (9.55) с функцией

$$F(R) = R + cR^2 + CR^3 \quad (9.78)$$

В этом случае вид потенциала оказывается весьма чувствительным к значениям параметров c и C . Этим можно воспользоваться, чтобы получить минимумы потенциала в картине Эйнштейна, приводящие к очень малым, но положительным значениям Λ_{eff} в картине Йордана, и, в отличие от квадратичной гравитации, это осуществимо в области, где $F' > 0$. Примеры такого поведения показаны на рис.9.3.5. Таким образом, кубическая

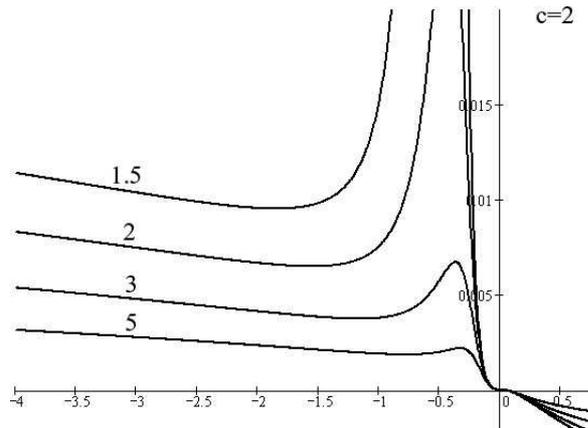


Рис. 9.3. Потенциал V_{Ein} для кубической гравитации с $d_1 = 3$ и $c = 2$. Кривые помечены значениями параметра C . Около минимумов $F'(\phi) > 0$.

гравитация в принципе пригодна для описания нашей Вселенной с компактными и стабильными дополнительными измерениями. При больших временах, когда можно пренебречь влиянием материи по сравнению с эффективной космологической постоянной, получим модель де Ситтера в трех наблюдаемых измерениях.

Область положительных кривизн, $\phi > 0$, оказывается менее интересной. Она также содержит минимумы, но с $V_{\text{min}} < 0$, см. рис.9.3.5. Состояния в этих минимумах метастабильны, т.к. минимумы лишь локальные, правда, время жизни таких состояний может быть сколь угодно большим при подходящих значениях параметров.

9.3.6. Обобщения

Описанный метод позволяет рассматривать более широкие классы лагранжианов. Продемонстрируем это, добавив к квадратичному лагранжиану члены, пропорциональные квадрату тензора Риччи $R_{AB}R^{AB}$ и скаляру Кречмана $\mathcal{K} = R_{ABCD}R^{ABCD}$. По су-

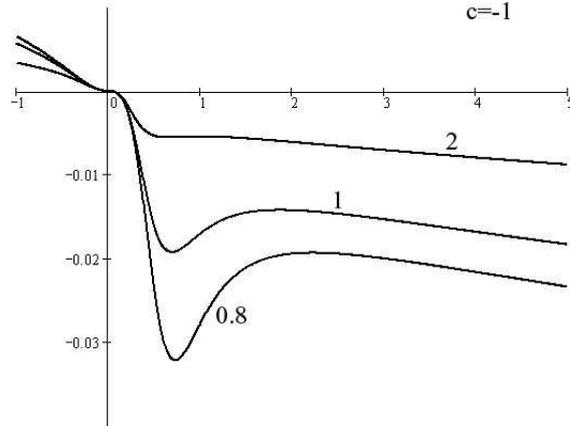


Рис. 9.4. $V_{\text{Ein}}(\phi)$ для кубической гравитации с $d_1 = 3$, $c = -1$. Кривые помечены значениями параметра C . Снова около минимумов $F'(\phi) > 0$.

шествующим взглядам, эти и другие инварианты кривизны высших порядков должны появляться за счет квантовых поправок, и их естественно включить в рассмотрение наряду с cR^2 . Действие принимает вид

$$S = \frac{\mathcal{V}[d_1]}{2\kappa^2} \int d^4x \sqrt{4g} e^{d_1\beta} (R + cR^2 + c_1 R_{AB}R^{AB} + c_2 \mathcal{K} - 2\Lambda), \quad (9.79)$$

где по внутренним переменным проведено интегрирование, аналогично (9.55). Для метрики (9.51) несложно получить выражения для $R_{AB}R^{AB}$ и \mathcal{K} :

$$R_{AB}R^{AB} = \bar{R}_{\mu\nu}\bar{R}^{\mu\nu} + 2d_1\bar{R}_{\mu\nu}B^{\mu\nu} + e^{-4\beta}\bar{R}_{ab}\bar{R}^{ab} + 2e^{-2\beta}\bar{R}[h][\square\beta + d_1(\partial\beta)^2] + d_1[\square\beta + d_1(\partial\beta)^2]^2, \quad (9.80)$$

$$\mathcal{K} = \bar{\mathcal{K}}[g] + 4d_1B_{\mu\nu}B^{\mu\nu} + e^{-4\beta}\bar{\mathcal{K}}[h] + 4e^{-2\beta}\bar{R}[h](\partial\beta)^2 + 2d_1(d_1 - 1)[(\partial\beta)^2]^2. \quad (9.81)$$

В приближении медленных изменений тем же способом, что и ранее, получаем четырехмерный лагранжиан

$$\sqrt{4g}L = \sqrt{4g}e^{d_1\beta} \left[R_4(1+2c\phi) + \phi + c_{\text{tot}}\phi^2 - 2\Lambda + (1+2c\phi)f_1 + 2c_1\phi \square \beta + 2(c_1d_1 + 2c_2)(\partial\beta)^2 \right], \quad (9.82)$$

где

$$c_{\text{tot}} = c + \frac{c_1}{d_1} + \frac{2c_2}{d_1(d_1-1)}.$$

Конформное преобразование (9.61) приводит к лагранжиану (9.65) в картине Эйнштейна с кинетическим и потенциальным членами

$$K_{\text{Ein}}(\phi) = K_{\text{Ein}}^{(2)}(\phi) + \frac{c_1 + c_2}{2\phi(1+2c\phi)}, \quad (9.83)$$

$$V_{\text{Ein}}(\phi) = -\text{sign}(1+2c\phi) \cdot \left[\frac{|\phi|}{d_1(d_1-1)} \right]^{d_1/2} \frac{c_{\text{tot}}\phi^2 + \phi - 2\Lambda}{(1+2c\phi)^2}, \quad (9.84)$$

где член $K_{\text{Ein}}^{(2)}(\phi)$ берется из (9.77).

Появление параметров c_1 и c_2 добавляет свободу выбора потенциала. Кинетический член также приобретает более сложный вид и может существенно влиять на динамику скалярного поля [203]: его нули могут представлять собой стационарные точки ϕ . Поле ϕ может захватываться окрестностями таких точек, в дополнение к минимумам потенциала.

В частности, как показано в [93], становятся жизнеспособными космологические модели со скалярным полем $\phi \rightarrow 0$ при $t \rightarrow \infty$, типа обсуждавшихся в разделе 9.3.4. В результате получается класс пространственно плоских космологических моделей, в которых как наблюдаемый масштабный фактор $a(\tau)$, так и масштабный фактор дополнительных измерений $b(\tau)$ растут экспоненциально в позднюю эпоху, однако $b(\tau)$ растет достаточно медленно, чтобы зависящие от него вариации эффективной

гравитационной постоянной G оставались в пределах, совместимых с наблюдениями. Такие модели предсказывают значительные изменения физических законов в отдаленном будущем за счет дальнейшего роста дополнительных измерений.

Чтобы удовлетворить жестким экспериментальным ограничениям на временную вариацию эффективной гравитационной постоянной G , параметры теории (9.79) должны удовлетворять ограничению [93]

$$12 \left[\frac{d_1 + 2}{d_1} + \frac{8(c_1 + c_2)}{d_1^2} \right] \gtrsim 100. \quad (9.85)$$

Таким образом, если в исходной теории оставить только $F(R)$, ограничению (9.85) удовлетворить невозможно: чтобы это сделать, необходимо привлечь квадрат тензора Риччи и/или скаляр Кречмана с исходными константами c_1 и c_2 , такими, что

$$c_1 + c_2 \gtrsim d_1^2. \quad (9.86)$$

При условии (9.86) модель потенциально жизнеспособна, поскольку она сочетает деситтеровское расширение наблюдаемой Вселенной с достаточно медленными вариациями гравитационной постоянной.

9.3.7. Интенсивное рождение частиц в постинфляционный период

Обратимся к еще одной проблеме, которая может быть успешно решена в рамках рассматриваемого класса теорий.

Согласно работе [130], быстрые осцилляции поля инфлатона сразу по окончании инфляции порождают частицы, которые и составляют материю современной Вселенной. Известно [196, 277], что механизм производства частиц и энтропии, а также нагрев Вселенной может быть неэффективным, если константа взаимодействия инфлатона с полями материи мала. С другой стороны, сильная связь приводит к значительным квантовым поправкам

к параметрам инфлатонного потенциала, что делает сомнительным предположение об их малости, вытекающее из сравнения с наблюдательными данными по флуктуациям температуры реликтового излучения. Эта проблема решается в рамках, например, гибридной инфляции, включающей еще одно гипотетическое скалярное поле (см. гл. 6). Во время инфляции плотность энергии медленно меняется, пока классическое значение поля не достигает точки бифуркации, после чего поле начинает быстро осциллировать около минимума потенциала, порождая требуемое число частиц. К сожалению, эта модель приводит также и к перепроизводству черных дыр, что является серьезным недостатком [265]. Другой перспективный механизм эффективного рождения частиц (параметрический резонанс) описан в [196]. Нелинейная многомерная гравитация предоставляет еще одну возможность.

Рассмотрим потенциал и кинетический член эффективного лагранжиана (9.65) с функцией $F(R)$ вида

$$F(R) = R + cR^2 - 2\Lambda. \quad (9.87)$$

Графики этих функций представлены на рис. 9.5. Нетривиальная зависимость кинетического члена от ϕ сильно влияет на динамику скалярного поля. Эффект особенно заметен, если минимум кинетического члена приближенно совпадает с минимумом потенциала, как на рис. 9.5.

Для иллюстрации рассмотрим упрощенную модель скалярного поля ϕ с потенциалом и кинетическим членом вида

$$\begin{aligned} V(\phi) &= \frac{1}{2}m^2\phi^2 \\ K(\phi) &= K_1 \cdot (\phi - \phi_{\min})^2 + K_{\min}, \quad K_1, K_{\min} > 0. \end{aligned} \quad (9.88)$$

(Функция $K(\phi)$ положительна и может быть сведена к единице переопределением поля $\phi \rightarrow \tilde{\phi}$, но такая эквивалентная формулировка теории менее удобна в данном случае.)

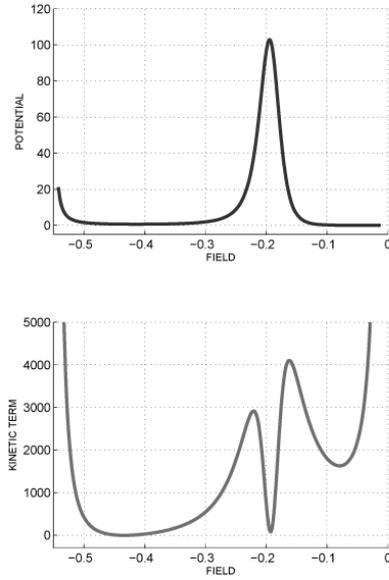


Рис. 9.5. Потенциал и кинетический член эффективной 4D теории для теории (9.55) с функцией $F(R)$ вида (9.87) и с параметрами $d = 3$, $c = 6.0$, $w_1\Lambda = 2$.

Численное решение соответствующих классических уравнений для пространственно плоской Вселенной

$$H^2 = \frac{\kappa^2}{3} \left[\frac{1}{2} K(\phi) \dot{\phi}^2 + V(\phi) \right],$$

$$K(\phi) \left(\ddot{\phi} + 3H\dot{\phi} \right) + \frac{1}{2} K_\phi(\phi) \dot{\phi}^2 + V_\phi(\phi) = 0 \quad (9.89)$$

(индекс ϕ означает $d/d\phi$, а H есть параметр Хаббла) представлено на рис.9.6 для квадратичного потенциала (9.88) и двух различных кинетических членов, с $K = 1/2$ и выписанного в (9.88). Значения параметров приведены в подписи к рисунку. Каждая

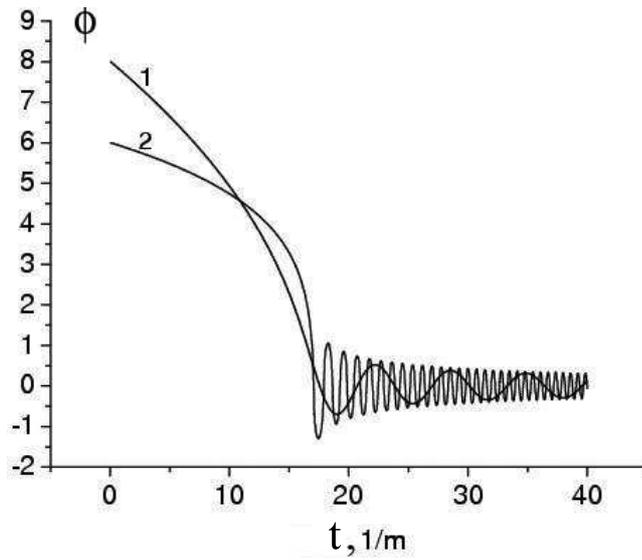


Рис. 9.6. Поведение инфлатона в конце инфляционной стадии для двух моделей. 1. Стандартный кинетический член, $K = 1/2$. 2. Кинетический член вида (9.88). Значения параметров $K_0 = 0.1$, $\phi_{\min} = 0.1$, $K_{\min} = 0.005$ даны в планковских единицах, время измеряется в обратных массах инфлатонного поля.

из кривых описывает поведение инфлатона до и после завершения инфляции. В нестандартном кинетическом члене $K(\phi)$ равно $1/2$ в конце инфляционной стадии (первая точка пересечения кривых на рис. 9.6). Очевидно, что в модели с переменным кинетическим членом инфляция начинается при меньших энергиях по сравнению с обычным случаем.

Отметим также разницу в частоте осцилляций инфлатонного поля. Число образованных частиц пропорционально этой частоте — см. обзор [132] и дополнительное обсуждение в [188]. Это заключение может быть дополнительно обосновано следующим образом.

По окончании инфляции амплитуда колебаний инфлатона

мала по сравнению с планковским масштабом, а коэффициент $K_{\text{eff}} \sim K_{\text{min}}$ в эффективном кинетическом члене мал благодаря выбору параметров. Эффективный Лагранжиан, содержащий взаимодействие инфлатона с неким скалярным полем χ

$$L_{\text{eff}} \simeq \frac{1}{2} K_{\text{min}} \dot{\phi}^2 + g \phi \chi \chi, \quad (9.90)$$

может быть преобразован к стандартному виду переопределением инфлатонного поля:

$$L_{\text{eff}} \simeq \frac{1}{2} \dot{\phi}^2 + \frac{g}{\sqrt{K_{\text{min}}}} \phi \chi \chi. \quad (9.91)$$

Малое значение K_{min} увеличивает константу связи (для выбранных параметров — на порядки) и, следовательно, приводит к интенсивному рождению частиц. Таким образом преодолевается трудность, обсуждавшаяся выше: слабая константа взаимодействия на инфляционной стадии не противоречит интенсивному производству частиц сразу по окончании инфляции.

9.3.8. Поведение поля вблизи $K = 0$

Функция $K(\phi)$ — коэффициент при кинетическом члене лагранжиана скалярного поля — вообще говоря, может иметь как нули, так и полюса, и их окрестности представляют особый интерес. Обсудим возможное состояние системы около нулевого значения $K(\phi)$ [203]. Если это простой нуль, при его переходе кинетический член меняет знак. Случай $K(\phi) = K_{\text{Ein}} > 0$ соответствует нормальному поведению скалярного поля, а случай $K(\phi) < 0$ отвечает фантомному полю с отрицательной кинетической энергией.

Вообще говоря, теории вида $F(R)$ позволяют получить лишь положительные значения $K(\phi)$. Действительно, как нетрудно видеть, выражение в квадратных скобках в формуле (9.66) всегда положительно с минимумом, равным $d(d+3)/3$.

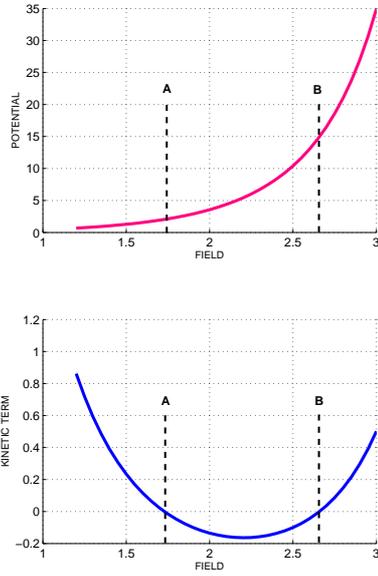


Рис. 9.7. Потенциал (9.94) и кинетический член (9.95) для теории (9.93) и $F(R)$ из (9.92). Значения параметров: $d = 3, c = -0.12, \Lambda = 2, c_1 + c_2 = -5.5, c' = -0.626$.

В более общих теориях, содержащих другие инварианты кривизны, это не так. В качестве примера рассмотрим простейшую нелинейную теорию с

$$F(R) = R + cR^2 - 2\Lambda, \quad (9.92)$$

включающую нелинейные слагаемые второго порядка: квадрат тензора Риччи $R_{AB}R^{AB}$ и скаляр Кречмана $\mathcal{K} = R_{ABCD}R^{ABCD}$, так что

$$S = \int d^D x \sqrt{|g|} [F(R) + c_1 R_{AB}R^{AB} + c_2 \mathcal{K}]. \quad (9.93)$$

Вычисления в рамках модели (9.93) с функцией (9.92) и метрикой (9.51), аналогичные проделанным выше, приводят к эффективной скалярно-тензорной теории с

$$V_{\text{Ein}}(\phi) = -\text{sign}(1 + 2c\phi)[d(d-1)]^{-d/2} \cdot |\phi|^{d/2} \frac{c'\phi^2 + \phi - 2\Lambda}{(1 + 2c\phi)^2},$$

$$c' = c + \frac{c_1}{d} + \frac{2c_2}{d(d-1)}, \quad (9.94)$$

$$K_{\text{Ein}}(\phi) = \frac{1}{\phi^2(1 + 2c\phi)^2} \cdot \left[c^2\phi^2(d^2 - 2d + 12) + d^2c\phi + \frac{1}{4}d(d+2) \right] + \frac{c_1 + c_2}{2\phi(1 + 2c\phi)}. \quad (9.95)$$

Пунктирные линии А и В на рис. 9.7 отмечают нули кинетического члена для квадратичной гравитации (9.92). Если флуктуация поля возникла слева от точки А, поле стремится к минимуму потенциала в точке $\phi = 0$, что соответствует увеличению размера дополнительного пространства. Некоторые флуктуации порождаются в области между точками А и В, где кинетический член отрицателен, поле ϕ , будучи фантомным, движется в сторону увеличения потенциала, к точке В. Очевидно, что точка В является аттрактором, не будучи минимумом потенциала. Вселенная, для которой реализовалась данная ситуация, имеет нетривиальное основное состояние, которое следует обсудить более детально. Для этого рассмотрим упрощенную ситуацию:

$$S = \int d^4x \left[\frac{1}{2}K(\phi)(\partial\phi)^2 - V(\phi) \right] \quad (9.96)$$

с $K(\phi_{\text{crit}}) = 0$ и без потери общности положим $\phi_{\text{crit}} = 0$. Тогда поведение $K(\phi)$ и $V(\phi)$ при малых ϕ есть $K(\phi) = k\phi + o(\phi)$ и $V(\phi) = V(0) + h\phi + o(\phi)$, где предполагается, что $k > 0$, $h > 0$, имея в виду точку В на рис. 9.7. Очевидно, что вблизи от критической точки поле стремится к ней вне зависимости

от того, справа или слева выбрано начальное значение. Тем не менее классическое движение около критической точки $\phi = 0$ отсутствует. Действительно, классическое уравнение для ϕ в искривленном пространстве-времени имеет вид

$$K(\phi) \square \phi + \frac{1}{2} K_\phi (\partial\phi)^2 = -V_\phi, \quad (9.97)$$

или, согласно приведенным выше выражениям для K и V для малых ϕ ,

$$k\phi \square \phi = - \left(h + \frac{1}{2} k (\partial\phi)^2 \right) + o(\phi). \quad (9.98)$$

Пусть $(\partial\phi)^2 > 0$ (как в случае космологии, где $\phi = \phi(t)$). Тогда правая часть уравнения (9.98) меньше, чем $-h < 0$ и не стремится к нулю при $\phi \rightarrow 0$. Значит, вторая производная в $\square \phi$ стремится к бесконечности, поскольку $\phi \rightarrow \phi_{\text{crit}} = 0$. Введение в лагранжиан дополнительных слагаемых с высшими степенями производных не улучшает ситуации. Действительно, в стационарном состоянии, когда все производные равны нулю, классическое уравнение (9.98) с $h > 0$ не имеет решений. Это означает, что кинетическая энергия однородного основного состояния не может равняться нулю.

9.3.9. Выводы

Показано, что чистая многомерная гравитация, без привлечения полей негеометрического происхождения и при весьма простом выборе геометрии, может быть ареной достаточно сложных явлений. К этому приводит предположение о нелинейности гравитации, однако такая нелинейность неизбежно возникает как результат квантовых поправок к теории Эйнштейна и в принципе должна включаться в теорию гравитации.

Введенное здесь приближение медленных изменений в многомерной гравитации существенно облегчает анализ. Это приближение оказывается справедливым в широком диапазоне явлений,

при которых кривизна и масштаб энергий далеки от планковских (правда, планковских в многомерном смысле, эти масштабы могут заметно отличаться от четырехмерных).

При редукции к четырем измерениям нелинейная многомерная гравитация приводит к сложной динамике эффективного скалярного поля, со стабильными и метастабильными минимумами потенциала и нетривиальным видом кинетического члена. Некоторые такие минимумы выглядят перспективно в смысле описания современной стадии эволюции Вселенной, с ускоренным расширением внешнего пространства и стабильными дополнительными измерениями. Последнее означает постоянство ФФК или их крайне медленное изменение, имеющее место, если система только приближается к своему стабильному состоянию. В других жизнеспособных моделях дополнительные измерения переменны, но их временные вариации достаточно малы и находятся в пределах современных экспериментальных ограничений.

Нерешенными проблемами остаются выбор физической конформной картины и необходимость тонкой настройки исходных параметров для получения наблюдаемого значения эффективной космологической постоянной.

9.4. Нелинейная D -мерная гравитация с несколькими дополнительными пространствами

Рассмотрим нелинейную многомерную гравитацию в пространствах более общего вида, допускающих редукцию к меньшей размерности, а именно — в D -мерном пространстве-времени \mathbb{M} со структурой

$$\mathbb{M} = \mathbb{M}_0 \times \mathbb{M}_1 \times \dots \times \mathbb{M}_n, \quad (9.99)$$

где $\dim \mathbb{M}_i = d_i$, и метрикой

$$ds_D^2 = g_{ab}(x)dx^a dx^b + \sum_{i=1}^n e^{2\beta_i(x)} g^{(i)}, \quad (9.100)$$

где (x) означает зависимость от первых d_0 координат x^a ; $g_{ab} = g_{ab}(x)$ — метрика в \mathbb{M}_0 , $g^{(i)}$ — не зависящие от x d_i -мерные метрики фактор-пространств \mathbb{M}_i , $i = \overline{1, n}$.

Исходное действие предполагаем в виде

$$S = \frac{1}{2} m_D^{D-2} \int \sqrt{D} g [F(R) + c_1 R_{AB} R^{AB} + c_2 \mathcal{K} + L_m], \quad (9.101)$$

где $F(R)$ — произвольная функция скалярной кривизны R пространства \mathbb{M} ; c_1, c_2 — константы; R_{AB} и $\mathcal{K} = R_{ABCD} R^{ABCD}$ — соответственно, тензор Риччи и скаляр Кречмана пространства \mathbb{M} ; L_m — лагранжиан материи. Заглавные латинские индексы охватывают все D координат, малые латинские (a, b, \dots) — координаты пространства \mathbb{M}_0 , а a_i, b_i, \dots — координаты пространств \mathbb{M}_i .

Будем полагать, что “калуца-клейновские” (КК) пространства \mathbb{M}_i компактны и имеют достаточно малый размер, чтобы исключить их наблюдаемость современными экспериментальными средствами.

D -мерный тензор Римана имеет следующие ненулевые компоненты:

$$\begin{aligned} R^{ab}_{cd} &= \overline{R}^{ab}_{cd}, \\ R^{aa_i}_{bb_i} &= \delta_{b_i}^{a_i} B_b^a{}_{(i)}, \quad B_b^a{}_{(i)} := e^{-\beta_i} \nabla_b (e^{\beta_i} \beta_i^a), \\ R^{a_i b_i}_{c_i d_i} &= e^{-2\beta_i} \overline{R}^{a_i b_i}_{c_i d_i} + \delta^{a_i b_i}_{c_i d_i} \beta_{i,a} \beta_i^a, \\ R^{a_i b_k}_{c_i d_k} &= \delta_{c_i}^{a_i} \delta_{d_k}^{b_k} \beta_{i,a} \beta_k^a, \quad i \neq k. \end{aligned} \quad (9.102)$$

Чертой сверху обозначены величины, полученные из метрик фактор-пространств g_{ab} и $g^{(i)}$, взятых по отдельности, $\beta_{,a} \equiv \partial_a \beta$, и обозначено $\delta^{ab}_{cd} \equiv \delta_c^a \delta_d^b - \delta_d^a \delta_c^b$ и аналогично для других видов индексов.

Ненулевые компоненты тензора Риччи и скалярная кривизна имеют вид

$$\begin{aligned}
R_a^b &= \bar{R}_a^b + \sum_i d_i B_{a^{(i)}}^b, \\
R_{a_i}^{b_i} &= e^{-2\beta_i} \bar{R}_{a_i}^{b_i} + \delta_{a_i}^{b_i} [\square \beta_i + \beta_{i,a} \sigma^{,a}], \quad \sigma := \sum_i d_i \beta_i, \\
R &= \bar{R}[g] + \sum_i e^{-2\beta} \bar{R}_i + 2\square \sigma + (\partial \sigma)^2 + \sum_i d_i (\partial \beta_i)^2, \quad (9.103)
\end{aligned}$$

где $(\partial \sigma)^2 \equiv \sigma_{,a} \sigma^{,a}$ и аналогично для других функций, $\square = g^{ab} \nabla_a \nabla_b - d_0$ -мерный оператор д'Аламбера, а $\bar{R}[g]$ и \bar{R}_i — скаляры Риччи, соответствующие g_{ab} и $g^{(i)}$. Здесь и далее \sum_i означает $\sum_{i=1}^n$.

Предположим, что $g^{(i)}$ — d_i -мерные метрики пространств ненулевой постоянной кривизны, т.е. сферы ($K_i = 1$) или компактные d_i -мерные гиперболические пространства ($K_i = -1$) с фиксированным радиусом кривизны r_0 , нормированным на D -мерный аналог m_D планковской массы, т.е. $r_0 = 1/m_D$ (мы пользуемся естественными единицами, полагая $c = \hbar = 1$). Тогда имеем:

$$\begin{aligned}
\bar{R}^{a_i b_i}{}_{c_i d_i} &= K_i m_D^2 \delta^{a_i b_i}{}_{c_i d_i}, \\
\bar{R}_{a_i}^{b_i} &= K_i m_D^2 (d_i - 1) \delta_{a_i}^{b_i}, \quad \bar{R}_i = K_i m_D^2 d_i (d_i - 1). \quad (9.104)
\end{aligned}$$

Тогда масштабные факторы $r_i(x) \equiv e^{\beta_i}$ в (9.100) безразмерны.

9.4.1. Приближение медленных изменений. Редукция к d_0 измерениям

Предположим, что все величины меняются медленно, т.е. рассматриваем каждую производную ∂_a (включая те, которые входят в определение \bar{R}) как выражения, содержащие малый параметр ε ; отбросим все величины порядков выше $O(\varepsilon^2)$.

Тогда имеем следующие разложения:

$$\begin{aligned}
R &= \phi + \bar{R}[g] + f_1 & f_1 &:= 2\Box\sigma + (\partial\sigma)^2 + \sum_i d_i(\partial\beta_i)^2; \\
F(R) &= F(\phi) + F'(\phi)(\bar{R}[g] + f_1) + O(\varepsilon^4); \\
R_{AB}R^{AB} &= \sum_i \frac{1}{d_i}\phi_i^2 + 2\sum_i d_i\phi_i[\Box\beta_i + (\partial\beta_i, \partial\sigma)] + O(\varepsilon^4); \\
\mathcal{K} &= 2\sum_i \frac{\phi_i^2}{d_1(d_1-1)} + 4\sum_i d_i\phi_i(\partial\beta_i)^2 + O(\varepsilon^4), \tag{9.105}
\end{aligned}$$

где

$$\phi_i := K_i m_D^2 (d_i - 1) e^{-2\beta_i}, \quad \phi := \sum_i d_i \phi_i. \tag{9.106}$$

Символ $(\partial\alpha, \partial\beta)$ означает $g^{ab}\alpha_{,a}\beta_{,b}$, и $F'(\phi) = dF/d\phi$.

В результате, пренебрегая $o(\varepsilon^2)$ и интегрируя по всем \mathbb{M}_i , получаем действие, редуцированное к d_0 измерениям:

$$\begin{aligned}
S &= \frac{1}{2}\mathcal{V} m_D^{d_0-2} \int \sqrt{g_0} d^{d_0}x \left\{ e^\sigma F'(\phi)\bar{R}_0 + [\text{Kin}] - \right. \\
&\quad \left. - 2V(\phi_i) + e^\sigma L_m \right\}, \\
[\text{Kin}] &= F'(\phi) e^\sigma f_1 + 2e^\sigma \sum_i d_i \phi_i \left[c_1 \Box\beta_i \right. \\
&\quad \left. + c_1(\partial\beta_i, \partial\sigma) + 2c_2(\partial\beta_i)^2 \right], \\
-2V(\phi_i) &= e^\sigma \left[F(\phi) + \sum_i d_i \phi_i^2 \left(c_1 + \frac{2c_2}{d_i-1} \right) \right], \tag{9.107}
\end{aligned}$$

где $g_0 = |\det(g_{\mu\nu})|$, а \mathcal{V} — произведение объемов n компактных d_i -мерных пространств \mathbb{M}_i единичной кривизны. Выражение (9.107) типично для (мульти)скалярно-тензорной теории (СТТ) гравитации в картине Йордана.

Вычитая полную дивергенцию, избавляемся от вторых производных в (9.107), и в результате кинетический член принимает вид

$$[\text{Kin}] = F' e^\sigma \left[-(\partial\sigma)^2 + \sum_i d_i (\partial\beta_i)^2 \right] - 2F'' e^\sigma (\partial\phi, \partial\sigma) + 4e^\sigma (c_1 + c_2) \sum_i d_i \phi_i (\partial\beta_i)^2, \quad (9.108)$$

где $F'' = d^2 F / d\phi^2$.

Переход к картине Эйнштейна

Для дальнейшего анализа удобно перейти к картине Эйнштейна с помощью конформного отображения (9.61). Выражение для скалярной кривизны в (9.107) преобразуется следующим образом:

$$\sqrt{g_0} e^\sigma \bar{R}_0 = \sqrt{g_0} f \bar{R}_0 = (\text{sign } f) \sqrt{\tilde{g}} \left[\tilde{R} + \frac{d_0 - 1}{d_0 - 2} \frac{(\tilde{\partial}f)^2}{f^2} \right] + \text{div}, \quad (9.109)$$

где тильдой обозначены величины, полученные из метрики $\tilde{g}_{\mu\nu}$ или с ее помощью, а div означает полную дивергенцию, не влияющую на уравнения поля. Действие (9.107) принимает вид

$$S = \frac{1}{2} \mathcal{V} m_D^{d_0-2} \int \sqrt{\tilde{g}} d^{d_0} x \left\{ [\text{sign } F'(\phi)] \left[\tilde{R} + [\text{Kin}]_E \right] - 2V_E(\phi_i) + L_{m(E)} \right\} \quad (9.110)$$

с кинетическим и потенциальным членами

$$[\text{Kin}]_E = \frac{1}{d_0 - 2} \left(\partial\sigma + \frac{F''}{F'} \partial\phi \right)^2 + \left(\frac{F''}{F'} \right)^2 (\partial\phi)^2 + \sum_i \left[1 + \frac{4}{F'} (c_1 + c_2) \phi_i \right] (\partial\beta_i)^2, \quad (9.111)$$

$$- 2V_E(\phi_i) = e^{-2\sigma/(d_0-2)} |F'|^{-d_0/(d_0-2)} \times$$

$$\times \left[F(\phi) + \sum_i d_i \phi_i^2 \left(c_1 + \frac{2c_2}{d_i - 1} \right) \right], \quad (9.112)$$

где мы опустили тильды, хотя используется метрика $\tilde{g}_{\mu\nu}$; лагранжиан материи есть

$$L_{m(E)} = e^{-2\sigma/(d_0-2)} |F'|^{-d_0/(d_0-2)} L_m. \quad (9.113)$$

Величины β_i и σ выражаются через n полей ϕ_i , число которых совпадает с числом дополнительных фактор-пространств.

9.4.2. Эффекты структуры дополнительных фактор-пространств

Выше обсуждались низкоэнергетические теории, соответствующие частным случаям метрики (9.100) при различном выборе начального действия. Неудивительно, что некоторые значения параметров подходят для описания отдельных свойств нашей Вселенной. Дополнительные возможности возникают при варьировании структуры дополнительных измерений, включающей число дополнительных фактор-пространств, их размерность и кривизну. При этом удается получить набор существенно различных низкоэнергетических теорий даже при фиксированном начальном лагранжиане.

Дополнительные измерения, масса инфлатона и Λ_{eff}

Вспомним еще раз известный недостаток хаотической инфляции в ее простейшей, квадратичной форме. Согласно наблюдаемым флуктуациям температуры космического микроволнового фона, масса инфлатона должна быть порядка $10^{-6} M_{\text{Pl}}$. Ее малость требует объяснения, до сих пор не существующего. Рассмотрим эффективный потенциал (9.112), генерируемый начальным действием (9.65), (9.92). Его вид представлен на рис. 9.8 для некоторых значений числа дополнительных измерений d . Очевидно, можно подобрать значения параметра d так, чтобы получить

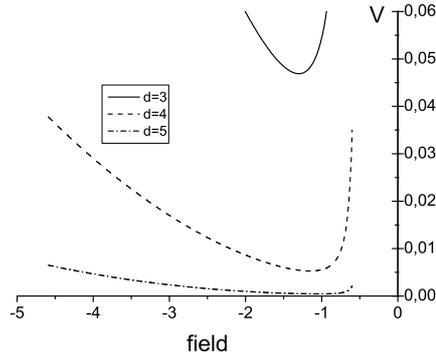


Рис. 9.8. Эффективный потенциал $V(\phi)$ для теории(9.65) с $F(R)$ вида (9.92) для различного числа дополнительных измерений. Значения параметров: $c = 1$, $\Lambda = -0.4$.

требуемую массу инфлатона. Численные оценки приводят к следующим результатам: вторая производная потенциала в минимуме (которая определяет массу инфлатона) равняется $\sim 0.2m_D$ при $d = 3$; $\sim 1.5 \cdot 10^{-3}m_D$ при $d = 5$ и $\sim 0.8 \cdot 10^{-5}m_D$ при $d = 7$. В исходном лагранжиане отсутствуют малые параметры. Тем не менее требуемое значение массы инфлатона возникает на классическом уровне при подходящем выборе параметра d .

Сильное влияние числа дополнительных измерений на вид низкоэнергетического лагранжиана неудивительно, поскольку потенциал (9.112) содержит фактор $\sim d^{-d}$. Так, если стационарное состояние есть $\phi = \phi_0$, безразмерный параметр $|\phi_0|m_D^{-2}$ и $F'(\phi_0)$ — величины порядка единицы, то эффективная космологическая константа $\Lambda_{\text{eff}} = V_{\text{in}}(\phi_0)$ связана с функцией $F(\phi_0)$ (которая может быть близка к m_D^2) следующим образом:

$$\Lambda_{\text{eff}}/F(\phi_0) \sim [d(d-1)]^{-d/2}. \quad (9.114)$$

Интересно, что $d^{-d} \approx 10^{-123}$ при $d = 67$. Это означает, что пространственно-временные флуктуации, порождающие $(67+4)$ -мерное пространство, могут служить причиной образования пространства с наблюдаемой плотностью энергии вакуума

$10^{-123} m_4$. Экстремальная малость Λ_{eff} связана лишь с числом дополнительных измерений пространства, случайно возникшего благодаря квантовым флуктуациям на планковском масштабе.

На рис. 9.9 представлен другой пример зависимости формы эффективного потенциала от параметра d . Видно, что само существование минимума потенциала, которое указывает на возможность стабилизации размеров дополнительного пространства, зависит от d . Так, если вселенная родилась вместе с дополнительным пространством отрицательной кривизны, то $\phi < 0$, см. рис. 9.9. Очевидно, что при начальном значении поля $\phi < -1$ его среднее значение в такой вселенной стремится к бесконечности при $d = 2$, к константе при $d = 4$ и к нулю для $d = 6$. Если же вселенная рождается при $-1 < \phi < 0$, то среднее значение поля локализуется в минимуме потенциала и размер дополнительного пространства остается фиксированным.

Изменение топологии дополнительного пространства также приводит к нетривиальным результатам. В качестве примера на рис. 9.10 представлен эффективный потенциал (9.94), возникающий при условии, что дополнительное пространство имеет топологию 3-сферы. Метастабильный минимум находится близко к нулю и вселенная, находящаяся в нем, может существовать длительное время, поскольку вероятность туннелирования в более низкий минимум мала.

Несколько фактор-пространств и переменный по пространству размер дополнительных измерений

Напомним, что мы априорно не предполагаем фиксированное число измерений дополнительного пространства, его геометрию и топологию. Вс это возникает на планковском масштабе благодаря квантовым флуктуациям. Чем сложнее возникающая структура дополнительного пространства, тем богаче возможности.

Рассмотрим пространство с метрикой (9.100) и двумя дополнительными фактор-пространствами: $\mathbb{M} = \mathbb{M}_4 \times \mathbb{M}_{d_1} \times \mathbb{M}_{d_2}$. Ис-

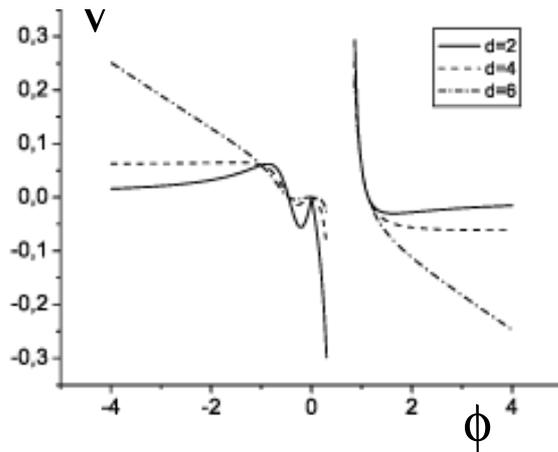


Рис. 9.9. Потенциалы при различных размерностях d дополнительного пространства, полученные из лагранжиана (9.65) с $F(R) = R + cR^2 + w_1R^3 + w_2R^4 - 2\Lambda$. Параметры: $c = 0$, $w_1 = 0$, $w_2 = -1$, $\Lambda = -0.25$. Кривые подогнаны к одному масштабу.

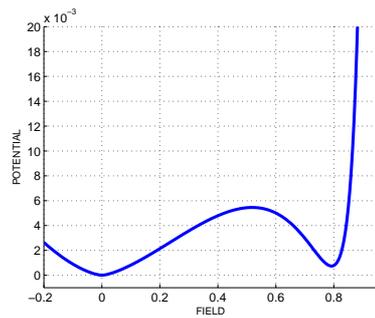


Рис. 9.10. Эффективный потенциал (9.94). Значения параметров: $c = -0.5$, $\Lambda = 0.2$, $c' = -0.626$ и $d = 2$.

пользуя то же самое действие (9.93), мы теперь должны ввести два скалярных поля, чтобы описать низкоэнергетический пре-

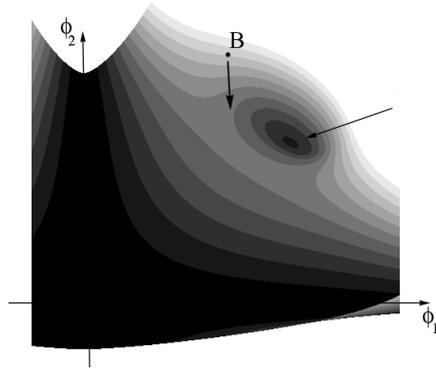


Рис. 9.11. Эффективный потенциал в случае дополнительного пространства $\mathbb{M}_{d_1} \times \mathbb{M}_{d_2}$, $d_1 = d_2 = 3$, с параметрами $c = -0.5$, $\Lambda = 0.2$, $c_1 = c_2 = -0.38$. Вид сверху, нижние уровни темнее. Локальный минимум отмечен длинной стрелкой.

дел теории. Эффективный потенциал в эйнштейновской картине имеет следующий вид:

$$\begin{aligned}
 V_{\text{Ein}}(\phi_1, \phi_2) = & -\frac{1}{2} \text{sign}(F'(d_1\phi_1 + d_2\phi_2)) \times \\
 & \times \frac{|\phi_1|^{d_1/2}}{[(d_1 - 1)]^{d_1/2}} \frac{|\phi_2|^{d_2/2}}{[(d_2 - 1)]^{d_2/2}} \times \\
 & \times \frac{F(d_1\phi_1 + d_2\phi_2) + d_1\phi_1^2 \left[c_1 + \frac{2c_2}{d_1-1} \right] + d_2\phi_2^2 \left[c_1 + \frac{2c_2}{d_2-1} \right]}{[F'(d_1\phi_1 + d_2\phi_2)]^2},
 \end{aligned} \tag{9.115}$$

На рис. 9.11 представлен потенциал с двумя долинами вдоль взаимно перпендикулярных направлений, $\phi_1 = 0$ и $\phi_2 = 0$, каждое из которых соответствует бесконечно большому размеру одного из фактор-пространств \mathbb{M}_{d_1} или \mathbb{M}_{d_2} . Наибольший интерес представляет локальный минимум, где оба фактор-пространства конечны. Вселенная может находиться достаточно долго в этом метастабильном состоянии, как и в случае дополнительного про-

странства с более простой структурой, обсуждавшемся выше.

Интересная возможность появляется, если предположить, что Вселенная образовалась в точке В на рис. 9.11. Во время инфляции поле движется от точки В вдоль стрелки. При этом квантовые флуктуации порождают различные значения поля в причинно не связанных доменах. Судьба домена зависит от значения поля внутри него. Даже если большинство доменов в конце инфляции окажутся в метастабильном минимуме, некоторая часть доменов устремится к одной из долин. В этом случае наша Вселенная должна содержать некоторое число доменов с макроскопически большим дополнительным пространством. Их число и размер сильно зависят от начальных условий.

Если в нашей Вселенной существуют области с дополнительным пространством большого размера, то законы низкоэнергетической физики будут отличаться от известных нам. Например, если звезда влетает в подобный домен, баланс сил внутри нее будет нарушен и звезда либо сколлапсирует, либо взорвется.

Итак, нами показано, что одна и та же теория может приводить к самым разным низкоэнергетическим теориям в зависимости от структуры дополнительного пространства и начальных условий.

9.5. Асимметричные браны в $d_0 = d + 2$ измерениях

В данном разделе мы применим описанную выше методику для построения некоторого класса моделей мира на бране.

Заметим, что к настоящему времени число публикаций в этой интенсивно развивающейся области исследований исчисляется тысячами, предложено огромное количество моделей (см., напр., обзоры [182, 219, 234, 250, 261]), различающихся по следующим основным признакам:

- Общая размерность D (в большинстве работ рассматрива-

ется $D = 5$).

- Модели с одной или несколькими бранами.
- Плоские или искривленные браны.
- Плоское или искривленное пространство вне браны.
- Компактное или некомпактное пространство вне браны.
- Возможное существование других полей, кроме гравитации (спинов 0, 1/2, 1), в пространстве вне браны.
- Тонкие или толстые браны. (Бесконечно тонкая брана может рассматриваться лишь как приближение, т.к. в любой фундаментальной теории существует предельная длина, такая, что на меньших масштабах классическое рассмотрение геометрии бессмысленно.)
- Различные симметрии как браны (или бран), так и окружающего пространства
- и т.д.

Уже эта достаточно грубая классификация показывает, насколько разнообразны должны быть наблюдательные предсказания: каждая из моделей, в принципе, дает свой набор эффектов.

Тонкие браны обычно рассматриваются как поверхностные распределения материи, происхождение которых не уточняется, например, в известных моделях Рэндалл и Сундрума [252, 253]. Более реалистические толстые браны строятся как распределения некоторых материальных, чаще всего скалярных полей негеометрического происхождения (в частности, как доменные стенки или другие топологические дефекты) [60, 124, 195, 225, 240, 260]). Здесь [99] мы пытаемся получить скалярное поле с потенциалом, подходящим для построения реалистической браны в

5-мерном пространстве-времени, из чисто геометрических величин — компонент метрики пространства еще более высокой размерности, причем измерения выше пятого будем предполагать компактными и достаточно малыми в духе концепции Калуцы-Клейна. Таким образом, возникающие модели можно назвать гибридными, так как они содержат и малые компактные, и сравнительно большие или бесконечные дополнительные измерения, аналогично упомянутой выше модели Хожавы-Виттена.

Итак, в D -мерном пространстве $D-5$ дополнительных измерений предполагаются малыми и компактными в духе Калуцы-Клейна, а пятое измерение — большим или бесконечным; брана получается из скалярного поля метрического происхождения, с эффективным потенциалом и кинетическим членом, зависящими от параметров исходного многомерного действия¹⁵.

Как мы убедимся, особенность наших моделей мира на бране заключается в том, что скалярное поле ϕ многомерного происхождения может обладать кинетическим членом переменного знака, т.е. его кинетическая энергия положительна в некоторых областях значений ϕ и отрицательна в других. Брана может образоваться в окрестности точки перехода между такими областями, и тогда она существенно асимметрична по построению. Впрочем, симметричные или почти симметричные браны также получаются. Кроме того, в отличие от многих авторов, мы рассматриваем модели мира на бране как с убывающим, так и с растущим вне браны масштабным фактором (warp factor); в последнем случае обеспечивается чисто гравитационный механизм удержания материи, что показано на примере массивных скалярных полей.

9.5.1. Геометрия

Для поиска решений в виде толстых бран предположим:

¹⁵Имеются примеры построения толстых бран с помощью скаляра геометрического происхождения в рамках интегрируемой геометрии Вейля [47, 52]. Мы, однако, работаем полностью в рамках римановой геометрии.

(i) Рассматривается D -мерное пространство-время со структурой $\mathbb{M}_0 \times \mathbb{M}_1$ и метрикой (9.100), с единственным дополнительным (КК) фактор-пространством \mathbb{M}_1 постоянной кривизны. Следовательно, имеется лишь одно эффективное скалярное поле $\phi = d_1\phi_1$. Тогда формулы (9.111) и (9.112) приводят к следующим выражениям для $[\text{Kin}]_E$ и V_E :

$$\begin{aligned} [\text{Kin}]_E &= K(\phi)(\partial\phi)^2, \\ K(\phi) &= \frac{(\text{sign } F')}{4\phi^2} \left[\frac{1}{d_0 - 2} \left(d_1 - 2\phi \frac{F''}{F'} \right)^2 + \right. \\ &\quad \left. + 4\phi^2 \left(\frac{F''}{F'} \right)^2 + d_1 + \frac{c_k}{F'} \right], \end{aligned} \quad (9.116)$$

$$-2V_E = \left(\frac{|\phi|}{d_1(d_1 - 1)} \right)^{d_1/(d_0-2)} |F'|^{-d_0/(d_0-2)} \left[F(\phi) + c_v \frac{\phi^2}{d_1} \right], \quad (9.117)$$

$$L_{m(E)} = \left(\frac{|\phi|}{d_1(d_1 - 1)} \right)^{d_1/(d_0-2)} |F'|^{-d_0/(d_0-2)} L_m, \quad (9.118)$$

$$c_k = 4(c_1 + c_2), \quad c_v = c_1 + \frac{2c_2}{d_1 - 1}. \quad (9.119)$$

На рис. 9.12 показан пример функций $V_E(\phi)$ и $K_E(\phi)$, получаемых из исходной функции $F(R)$ в (9.101) вида

$$F(R) = R + w_2 R^2 + w_3 R^3 + w_4 R^4 - 2\Lambda_D \quad (9.120)$$

со следующими значениями параметров:

$$\begin{aligned} d_0 = 5, \quad d_1 = 6, \quad w_2 = 1.53, \quad w_3 = 0.79, \quad w_4 = 0.15, \\ \Lambda = 6, \quad c_k = 8, \quad c_v = 4.3. \end{aligned} \quad (9.121)$$

Заглядывая вперед, скажем, что этот набор параметров приводит к желательному для построения асимметричной браны виду решения для скалярного поля (кинку). Поле ϕ изменяется

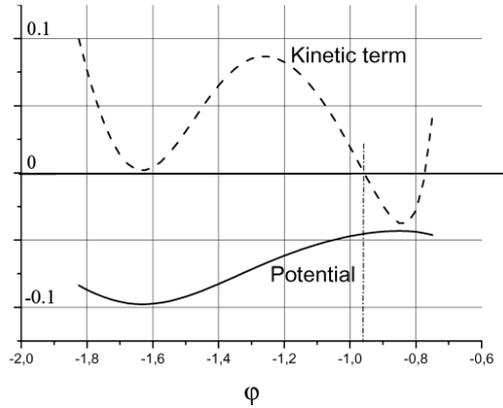


Рис. 9.12. Функции $V_E(\phi)$ и $K_E(\phi)$ в пространстве $M_5 \otimes M_d, d = 6$, с параметрами исходного лагранжиана, заданными в (9.120), (9.121). Вертикальной чертой отмечена середина браны.

между максимумом V_E в “фантомной” области, где $K_E < 0$, и минимумом V_E в “нормальной” области ($K_E > 0$).

Можно получить и другой, более обычный вид кинков, в которых поле ϕ меняется между двумя минимумами потенциала $V_E(\phi)$, а функция $K(\phi)$ всюду положительна. Пример таких функций $V(\phi)$ и $K(\phi)$ соответствует формуле (9.120) и параметрам

$$d_0 = 5, d_1 = 6, w_2 = 1.53, w_3 = 0.79, w_4 = 0.15, \\ \Lambda = 6, c_k = 1, c_v = -8. \quad (9.122)$$

Можно убедиться, что в результате возникнет асимметричный вариант толстой браны типа RS2 в пятимерном пространстве-времени, с двумя антидеситтеровскими горизонтами по разные стороны от браны.

(ii) Будем рассматривать пространство M_0 как $(d_0 = d + 2)$ -мерное пространство-время со структурой $\mathbb{R}_t \times \mathbb{R}_u \times \mathbb{R}^d$ и мет-

рикой в картине Эйнштейна

$$ds^2 = e^{2\gamma(u)} dt^2 - e^{2\beta(u)} d\vec{x}^2 - e^{2\alpha(u)} du^2. \quad (9.123)$$

Здесь \mathbb{R}_t — временная ось, \mathbb{R}^d — d -мерное евклидово пространство, α , β и γ — функции координаты u , интерпретируемой как координата, меняющаяся в направлении перпендикулярно бране, с пока не фиксированной областью определения $\mathbb{R}_u \subseteq \mathbb{R}$. Любое подпространство $u = \text{const}$ в \mathbb{M}_0 есть пространство Минковского.

Под (толстой) браной будет пониматься концентрация плотности энергии скалярного поля в некоторой области значений координаты u .

Тензор Римана в \mathbb{M}_0 (мы здесь опускаем тильду над R без риска путаницы) диагонален по отношению к парам индексов, и его ненулевые компоненты имеют вид

$$\begin{aligned} R^{tu}_{tu} &= -e^{-\gamma-\alpha} (e^{\gamma-\alpha} \gamma')', \\ R^{it}_{kt} &= -\delta_k^i e^{-2\alpha} \beta' \gamma', \\ R^{ui}_{uk} &= -\delta_k^i e^{-\beta-\alpha} (e^{\beta-\alpha} \beta')', \\ R^{ij}_{kl} &= -e^{-2\alpha} \beta'^2 \delta^{ij}_{kl}, \end{aligned} \quad (9.124)$$

где штрих обозначает d/du ; индексы i, j, \dots соответствуют d координатам в \mathbb{R}^d . Мы имеем в виду в основном размерность $d = 3$, но оставляем d произвольным для общности.

Необходимое условие регулярности многообразия \mathbb{M}_0 — конечность его скаляра Кречмана $\mathcal{K}_0 = R^{ab}_{cd} R^{cd}_{ab}$. В силу попарной диагональности тензора R^{ab}_{cd} , \mathcal{K} есть сумма квадратов ненулевых компонент последнего. Таким образом, в регулярных конфигурациях все компоненты (9.124) конечны.

Для тензора Риччи имеем

$$\begin{aligned} R^t_t &= -e^{-2\alpha} [\gamma'' + \gamma'(-\alpha' + \gamma' + d\beta')], \\ R^u_u &= -e^{-2\alpha} [\gamma'' + \gamma'^2 - \alpha'\gamma' + d(\beta'' + \beta'^2 - \alpha'\beta')], \\ R^k_i &= -\delta_i^k e^{-2\alpha} [\beta'' + \beta'(-\alpha' + \gamma' + d\beta')]. \end{aligned} \quad (9.125)$$

9.5.2. Уравнения поля

Рассмотрим действие (9.110) в картине Эйнштейна с $L_m = 0$ для метрики (9.123) и скалярного поля $\phi(u)$. Лагранжиан скалярного поля имеет вид

$$L = K(\phi)g^{ab}\partial_a\phi\partial_b\phi - 2V(\phi). \quad (9.126)$$

В уравнениях Эйнштейна $G_a^b = -T_a^b$ ненулевые компоненты тензора энергии-импульса T_a^b есть

$$\begin{aligned} T_t^t &= T_i^i = \frac{1}{2}e^{-2\alpha}K\phi'^2 + V, \\ T_u^u &= -\frac{1}{2}e^{-2\alpha}K\phi'^2 + V \end{aligned} \quad (9.127)$$

(по i нет суммирования).

Уравнения скалярного поля и некоторые компоненты уравнений Эйнштейна можно записать в виде

$$2Ke^{-\alpha-d\beta-\gamma}(e^{\gamma+d\beta-\alpha}\phi')' + K_\phi e^{-2\alpha}\phi'^2 = 2V_\phi, \quad (9.128)$$

$$\gamma'' + \gamma'(-\alpha' + \gamma' + d\beta') = -(2/d)V e^{2\alpha}, \quad (9.129)$$

$$\beta'' + \beta'(\beta' - \alpha' - \gamma') = -(K/d)\phi'^2, \quad (9.130)$$

$$\beta'' - \gamma'' + (\beta' - \gamma')(-\alpha' + d\beta' + \gamma') = 0, \quad (9.131)$$

$$d(d-1)\beta'^2 + 2d\beta'\gamma' = K\phi'^2 - 2Ve^{2\alpha}. \quad (9.132)$$

Здесь (9.129) — уравнение $R_t^t = \dots$, (9.130) — разность $R_t^t - R_u^u = \dots$, а (9.131) — разность $R_t^t - R_i^i = \dots$ при фиксированном i .

Только три из этих пяти уравнений независимы: скалярное уравнение (9.128) следует из уравнений Эйнштейна, а (9.132) является первым интегралом остальных.

Eq. (9.131) легко интегрируется:

$$2\gamma' - 2\beta' = C_1 e^{\alpha-\gamma-d\beta}, \quad C_1 = \text{const}. \quad (9.133)$$

У нас еще остается произвол в выборе координатного условия, способного упростить уравнения. Мы воспользуемся двумя вариантами выбора координаты u для различных целей.

1. При выборе "квазиглобальной" координаты $u = \rho$, соответствующей условию $\alpha + \gamma = 0$, обозначая

$$e^{2\gamma} = e^{-2\alpha} =: A(\rho), \quad e^\beta =: r(\rho), \quad (9.134)$$

получим полевые уравнения в виде

$$2K(Ar^d\phi')' + K_\phi Ar^d\phi'^2 = 2r^dV_\phi, \quad (9.135)$$

$$(A'r^d)' = -(4/d)r^dV, \quad (9.136)$$

$$dr''/r = -K\phi'^2, \quad (9.137)$$

$$A(r^2)'' - r^2A'' + (d-2)r'(2Ar' - A'r) = 0, \quad (9.138)$$

$$d(d-1)Ar'^2 + dA'rr' = KAr^2\phi'^2 - 2r^2V. \quad (9.139)$$

Формула (9.133) переписывается как

$$(A/r^2)' = C_1/r^{d+2}. \quad (9.140)$$

Условие конечности всех компонент (9.124) в терминах $A(\rho)$ и $r(\rho)$ имеет вид

$$A \frac{r'^2}{r^2} < \infty, \quad |A''| < \infty, \quad \frac{|A'r'|}{r} < \infty, \quad \frac{|Ar''|}{r} < \infty. \quad (9.141)$$

2. Другое полезное координатное условие — $\alpha = \gamma$, при котором координату (обозначим ее x) можно назвать волновой, так как она удобна для анализа волновых уравнений на фоне исследуемых метрик. Она будет использована в разделе 9.5.6.

9.5.3. Асимптотическое поведение решений

Наша дальнейшая цель — получение регулярных решений уравнений поля, пригодных для описания толстых бран. Для этого рассмотрим возможное асимптотическое поведение нетривиальных решений, соответствующее стремлению $\phi \rightarrow \phi_0 = \text{const}$ при больших значениях квазиглобальной координаты ρ .

Тогда в силу (9.137)

$$r \approx r_0 + r_1 \rho, \quad r_0, r_1 = \text{const.} \quad (9.142)$$

Таким образом, возможны два режима: $r \rightarrow r_0$ и (без потери общности) $r \approx \rho$ при $\rho \rightarrow \infty$.

В случае $r \rightarrow r_0 > 0$ уравнение (9.140) дает $A' \rightarrow \text{const}$ и, следовательно, $V \rightarrow 0$. Наконец, уравнение (9.139) тогда приводит к $A' \rightarrow 0$, и мы приходим к плоской асимптотике.

В случае $r \approx \rho \rightarrow \infty$ уравнение (9.140) при дальнейшем интегрировании дает

$$A \approx A_1 \rho^2 - \frac{c_1}{d+1} \rho^{1-d}, \quad A_1 = \text{const}, \quad (9.143)$$

и подстановка в (9.136) связывает A_1 с асимптотическим значением потенциала:

$$V_0 = V(\phi_0) = -\frac{1}{2}d(d+1)A_1. \quad (9.144)$$

Полагая $V_0 \neq 0$, следовательно, $A_1 \neq 0$, мы можем пренебречь вторым членом в (9.143) и, таким образом, приходим к двум возможностям:

- 1) $V_0 > 0$, $A < 0$: пространство де Ситтера dS_{d+2} в Т-области, т.е. в космологическом виде, и
- 2) $V_0 < 0$, $A > 0$: пространство анти-де Ситтера AdS_{d+2} в статическом виде.

Итак, единственная нетривиальная статическая асимптотика при $\phi \rightarrow \text{const}$ и $V \neq 0$ есть пространство анти-де Ситтера. Возможна также плоская асимптотика ($A_1 = 0$, $V \rightarrow 0$), которую мы в дальнейшем не рассматриваем в силу ее бесперспективности с точки зрения удержания частиц на бране.

Отметим, что проведенный анализ асимптотик не зависел от вида функции $K(\phi)$.

9.5.4. Симметричные и несимметричные браны

Обсудим теперь глобальные свойства решений.

Если (временно) допустить, что рассматриваемая конфигурация обладает зеркальной (\mathbb{Z}_2) симметрией относительно подпространства $\rho = 0$, то метрические функции $A(\rho)$ и $r(\rho)$ четные, а функция $\phi(\rho)$ носит характер кинка. Без потери общности,

$$\phi(0) = 0, \quad A(0) = r(0) = 1, \quad A'(0) = r'(0) = 0. \quad (9.145)$$

Подставляя (9.145) в (9.140), немедленно убеждаемся, что $c_1 = 0$, так что $A(\rho) \equiv r^2(\rho)$. Таким образом, метрика характеризуется единственным масштабным фактором (называемым иногда фактором искажения, warp factor) $A(\rho)$, и задача сводится к рассмотренной ранее (см. [38, 95] и приведенные там ссылки). Так, было получено, что при условии $K(\phi) > 0$ единственный вид глобально регулярных конфигураций имеет две АдС-асимптотики по обе стороны от браны; более того, по обе стороны имеются горизонты типа антидеситтеровских, на которых $A \equiv r^2 \rightarrow 0$, при конечном ρ [38], и лишь за такими горизонтами располагаются асимптотики с большими ρ , где $A \sim \rho^2$. Это легко увидеть из уравнения (9.137), которое приводит к $r'' \leq 0$, и в любом решении, удовлетворяющем условию (9.145) и такому, что $r \neq \text{const}$, $r(\rho)$ обращается в ноль при некотором конечном ρ .

Таким образом, все браны с $K(\phi) > 0$ — типа RS2, или, другими словами, являются сглаженными вариантами тонких бран с положительным натяжением и убывающими (в направлении от браны) масштабными факторами. Все такие конфигурации содержат горизонты $\rho = \rho_h$, где $A \equiv r^2 = A_1(\rho - \rho_h)^2$. В то же время такие горизонты неустойчивы к появлению материи, так как на них обращается в ноль не только g_{tt} , но и пространственный объем, что приводит к бесконечному росту плотности материи. Этот вывод подтвержден в работах [38, 95] исследованием поведения массивного пробного скалярного поля, которое не только само расходится на АдС-горизонтах, но и, в силу этой расходимости, оказывается бесконечной энергия поля на единицу

объема браны. Таким образом, по меньшей мере этот вид полей не удерживается симметричными бранами типа RS2.

Отметим, что \mathbb{Z}_2 -симметрия возможна только если как $K(\phi)$, так и $V(\phi)$ являются четными функциями. Это, разумеется, не так в нашей общей модели. Однако нетрудно показать, что отсутствие \mathbb{Z}_2 -симметрии не изменяет основных свойств бран при $K > 0$: имеется два горизонта АдС-типа по обе стороны от середины браны, хотя и расположенные несимметрично, но приводящие к тем же проблемам. По этой причине, по нашему мнению, больший интерес представляют браны с растущим масштабным фактором.

Отсутствие зеркальной симметрии и поведение вблизи нуля $K(\phi)$

Мы видели, что искомые решения не существуют в случае нормального скалярного поля ($K > 0$). Для наших целей поэтому требуется либо всюду фантомное скалярное поле $K(\phi) < 0$, либо поле с переменным знаком $K(\phi)$. Остановимся на этой возможности. В этом случае система, вообще говоря, лишена \mathbb{Z}_2 -симметрии.

Рассмотрим ситуацию, когда $K(\phi)$ имеет разные знаки при (без потери общности) $\phi > 0$ и $\phi < 0$. В общем случае функции A и r^2 различны и связаны соотношением (9.140).

Пусть $K(\phi)$ имеет простой ноль при $\phi = 0$, т.е.

$$K(\phi) = K_1\phi + \frac{1}{2}K_2\phi^2 + \dots, \quad K_1, K_2 = \text{const}, \quad (9.146)$$

где $K_1 \neq 0$, и проверим, может ли вся система вести себя регулярно при соответствующем значении ρ , скажем, $\rho = 0$. Предположим, что метрические функции A и r могут быть разложены в ряды Тейлора около $\rho = 0$:

$$\begin{aligned} A &= A_0 + A_1\rho + \frac{1}{2}A_2\rho^2 + \dots, \\ r &= r_0 + r_1\rho + \frac{1}{2}r_2\rho^2 + \dots \end{aligned}$$

и что $V(0) = V_0$. Тогда уравнения (9.136) и (9.140) приводят в своих главных порядках к соотношениям между коэффициентами разложения:

$$\begin{aligned} A_2 + \frac{dr_1}{r_0} A_1 &= -\frac{4}{d} V_0, \\ A_0 \left(\frac{A_1}{A_0} - 2 \frac{r_1}{r_0} \right) &= \frac{c_1}{r_0^d}. \end{aligned} \quad (9.147)$$

Уравнение (9.137) принимает вид

$$\frac{d}{r_0} (r_2 + r_3 \rho + \dots) = -(K_1 \phi + \dots) \phi'^2. \quad (9.148)$$

Если $r_2 \neq 0$, то в главном порядке имеем $\phi \phi'^2 = dr_2 / (r_0 K_1)$, т.е. ϕ имеет фиксированный знак и не может пройти через нулевое значение при $\rho = 0$. Если мы хотим построить модель, где ϕ переходит от нормального к фантомному режиму, этот случай надо отбросить и положить $r_2 = 0$. Тогда (9.148) приводит к уравнению

$$d \frac{r_3}{r_0} \rho \approx K_1 \phi \phi'^2, \quad (9.149)$$

которое решается в случае $\phi \propto \rho$ около $\rho = 0$ при условии $r_3 \neq 0$.

Итак, наша система допускает гладкий переход от нормального к фантомному режиму поля ϕ . При таком переходе ϕ имеет простой ноль при $\rho = 0$ (при наших допущениях). Величина r'' тоже имеет там простой ноль, что вполне естественно, поскольку, согласно (9.137), r'' должно менять знак вместе с $K(\phi)$.

9.5.5. Конфигурации с возрастающими масштабными факторами

Покажем, что существуют такие функции $V(\phi)$ и $K(\phi)$, которые приводят к конфигурациям с возрастающими масштабными факторами.

Если функция $r(\rho)$ известна (или выбрана), то решение системы уравнений (9.135)–(9.139) известно полностью: $A(\rho)$ находится из (9.140), $\phi(\rho)$ из (9.137), а $V(\phi(\rho))$ из (9.136), и это приводит к определенной функции $V(\phi)$ при условии, что ϕ' не меняет знака. Воспользуемся этим методом обратной задачи и покажем, как задать $r(\rho)$, чтобы получить желаемые асимптотики.

На обоих пределах $\rho \rightarrow \pm\infty$ следует выбрать $r \sim a_{\pm}|\rho|$ с некоторыми $a_{\pm} = \text{const} > 0$. В области, где (по предположению) $K > 0$, следует взять $r(\rho)$ такое, что $r'' < 0$ и наоборот, соблюдая при этом условие $r > 0$.

Повторное интегрирование уравнения (9.140) дает

$$B(\rho) := \frac{A}{r^2} = c_1 \int \frac{d\rho}{r^{d+2}}, \quad (9.150)$$

где интеграл сходится при $\rho \rightarrow \pm\infty$. Следовательно, можно выбрать константу интегрирования в (9.150) таким образом, что $B > 0$ всюду и стремится к положительным конечным пределам при $\rho \rightarrow \pm\infty$. Таким образом, получено искомое поведение решений при больших ρ с (вообще говоря) различными отрицательными значениями V на двух асимптотиках. В частном случае $c_1 = 0$ просто полагаем $B \equiv 1$.

Далее, если известно $K(\phi)$, получаем $\phi(\rho)$ из (9.137) с точностью до аддитивной постоянной, которую следует выбрать так, что $K(\phi)$ меняет знак одновременно с r'' ; это всегда возможно, если K имеет один ноль. Можно заметить, что если $K(\phi)$ имеет простой ноль и, с другой стороны, $r''' \neq 0$ в точке перегиба, в которой $r'' = 0$, то все функции ведут себя гладко в этой точке (см. раздел 4) и, кроме того, ϕ' имеет всюду один знак, что приводит к правильно заданной функции $V(\phi)$. Этим завершается наше построение.

Для численного решения уравнений поля вместо найденных функций $V(\phi)$ и $K(\phi)$, показанных на рис.9.13, будем для простоты использовать функции

$$V_E = [(\phi - 0.5)^2 - 1]^2 - 1.1, \quad K_E = -5\phi, \quad (9.151)$$

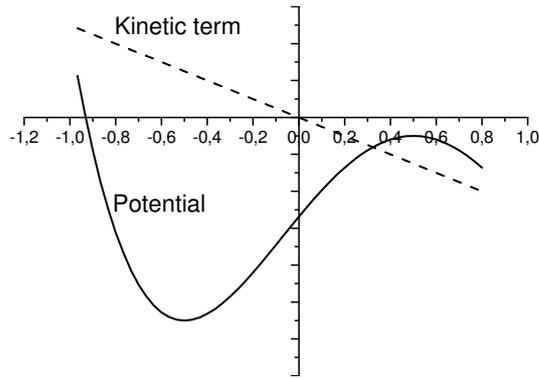


Рис. 9.13. Функции $V(\phi)$ и $K(\phi)$ в модели, полученной из многомерной гравитации в приближении медленного движения.

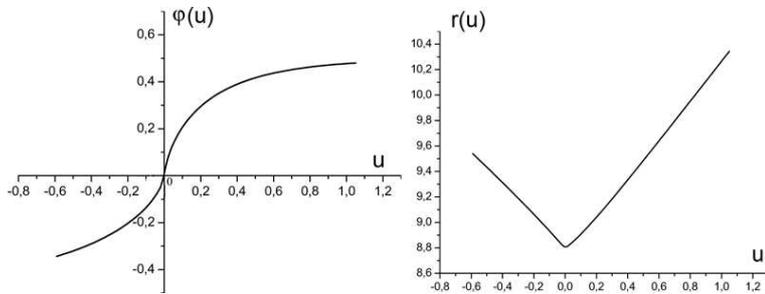


Рис. 9.14. Решения уравнений (9.135)–(9.139) с $V(\phi)$ и $K(\phi)$, заданным согласно (9.151)

вид которых показан на рис.9.13. Пример соответствующего им регулярного решения уравнений поля с растущим масштабным фактором $A \equiv r^2 \sim \rho^2$ представлен на рис.9.14.¹⁶

¹⁶Мы здесь отвлекаемся от геометрического смысла ϕ , связанного с масштабом КК-измерений, так что переход через $\phi = 0$ — не причина для беспокойства. Напомним, что на рис. 9.12 все представляющие интерес значения ϕ отрицательны.

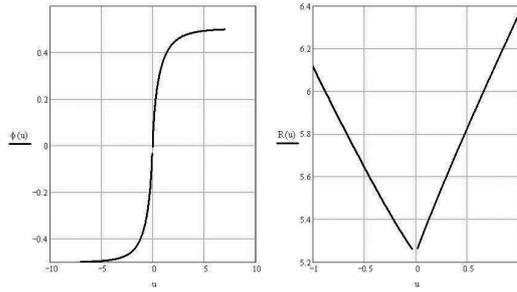


Рис. 9.15. Асимметричная брана с растущим масштабным фактором

9.5.6. Удержание материи

Предположим, что d_0 -мерная метрика обладает растущим при удалении от браны масштабным фактором $A \equiv r^2$, и рассмотрим проблему удержания материи на примере пробного скалярного поля.

Лагранжиан материи L_m , заданный изначально в D -мерном пространстве, в нашей d_0 -мерной картине Эйнштейна принимает вид (9.113) или (9.118). Если L_m описывает линейное скалярное поле,

$$L_m = g^{AB} \partial_A \chi \partial_B \chi - m^2 \chi^2, \quad (9.152)$$

то в d_0 -мерной картине Эйнштейна, в предположении $\chi = \chi(x)$,

$$S_m \sim \int d^{d_0} x \sqrt{\tilde{g}} \left[|F'(\phi)|^{-1} \tilde{g}^{ab} \partial_a \chi \partial_b \chi - e^{-2\sigma/(d_0-2)} |F'(\phi)|^{-d_0/(d_0-2)} m^2 \chi^2 \right]. \quad (9.153)$$

Возьмем $[d_0 = (d+2)]$ -мерную метрику в общем виде (9.123), не фиксируя “калибровку”, т.е., координату u , но полагая, с учетом приведенной выше аргументации, $\gamma \equiv \beta$, что эквивалентно $r^2 \equiv A$. Действие пробного поля переписывается в виде

$$S_m = \int d^{d+2} x Q(u) \left[\tilde{g}^{ab} \partial_a \chi \partial_b \chi - M^2(u) \chi^2 \right]. \quad (9.154)$$

Отметим, что Q включает $\sqrt{\tilde{g}}$. Здесь \tilde{g}_{ab} — метрика в картине Эйнштейна в рассматриваемых моделях, а функции $Q(u)$ и $M^2(u)$ задаются формулами

$$\begin{aligned} Q(u) &= e^{\alpha+d\beta+\gamma}|F'(\phi)|^{-1}, \\ M^2(u) &= m^2 e^{-2\sigma/d}|F'(\phi)|^{-2/d}, \end{aligned} \quad (9.155)$$

где в нашем случае $\sigma = d_1\beta_1$. Уравнение поля для χ есть

$$\partial_a[Q\tilde{g}^{ab}\partial_a\chi] + QM^2\chi = 0. \quad (9.156)$$

Рассмотрим отдельную моду χ — плоскую волну в пространстве Минковского (вспомним, что сечения постоянного u пространства M_0 плоские)

$$\chi = X(u) e^{i\omega t - i\vec{p}\vec{x}}. \quad (9.157)$$

Тогда уравнение (9.156) дает для $X(u)$ (штрих = d/du):

$$[Q e^{-2\alpha} X']' + [\omega^2 e^{-2\gamma} - \vec{p}^2 e^{-2\beta} - M^2(u)] Q X = 0, \quad (9.158)$$

где $p^2 = \omega^2 - \vec{p}^2$ — это фактически квадрат массы поля χ в пространстве Минковского.

Дальнейшая подстановка

$$X(u) = y(u)/\sqrt{f}, \quad f(u) := Q e^{-2\alpha} \quad (9.159)$$

приводит это уравнение к виду, близкому к шредингеровскому:

$$y'' + [p^2 e^{2\alpha-2\beta} - V_{\text{eff}}(u)]y = 0, \quad (9.160)$$

с эффективным потенциалом

$$V_{\text{eff}}(u) = M^2 e^{2\alpha} + \frac{1}{2} \frac{f''}{f} - \frac{1}{4} \frac{f'^2}{f^2}. \quad (9.161)$$

Вид уравнения (9.160) показывает, что правильное уравнение Шредингера, в котором p^2 — искомое собственное значение краевой задачи, возникает, если координата u выбрана при помощи

условия $\alpha = \beta$. (Напомним, что в обсуждаемых решениях, в терминах общей метрики (9.123), выполнено еще и условие $\beta = \gamma$.) Выбранную таким образом координату можно назвать “волновой”, так как она удобна для анализа волновых уравнений на фоне рассматриваемой метрики; она аналогична известной “черепашной” координате, используемой при анализе возмущений черных дыр.

Проблема удержания частиц на бране сводится к вопросу о существовании мод с конечной энергией E_χ на единицу объема браны, где

$$\begin{aligned} E_\chi[X] &= \int du Q(u) T_t^t[X], \\ 2T_t^t[X] &= e^{-2\alpha} X'^2 + e^{-2\beta} (\omega^2 + \vec{p}^2 + M^2) X^2. \end{aligned} \quad (9.162)$$

Требование сходимости интеграла (9.162) приводит к граничным условиям для решений уравнения (9.158) или (9.160), описывающих захваченные моды.

В представляющих интерес конфигурациях, показанных на рис. 9.15, имеем $e^\beta \equiv r \sim |\rho|$ при $\rho \rightarrow \pm\infty$, а ϕ стремится к постоянным значениям, при которых величина F' конечна. (Крайние значения $\rho \rightarrow \pm\infty$ соответствуют экстремумам потенциала V_E , заданного формулой (9.117), а величина F' конечна при всех значениях ρ , так как иначе потенциал V_E обращался бы в бесконечность.) Таким образом, из D -мерной получается фактически d_0 -мерная система.

Волновая координата $u = x$ связана с ранее использованной координатой ρ соотношением

$$dx^2 = d\rho^2 / r^4(\rho), \quad (9.163)$$

следовательно, x принимает конечные значения при $\rho \rightarrow \pm\infty$, и $\rho \in \mathbb{R}$ соответствует интервалу $x \in (x_-, x_+)$, $x_\pm \text{const}$.

Рассмотрим для определенности $x \rightarrow x_+$, где $\rho \rightarrow +\infty$. Тогда $Q \sim r^{d+2} \sim \xi^{-(d+2)}$, а $M \rightarrow m_1$, где m_1 — постоянная, “перенор-

мированная” масса m и $\xi = x_+ - x$. Тогда необходимое условие сходимости E_χ имеет вид

$$|X|^2, |dX/dx|^2 = o(\xi^{d-1}). \quad (9.164)$$

Кроме того, легко убедиться, что эффективный потенциал в (9.160) ведет себя как $V_{\text{eff}} \sim \xi^{-2} \rightarrow \infty$ и содержит положительные вклады как от функции $f(u)$, так и от массового слагаемого.

При $x \rightarrow x_-$ имеют место те же условия сходимости E_χ и то же самое поведение V_{eff} . Можно сделать вывод, что полученная краевая задача вполне аналогична квантово-механической задаче для уравнения (9.160) с потенциалом типа ящика, а следовательно, захваченные моды образуют дискретный спектр p^2 при любых значениях исходной массы m скалярного поля χ .

Представляет интерес и возможное существование безмассовой моды ($p = 0$). В случае $m = 0$ (так что $m_1 = 0$) линейно независимые решения уравнения (9.158) есть

$$X_1 = \text{const}, \quad X_2 = \int \frac{r^2}{Q} dx. \quad (9.165)$$

С решением X_1 интеграл E_χ расходится на обоих пределах. С решением X_2 мы можем потребовать его сходимости на одном из пределов, например, при x_- , тогда интегрирование в (9.165) происходит от x_- до x и приводит к некоторому конечному значению $X_2(x_+) > 0$. Но при этом интеграл (9.162) расходится при $x = x_+$.

Таким образом, на рассматриваемой бране скалярные поля (независимо от их исходной массы) не образуют безмассовых мод. Этот вывод подтверждается видом V_{eff} , полученным для приведенных выше вариантов V_E , K_E , см. рис.9.16: поскольку $V_{\text{eff}} > 0$, наименьшие собственные значения p^2 заведомо положительны.

9.5.7. Гравитация на бране

Приведем оценку модифицированного ньютоновского потенциала для гравитационного поля в полученных моделях мира на

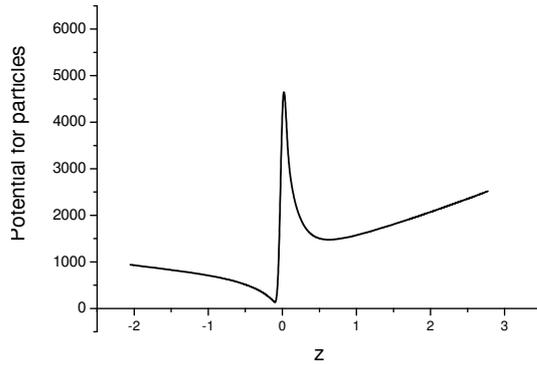


Рис. 9.16. Эффективный потенциал V_{eff} скалярного поля с исходной массой $m = 0$ для конфигурации (9.15). Калибровка $\alpha = \gamma$.

бране (полагая, в соответствии с реальностью, $d = 3$, $d_0 = 5$). Таких моделей найдено два вида:

- (А) с убывающими масштабными факторами, типа RS2, и
- (Б) с растущими масштабными факторами — нового типа, с эффективным скалярным полем, меняющим свой характер от фантомного к нормальному.

Приведенные выше конкретные примеры являются лишь представителями указанных классов, и, очевидно, в общем случае параметры моделей произвольны; необходимо лишь, чтобы все их характерные масштабы были далеки от планковских.

В случае А модели отличаются от RS2 в первую очередь конечной толщиной, которую можно охарактеризовать шириной l_m максимума плотности энергии скалярного поля ϕ , образующего брану. В случае $l_m \ll l_c$, где l_c — радиус кривизны АдС-пространства вокруг браны, при расчете гравитации на бране можно применять приближение тонкой браны. В обсуждаемых моделях условие $l_m \ll l_c$ выполняется, так как l_m имеет, по построению, порядок лишь на несколько единиц меньше планковского (имеется в виду многомерный планковский масштаб, см.

(9.104), $m_D \gtrsim 1$ ТэВ, соответствующая длина — $l_D \lesssim 10^{-17}$ см), а радиус кривизны l_c может, по оценкам, достигать десятых долей миллиметра. Во всяком случае, интерес для эксперимента в обозримом будущем могут представлять лишь конфигурации с достаточно большими l_c — как минимум, больше атомных размеров, $\sim 10^{-8}$ см. При этих условиях можно воспользоваться имеющимися оценками для модели RS2 (см. обзор [92] и приведенные там ссылки), модифицируя их применительно к нашему случаю. Если считать малой имеющуюся в общем случае \mathbb{Z}_2 -асимметрию (разные значения l_c по разные стороны от браны), то, аналогично [126], получаем следующее приближенное выражение для обобщенного ньютоновского потенциала V_N как функции расстояния R от тяготеющей массы M :

$$\begin{aligned} V_N(R) &\approx -\frac{GM}{R} \left[1 + \frac{4}{3\pi} U(z) \right], \\ z^2 U(z) &= (z \cos z - \sin z) \operatorname{ci} z + (\cos z + z \sin z) \operatorname{si} z + \pi/2, \end{aligned} \quad (9.166)$$

где $z = R/l_c$, а функции $\operatorname{si} z$ и $\operatorname{ci} z$ — соответственно, интегральные синус и косинус. Выражение (9.166) представляет собой плавный переход между асимптотическими выражениями для V_N при больших и малых радиусах R [107]:

$$\begin{aligned} V_N(R) &\approx -\frac{GM}{R} \left(1 + \frac{2l_c}{3R} \right), & R \ll l_c, \\ V_N(R) &\approx -\frac{GM}{R} \left(1 + \frac{2l_c}{3R^2} \right), & R \gg l_c. \end{aligned} \quad (9.167)$$

Несложно убедиться, что на радиусах больше $10l_c$ потенциал практически неотличим от ньютоновского, а на малых по сравнению с l_c радиусах (но больших по сравнению с толщиной браны, иначе не работает наше приближение) гравитация эффективно становится пятимерной.

Ньютоновская гравитационная постоянная G в этих формулах связана с пятимерной планковской массой m_5 соотношением

$$8\pi G = (m_5^3 l_c)^{-1}. \quad (9.168)$$

В свою очередь, величина m_5 связана с исходной многомерной планковской массой m_D коэффициентом порядка единицы: $m_5^3 = m_D^{D-2} / \mathcal{V}$ [см. (9.101) и (9.110)].

Во случае Б браны представляют собой сглаженный вариант тонких бран с отрицательным натяжением. Такие браны рассматривались прежде всего в модели RS1 [252], причем в этой модели именно на них, по предположению, сконцентрирована видимая материя. Предположим, как и в случае А, что толщина браны мала по сравнению с радиусами кривизны АдС-пространств по обе стороны от браны, и, подобно случаю А, будем считать малым отклонение от \mathbb{Z}_2 -симметрии. Тогда можно модифицировать для нашего случая результаты, полученные ранее для гравитационного поля на наблюдаемой (планковской) бране в моделях типа RS1 [49, 280]. В качестве исходной воспользуемся формулой [280]

$$V_N(R) = -\frac{G_1 M}{R} \left[-1 + \frac{8}{3\pi} \frac{\text{si } z}{1 - e^{-2z_0}} + \frac{2}{3\pi z^2} \frac{z \cos z + \sin z}{1 - e^{-2z_0}} \right], \quad (9.169)$$

где по-прежнему $z = R/l_c$, $8\pi G_1 = 1/(m_5^3 l_c)$, $z_0 = l_b/l_c$, l_b — расстояние между двумя бранами с разными знаками натяжения вдоль дополнительного пространственного измерения. Наша конфигурация с единственной браной, имеющей отрицательное натяжение, соответствует пределу $l_b \rightarrow \infty$, т.е. $z_0 \rightarrow \infty$, и формула (9.169) несколько упрощается. При больших и малых значениях отношения R/l_c получаем асимптотические формулы, отличные от (9.167):

$$\begin{aligned} V_N(R) &\approx -\frac{G_1 l_c M}{3\pi R^2}, & R \ll l_c, \\ V_N(R) &\approx -\frac{G_1 M}{3R}, & R \gg l_c. \end{aligned} \quad (9.170)$$

Таким образом, при $R \gg l_c$ выполняется закон Ньютона с гравитационной постоянной $G = G_1/3$ (тогда как в случае А было $G = G_1$), а на малых радиусах, как и в случае А, потенциал пропорционален r^{-2} , хотя и с другим коэффициентом.

Радиус кривизны l_c АдС-пространства вокруг бран остается произвольным параметром построенных моделей.

Заключение

Исходя из многомерной теории гравитации с произвольным числом измерений и различными видами нелинейных по кривизне лагранжианов, построен новый класс моделей мира на бране смешанного характера: часть ($D - d_0 = d_1$) дополнительных измерений компактна и ненаблюдаема за счет своих малых размеров, т.е. носит калуца-клейновский (КК) характер, а из оставшихся $d_0 = 5$ измерений одно имеет бесконечную протяженность и ненаблюдаемо за счет удержания материи в окрестности выделенной 4-поверхности (браны).

Для построения моделей использовано приближение медленных изменений по сравнению с планковскими масштабами, показавшее свою плодотворность при решении ряда космологических проблем — см. раздел 9.3.

Существенной особенностью построенных моделей мира на бране является отсутствие каких-либо негравитационных источников. Скалярное поле ϕ , необходимое для описания толстой браны как топологического дефекта (доменной стенки), возникает как функция масштабного фактора КК-измерений.

Предполагая, что исходный лагранжиан содержит полином четвертой степени от многомерной скалярной кривизны, а также квадрат тензора Риччи и скаляр Кречмана, мы приходим к двум типам моделей толстых бран: одни (А) с убывающим, другие (Б) с растущим масштабным фактором в направлении от браны. В обоих случаях пространство-время вдали от браны асимптотически антидеситтеровское, эффективная толщина браны (которую можно определить, например, как ширину пи-

ка плотности энергии поля ϕ на половине высоты пика) значительно меньше, чем радиус кривизны АдС-пространства. Браны типа А подобны бранам в модели RS2, браны типа Б — бранам с отрицательным натяжением в модели RS1. В последнем случае эффективное скалярное поле носит фантомный характер по одну сторону от браны и обычный по другую. В общем случае браны типа А не обладают \mathbb{Z}_2 -симметрией относительно центральной плоскости, а браны типа Б всегда \mathbb{Z}_2 -асимметричны по построению. Однако для обоих типов бран не исключено (хотя бы приближенное) равенство АдС-радиусов кривизны l_c по разные стороны от браны.

Полученные приближенные формулы описывают модифицированный закон тяготения Ньютона, которому подчиняется материя, сконцентрированная на бране. Эти формулы подобны тем, которые действуют в моделях RS1 и RS2, поскольку в них отражаются общие закономерности поправок к закону Ньютона за счет существования дополнительного (пятого) измерения, а существование КК-измерений на этих поправках не отражается.

Глава 10

Тонкая настройка параметров Вселенной

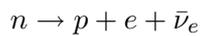
Одним из впечатляющих наблюдательных фактов является тонкая настройка параметров Вселенной. Под этим подразумевается, что все микроскопические параметры — массы частиц и константы взаимодействия — не только имеют такие значения, при которых во Вселенной может возникнуть разумная жизнь, но такие допустимые значения заключены в очень узких пределах. Незначительные отклонения от реальных значений привели бы к катастрофическим последствиям для существования сложных структур. Рассмотрим несколько примеров.

10.1. Зачем нужно слабое взаимодействие?

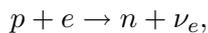
10.1.1. Масса электрона и свойства Вселенной

Характерные массы фермионов порядка 1 ГэВ. Лишь электрон имеет массу на три порядка меньше. Представим себе, что его масса увеличилась в несколько раз. Казалось бы, не должно произойти ничего существенного — все равно он останется очень легкой частицей. Однако увеличение массы электрона всего лишь

второе привело бы к катастрофическим последствиям для Вселенной. Действительно, разность масс нейтрона и протона $\Delta m = m_n - m_p = 1.28$ МэВ, в то время как масса “нового электрона” составляет $m_E = 3m_e = 1.53$ МэВ. Это означает, что нейтрон становится стабильным, поскольку распад за счет слабого взаимодействия



энергетически запрещен. Кроме того, оказывается разрешенной реакция



из-за которой нарушается стабильность атомов водорода. Следовательно, сразу после периода рекомбинации практически все протоны и электроны конвертируются в нейтральные частицы. Заряженные частицы исчезнут задолго до образования первых звезд. В этом случае звезд не появится вовсе, поскольку гравитационное сжатие облаков частиц эффективно только при наличии процесса диссипации энергии.

Итак, увеличение массы электронов второе приведет к отсутствию звезд во Вселенной и, следовательно, отсутствию разумной жизни.

10.1.2. Углеродный уровень

Тонкая настройка работает настолько безоговорочно, что с ее помощью можно оценивать неизвестные параметры. Наиболее известный пример — предсказание Фреда Хойла.

В звездах углерод образуется в две ступени. Сначала сливаются две альфа-частицы, образуя нестабильный изотоп бериллий-8. Затем к бериллию добавляется еще одна, третья альфа-частица и появляется ядро углерода. Но бериллий-8 быстро распадается. Поэтому при образовании Вселенной надо бы предусмотреть резонансное взаимодействие бериллия-8 и альфа-частицы. Природа так и сделала — энергетический уровень, равный 7,65 МэВ,

замечателен тем, что суммарная энергия возбужденного состояния ядра углерода всего на 0.3 МэВ выше суммарной массы альфа-частицы и ядра бериллия. Эти 0.3 МэВ компенсируются кинетической энергией сталкивающихся частиц, резонансно увеличивая эффективность реакции, что было теоретически предсказано Фредом Хойлом в 1953 году. Когда наша Вселенная только зарождалась, Природа уже должна была “знать” о будущей необходимости этого уровня.

10.1.3. Медленные реакции в звездах

Рассмотрим еще один интересный пример тонкой настройки параметров Вселенной. Зададимся вопросом: какой должна быть константа Ферми G_F , чтобы возникла разумная жизнь нашего типа. Опыт показывает, что для ее возникновения требуется порядка 5 млрд лет. Именно столько — не меньше — должны длиться ядерные реакции в звезде типа Солнца. Таким образом, время первой реакции в протон-протонном цикле



должно быть $t = 1/(\sigma v n) \sim 5 - 10$ млрд лет, где $v \sim \sqrt{2T/m_p}$ — средняя скорость протонов в звезде, а средняя концентрация протонов в звезде $n \sim 1.5 \text{ г/см}^3$. Ядерные реакции начинаются при температурах $T \approx 10^7 \text{ К}$. Сечение первой реакции протон-протонного цикла, идущей за счет слабого взаимодействия, $\sigma \sim G_F^2 E^2 \sim G_F^2 T^2$.

Собирая все вместе, получаем выражение для фермиевской константы

$$G_F \sim \frac{m^{1/4}}{(nt)^{1/2} T^{5/4}}, \quad (10.2)$$

связывающее ее с такими параметрами как средние концентрация протонов и температура в звезде, масса протона и время возникновения разумной жизни, а точнее, время жизни звезды. Подставляя численные значения, получаем $G_F \sim 10^{-5} \text{ ГэВ}^{-2}$,

что по порядку величины совпадает с экспериментальным значением.

Отметим, что ядерные реакции обычно протекают быстро. Малая величина фермиевской константы позволяет резко замедлить ядерные процессы в звездах, начинающиеся с реакции (10.1). Просмотрим, насколько эффективно это замедление. Светимость Солнца известна, поэтому энергия, выделяемая Солнцем в секунду, есть $L_{\odot} = 3 \cdot 10^{26}$ Дж/с. Следовательно, 1 кг Солнечной материи излучает

$$\varepsilon_{\odot} \sim L_{\odot}/M_{\odot} \sim 10^{-7} \text{ Дж}/(\text{с кг}).$$

Сравним с удельным энерговыделением человека весом 80 кг:

$$\varepsilon_{\text{homo}} \sim \frac{2000 \cdot 10^3 \cdot 0.01 \cdot 4.18}{24 \cdot 3600 \cdot 80} \sim 0.02 \text{ Дж}/(\text{с} \cdot \text{кг}).$$

Здесь для оценки предполагается, что лишь 1% энергии, поглощаемой человеком, 2000 ККал, переходит в тепловую. Благодаря малости фермиевской константы ядерные реакции в звездах оказываются на много порядков менее эффективными по сравнению с химическими реакциями.

Вопрос. Если химические реакции внутри человека проходят гораздо быстрее, чем ядерные реакции в звездах, то почему температура человеческого тела так низка по сравнению с температурой Солнца?

Слабое взаимодействие должно быть действительно очень слабым, чтобы нейтрон долго жил. Кроме того, при взрывах сверхновых нейтрино, благодаря слабому взаимодействию, способны унести основную энергию от центральной части звезды, позволяя ее внутренним слоям резко сжиматься, коллапсировать. Но с другой стороны, константа слабого взаимодействия не может быть слишком малой. В противном случае, во-первых, число нейтрино, рождающихся в звезде, было бы слишком малым, и, во-вторых, внешние слои взрывающейся звезды не получили бы от нейтрино достаточной энергии, чтобы разлететься в космическом пространстве.

10.1.4. Примеры тонкой настройки Вселенной

Ниже перечислены параметры, которые должны находиться в узком интервале, чтобы во Вселенной появились сложные структуры.

- 1) Константа сильного взаимодействия.
Больше: отсутствует водород и легкие ядра, например углерод.
Меньше: ядра тяжелее водорода неустойчивы.
- 2) Константа слабого взаимодействия.
Больше: а) Слишком много первичного гелия в первых звездах \Rightarrow слишком много тяжелых элементов вырабатывается в звездах, б) Отсутствуют взрывы сверхновых. Ядра остаются внутри остатков звезд.
Меньше: а) Слишком мало первичного гелия в первых звездах \Rightarrow слишком мало тяжелых элементов вырабатывается в звездах, б) Отсутствуют взрывы сверхновых. Ядра остаются внутри остатков звезд.
- 3) Гравитационная постоянная.
Больше: звезды слишком горячие и вследствие этого сгорают слишком быстро (для появления разума надо 5 млрд лет).
Меньше: холодные звезды. Следовательно, отсутствуют ядерные реакции и их продукт — тяжелые элементы.
- 4) Константа электромагнитного взаимодействия.
Больше: атомы меньше (все структуры более плотные). Ядра нестабильны.
Меньше: слишком слабые химические связи.
- 5) Отношение константы электромагнитного взаимодействия к гравитационной константе.
Больше: массы звезд велики и поэтому их время жизни мало.

Меньше: нет массивных звезд и поэтому нет тяжелых элементов.

- 6) m_e/m_p . **Больше:** слабые химические связи.
Меньше: слабые химические связи.
- 7) Скорость расширения Вселенной.
Больше: галактики не успевают сформироваться
Меньше: Вселенная коллапсирует до образования звезд.
- 8) Средняя плотность энергии барионов.
Больше: слишком много первичного гелия, звезды сгорают слишком быстро.
Меньше: слишком мало первичного гелия, значит, слишком мало тяжелых элементов.
- 9) Начальные флуктуации плотности.
Больше: Вселенная состояла бы из черных дыр.
Меньше: подавлен процесс звездообразования.
- 10) Среднее расстояние между звездами.
Больше: плотность тяжелых элементов недостаточна для образования твердых планет.
Меньше: нестабильные орбиты планет.
- 11) Скорость распада бериллия-8.
Больше: отсутствуют элементы тяжелее бериллия.
Меньше: скорость ядерных реакций возрастает, что приводит к малому времени жизни звезд.
- 12) Начальный избыток нуклонов по сравнению с антинуклонами.
Больше: плотность барионов слишком высока. Образуются в основном черные дыры.
Меньше: недостаточно барионов для образования звезд.
- 13) Частота взрывов сверхновых.
Больше: уничтожается жизнь на планетах.

Меньше: недостаточно тяжелых элементов для образования твердых планет.

14) Плотность темной материи.

Больше: ранний коллапс Вселенной.

Меньше: Образование галактик подавлено.

15) Разность масс нейтрона и протона.

Больше: малое время жизни нейтрона, значит, слишком мало первичного гелия.

Меньше: Быстрый распад нейтронов. Отсутствуют тяжелые элементы.

10.1.5. Какой должна быть Окончательная Теория?

Предположим, мы пришли к цели и создали окончательную “теорию всего”. Что она собой должна представлять? Похоже, есть три варианта: А. Законы созданы неким высшим разумом для каких-то своих целей. Б. Законы природы строго выводятся из начального постулата и просто не могут быть другими. Вся история науки видится как приближение к этому постулату. В. Существует много вселенных с различными законами, и мы живем в одной из них. Этот подход приобретает все больше сторонников, но конструктивен ли он?

Число измерений

Если число измерений больше трех, то не удастся удержать планеты около звезд — нет стационарных орбит. Если пространство имеет всего два измерения, то трудно обеспечить жизнеспособность сложной структуры. Итак, только 3 пространственных измерения.

Подбор частиц и взаимодействий

Основными являются протоны и электроны, которые уже позволяют создать простейший из атомов — водород. Понадобятся более массивные ядра, например углерод, основа органических соединений. Но массивные ядра из одних протонов мгновенно распадаются из-за сильного электрического отталкивания протонов. Вводим новое, сильное взаимодействие. Слишком большие ядра — тоже плохо — слишком высока плотность вещества, сделанного из них. Поэтому сильное взаимодействие должно быть короткодействующим. Ситуация существенно улучшилась, но ядра, хоть и медленно, но все-таки распадаются. Если за счет флуктуаций протон удалится от остальных, сильное взаимодействие уже не в состоянии противодействовать электрическому отталкиванию, протон улетает, ядро распадается. Видно, что цель близка — надо совсем немного, чтобы ядра были и достаточно тяжелыми и стабильными. Последний штрих — вводим нейтроны, заряда не имеющие, но участвующие в сильном взаимодействии. Теперь ядра состоят из протонов и нейтронов и стабильны вплоть до ядер железа. Все это должна была предусмотреть Природа заранее.

Для создания сложных структур нужен большой объем, и расширение пространства решает эту проблему. Но скорость расширения не может быть произвольной — от нее зависит множество эффектов. Главным можно считать образование звезд.

Расширение пространства означает такое охлаждение среды, при котором разум существовать не может. Противоречие можно устранить, создав локальный нагрев среды. Эту функцию выполняют звезды. Кроме того, звезды являются фабриками тяжелых элементов.

Время жизни звезд

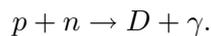
Тяжелые элементы образуются в звездах и должны быть выведены в конечном итоге в космос, чтобы успели родиться звез-

ды следующего поколения с планетарными системами. Поэтому параметры теории должны быть подобраны так, чтобы в конце жизни хотя бы часть первых звезд имела короткий жизненный цикл, оканчивающийся взрывом. Но последующие поколения звезд должны жить долго (10 млрд лет), чтобы зародился разум. Природе удалось удовлетворить этим противоречивым требованиям.

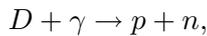
Долгая жизнь звезд следующих поколений основана на медленном протон-протонном цикле (вот зачем нужно слабое взаимодействие). Короткая жизнь первых звезд — большая масса + горение первичного гелия. Но как создать первичный гелий?

Проход через угольное ушко

Для создания гелия необходимы нейтроны. Последние, рожденные, когда Вселенная была горячей, должны распасться в дальнейшем. Как сохранить нейтроны? Сохраняем нейтроны внутри дейтерия?



Мешает обратная реакция



эффективная при высоких температурах. Энергия связи дейтерия невелика — 2.234 МэВ, и при температурах выше 10^9 К гелий быстро разрушается. Пороговая температура Вселенной 10^9 К достигается через 300 с после Начала. Если бы нейтроны имели “обычное” время жизни порядка 10^{-6} с, то к этому времени их концентрация была бы ничтожно мала.

Выход из положения — долгоживущий нейтрон (900 с)! За это время температура успевает уменьшиться, так что энергия фотонов становится недостаточной для разрушения ядер дейтерия. Оставшихся 24% нейтронов оказывается достаточно.

10.2. Механизмы тонкой настройки

10.2.1. Каскадная редукция

Динамика нашей Вселенной хорошо описывается современной теорией, содержащей 30-40 параметров. Число этих параметров, значения которых определяются экспериментально, слишком велико, чтобы считать теорию окончательной. Более того, формулировка будущей теории, претендующей на эту роль, не должна содержать конкретных численных значений. В противном случае потребуются еще более общая теория, объясняющая их происхождение. Кроме того, известно, что для рождения и существования нашей Вселенной область допустимых значений параметров должна быть крайне узкой (тонкая настройка параметров), что крайне труднообъяснимо. Обсуждению данной проблемы посвящена обширная литература [42, 59, 71, 125, 170, 181, 216, 254, 290, 303]. Одним из путей ее решения является предположение о множественности вселенных с различными свойствами [178, 207, 264, 266]. Богатые возможности для обоснования этого предположения содержатся в идее многомерности нашего пространства. Число дополнительных измерений давно является предметом дискуссий. Модель Калуцы — Клейна первоначально содержала одно дополнительное измерение. Сейчас уже обсуждаются, например, пространства бесконечной размерности [115] и даже пространства переменной размерности [69].

В статье [31] предложен механизм образования вселенных с существенно различающимися свойствами. Понятие суперпространства расширяется до набора суперпространств с различным, не ограниченным сверху, числом измерений. В каждом из суперпространств происходят процессы редукции в подпространства меньшей размерности. Точнее, определим суперпространство $\mathcal{M}_D = (M_D; g_{ij})$ как набор метрик g_{ij} в пространстве M_D с точностью до диффеоморфизмов. На пространственно-подобном сечении Σ введем метрику h_{ij} (см. детали в [227] и обзоре [309])

и определим пространство всех римановых $(D - 1)$ -метрик

$$\text{Riem}(\Sigma) = \{h_{ij}(x) | x \in \Sigma.\}$$

Амплитуда перехода от Σ_{in} к Σ_{f} есть

$$A_{\text{f, in}} = \langle h_{\text{f}}, \Sigma_{\text{f}} | h_{\text{in}}, \Sigma_{\text{in}} \rangle = \int_{h_{\text{in}}}^{h_{\text{f}}} Dg \exp[iS(g)]. \quad (10.3)$$

Напомним, что мы используем единицы $\hbar = c = 1$. Топологии сечений Σ_{in} и Σ_{f} могут быть различны. Нас будут интересовать квантовые переходы, при которых топология гиперповерхности Σ_{f} представляет собой прямое произведение подпространств $M_{D-1-d} \otimes M_d$. Пространство M_d предполагается компактным. Далее будет исследовано, какой класс геометрий на гиперповерхности Σ_{f} может инициировать классическую динамику в дальнейшем.

Всё рассмотрение проводится в рамках нелинейной гравитации в пространстве $D > 4$ измерений без включения полей материи. Мы обсудим возникновение планковского масштаба и его связь с массой инфлатона. Редукция в пространства меньшей размерности осуществляется в несколько этапов, образуя каскад. Различные каскады приводят к формированию четырехмерных пространств с различными эффективными теориями и различным числом дополнительных измерений.

Параметры низкоэнергетической теории оказываются зависящими от топологии дополнительных пространств и варьируются в широких пределах, см. также [122], несмотря на то что параметры исходной теории фиксированы. Это касается и таких фундаментальных величин, как, например, масса Планка и топология дополнительного пространства.

Отсутствие полей материи, постулируемых в статье изначально, является принципиальным моментом. Предполагается, что при низких энергиях компоненты метрического тензора дополнительного (супер-)пространства будут интерпретироваться как поля материи в духе теорий типа Калуцы — Клейна.

Образование пространства-времени совместно с параметрами теории

Первоначально понятие суперпространства означало набор различных геометрий [306], затем в него был включен набор всевозможных топологий [227]. Сделаем следующий шаг и расширим суперпространство за счет добавления пространств различной размерности. Более точно, определим расширенное суперпространство E как прямое произведение суперпространств M различных размерностей:

$$E = M_1 \otimes M_2 \otimes M_3 \otimes \dots \otimes M_D \dots \quad (10.4)$$

Здесь M_d — суперпространство размерности $d = 1, 2, \dots$, являющееся набором всевозможных геометрий (с точностью до диффеоморфизмов) и топологий.

Квантовые флуктуации порождают различные геометрии в каждом из суперпространств (пространственно-временная пена) [227, 306]. Вероятность квантового рождения “долгоживущих” 3-геометрий и условия, при которых это происходит, обсуждаются в разделе 10.2.1. Сейчас же рассмотрим следствия гипотезы о существовании расширенного суперпространства.

Выберем пространство некоторой размерности M_D . Его топологическая структура может меняться под влиянием квантовых флуктуаций, [200, 227]. В частности, возможны и топологии, допускающие расслоения пространства пространственно-подобными поверхностями Σ , что и подразумевается в выражении для амплитуды перехода (10.3). В дальнейшем рассматриваются пространства, допускающие разбиение вида

$$M_D = \mathbb{R} \otimes M_{D-1}^{(space)}, \quad (10.5)$$

где \mathbb{R} представляет временноподобное направление.

Конкретизируем топологию и метрику на пространственно-подобном сечении Σ_f амплитуды (10.3), подчинив их следующим условиям:

а) топология сечения Σ_f имеет вид прямого произведения

$$\Sigma_f = M_{D-1}^{(\text{space})} = M_{D_1} \otimes M_{d_1}, \quad (10.6)$$

где D_1 и d_1 — размерности соответствующих подпространств. Здесь и далее компактное подпространство обозначается как M_{d_k} , $k = 1, 2, \dots$

б) выполняется условие на кривизну подпространств M_{D_1} и M_{d_1}

$$R_{D_1}(g_{ab}) \ll R_{d_1}(\gamma_{ij}), \quad (10.7)$$

т.е. кривизна подпространства M_{D_1} мала по сравнению с кривизной подпространства M_{d_1} .

с) В наборе подпространств M_{d_1} выберем максимально симметричные пространства постоянной кривизны R_{d_1} , которая связана с параметром кривизны k обычным образом,

$$R_{d_1}(\gamma_{ij}) = kd_1(d_1 - 1). \quad (10.8)$$

В остальном топология и геометрия подпространства M_{D_1} произвольны.

В разделе 10.2.1 обсуждается вероятность квантовых переходов, конечное состояние которых характеризуется определенной выше топологией и метрикой. Сейчас же рассмотрим следствия таких переходов. Выберем динамические переменные и форму лагранжиана. Метрику пространства M_D запишем в виде [110]

$$\begin{aligned} ds^2 &= G_{AB}dX^A dX^B = g_{ab}(x)dx^a dx^b - b^2(x)\gamma_{ij}(y)dy^i dy^j = \\ &= dt^2 - h_{\alpha\beta}dx^\alpha dx^\beta - b^2(x)\gamma_{ij}(y)dy^i dy^j. \end{aligned} \quad (10.9)$$

Здесь g_{ab} — метрика подпространства $R \otimes M_{D_1}$ с сигнатурой $(+ - - - \dots -)$, $b(x)$ — радиус кривизны компактного подпространства M_{d_1} , а $\gamma_{ij}(y)$ — его положительно определенная метрика. Для заданного расслоения пространства пространственно-подобными поверхностями, всегда можно выбрать нормальные гауссовы координаты, что и использовано в последнем равенстве (10.9).

Действие Эйнштейна-Гильберта для гравитационного поля, линейное по кривизне R , полностью описывает физические явления при низких энергиях, в которых важна гравитация. Тем не менее, очевидно, что квантовые эффекты неизбежно приводят к нелинейным поправкам в действии [135]. При этом действие содержит слагаемые с высшими производными в виде полиномов различных степеней от скаляра Риччи и других инвариантов. Таким образом, какое бы гравитационное действие мы не приняли за основу, после учета квантовых поправок оно приобретает вид

$$S = \int d^N x (R + \varepsilon_1 R^2 + \varepsilon_2 R^3 + \varepsilon_3 R^4 + \dots + \alpha_1 R_{AB} R^{AB} + \dots)$$

с набором неизвестных коэффициентов, зависящих от геометрии пространства [106, 259, 285]. Тем не менее проблема не так остра, поскольку нелинейные по скаляру Риччи теории могут быть сведены конформным преобразованием к линейной теории [58, 184]. Более того, в статьях [94, 98] предлагается более общий способ приведения произвольных лагранжианов к стандартной форме Эйнштейна-Гильберта в пределе низких энергий. При этом оказывается, что проблема стабилизации размеров дополнительных измерений [110] становится вполне решаемой — см. главу 8.

Итак, теории с высшими производными неизбежны при учете квантовых явлений, способны давать те же предсказания, что и общая теория относительности, и при этом обладают более богатыми возможностями. Для простоты ограничимся действием, квадратичным по скаляру Риччи,

$$S_D = \frac{1}{2} \int d^D X \sqrt{-G} [R_D(G_{AB}) + C R_D(G_{AB})^2 - 2\Lambda] + \int_{\partial M_D} K d^{D-1} \Sigma. \quad (10.10)$$

Вклад от границы ∂M_D представляет собой слагаемое, введенное Хокингом и Гиббонсом. В дальнейшем анализе параметры

C и Λ полагаются фиксированными, а все многообразие низкоэнергетических теорий, порождаемых действием (10.10), возникает благодаря различным способам редукции к пространствам меньшей размерности.

Поскольку рассуждения проводятся в рамках чисто геометрического подхода, в теории (10.10) изначально масштаб не фиксирован и произвольные параметры C и Λ безразмерны. В следующем разделе обсуждается возникновение планковского масштаба при низких энергиях.

Как было показано в [94, 98], теория (10.10) позволяет получить компактное стабильное пространство дополнительных измерений с ненулевой плотностью энергии вакуума. Наблюдаемое малое значение энергии вакуума требует чрезвычайно тонкой подгонки параметров C и Λ . Ниже предлагается механизм, позволяющий естественным образом варьировать параметры эффективной теории в широких пределах. При этом в каждой вселенной, образующейся в результате квантовой флуктуации, значения параметров уникальны. Это означает, что существует доля вселенных с требуемыми значениями параметров и, следовательно, энергии вакуума. Таким образом, малость плотности темной энергии в нашей Вселенной означает малость доли вселенных, похожих на нашу.

Вернемся к вопросу о механизме вариации параметров. Следующие соотношения имеют место благодаря специальной форме выбранной метрики (10.9), [110]:

$$\sqrt{-G} = b^{d_1} \sqrt{-g} \sqrt{\gamma}, \quad (10.11)$$

$$R_D(G_{AB}) = R_{D_1}(g_{ab}) + b^{-2} R_{d_1}(\gamma_{ij}) + K_b, \quad (10.12)$$

$$K_b \equiv -2d_1 b^{-1} g^{ab} \nabla_a \nabla_b b - d_1(d_1 - 1) b^{-2} g^{ab} (\nabla_a b)(\nabla_b b). \quad (10.13)$$

Объем V_{d_1} внутреннего пространства M_{d_1} зависит от его метрики и топологии,

$$V_{d_1} = \int d^{d_1} y \sqrt{\gamma}. \quad (10.14)$$

Для удобства введем скалярное поле $\phi(x)$:

$$\phi(x) = b(x)^{-2} R_{d_1}(\gamma_{ij}), \quad (10.15)$$

в терминах которого действие принимает вид

$$\begin{aligned} S_D = & \frac{V_{d_1}}{2} \int d^{D_1} x \sqrt{-g} \left(\frac{R_{d_1}}{\phi(x)} \right)^{d_1/2} \times \\ & \times \left\{ R_{D_1}(g_{ab}) + 2C\phi(x)R_{D_1}(g_{ab}) + CR_{D_1}^2(g_{ab}) - \right. \\ & - 2\Lambda + \phi(x) + C\phi^2(x) + \\ & \left. + 2CK_b [R_{D_1}(g_{ab}) + \phi(x)] + K_b + CK_b^2 \right\}. \quad (10.16) \end{aligned}$$

Поскольку поле ϕ однозначно связано с радиусом кривизны $b(x)$ компактного пространства M_{d_1} , наличие стационарных решений $\phi(x) = \phi_m$, минимизирующих действие, означало бы стабильность размеров этого пространства.

В картине Йордана, согласно (10.16), имеем

$$\begin{aligned} S_D = & \frac{1}{2} \int d^{D_1} x \sqrt{-g} \{ f(\phi) R_{D_1}(g_{ab}) + C_J(\phi) R_{D_1}^2(g_{ab}) - \\ & - 2U_J(\phi) + K_J(\phi) \}, \quad (10.17) \end{aligned}$$

$$f(\phi) = \mathcal{V}_{d_1} \phi^{-d_1/2}(x) [1 + 2C\phi(x)], \quad (10.18)$$

$$U_J(\phi) = \mathcal{V}_{d_1} \phi^{-d_1/2}(x) \left[\Lambda - \frac{1}{2} (\phi(x) + C\phi^2(x)) \right], \quad (10.19)$$

$$C_J(\phi) = \mathcal{V}_{d_1} \phi^{-d_1/2}(x) C, \quad (10.20)$$

$$K_J(\phi) = \mathcal{V}_{d_1} \phi^{-d_1/2}(x) [K_b + 2CK_b(R_{D_1} + \phi(x)) + CK_b^2], \quad (10.21)$$

$$\mathcal{V}_{d_1} = V_{d_1} R_{d_1}^{d_1/2}. \quad (10.22)$$

Отметим явную зависимость параметров получившейся эффективной теории от геометрии дополнительного пространства. Масса m_ϕ скалярного поля $\phi(x)$ пропорциональна второй производной потенциала в минимуме (в картине Эйнштейна) и может

варьироваться в широких пределах. Отметим, что определенная таким образом масса является безразмерной величиной. Наиболее интересны три ситуации:

- (i) Минимум потенциала отсутствует, что означает нестационарность размера дополнительного пространства;
- (ii) Минимум потенциала существует и выполняется условие

$$m_\phi^2 \leq R_{D_1}(g_{ab}), \quad (10.23)$$

при котором скалярное поле эволюционирует вместе с метрикой “основного” подпространства M_{D_1} . Именно такая ситуация реализуется в инфляционных моделях и обсуждается в следующем разделе;

- (iii) Минимум потенциала существует и выполняется условие

$$m_\phi^2 \gg R_{D_1}(g_{ab}). \quad (10.24)$$

В этом случае скалярное поле $\phi(x)$ быстро оказывается в минимуме потенциала,

$$\phi(x) = \phi_m = \text{const}, \quad (10.25)$$

и не меняется со временем при низкоэнергетических процессах.

Последний вариант наиболее естествен, поскольку время релаксации пропорционально масштабу дополнительного пространства M_{d_1} , малому по сравнению с масштабом пространства M_{D_1} . Обсудим подробнее эту ситуацию. Полагая выполненным условие (10.25), проведем конформное преобразование вида

$$\begin{aligned} g_{ab} &= |f(\phi_m)|^{-2/(D_1-2)} \tilde{g}_{ab}, \\ R_{D_1} &= |f(\phi_m)|^{2/(D_1-2)} \tilde{R}_{D_1}, \\ \sqrt{-g} &= |f(\phi_m)|^{-D_1/(D_1-2)} \sqrt{-\tilde{g}}, \end{aligned} \quad (10.26)$$

которое, будучи примененным к выражению (10.17), возвращает нас к начальному виду действия

$$\begin{aligned}
S_{D_1} &= \frac{1}{2} \int d^{D_1} x \sqrt{-g} \{ R_{D_1}(g_{ab}) + C_{D_1} R_{D_1}(g_{ab})^2 - 2\Lambda_{D_1} \}, \\
C_{D_1} &= C_J(\phi_m) |f|^{(4-D_1)/(D_1-2)} = \\
&= (\text{sign } f) \mathcal{V}_{d_1} |f(\phi_m)|^{(4-D_1)/(D_1-2)} \phi_m^{-d_1/2} C, \\
\Lambda_{D_1} &= V_{D_1}(\phi_m) = \\
&= (\text{sign } f) \mathcal{V}_{d_1} |f(\phi_m)|^{-D_1/(D_1-2)} \phi_m^{-d_1/2} \left[\Lambda - \frac{1}{2}(\phi_m + C\phi_m^2) \right].
\end{aligned} \tag{10.27}$$

Поскольку рассматривается случай (10.25), когда поле ϕ уже находится в минимуме, кинетическое слагаемое проигнорировано.

Вид действия (10.27) совпадает с исходным (10.10), но уже в подпространстве M_{D_1} и с перенормированными параметрами C_{D_1} и Λ_{D_1} , зависящими от объема V_{d_1} и кривизны R_{d_1} дополнительного пространства M_{d_1} . В Таблицах 10.2.1, 10.2.1 приведены результаты численных расчетов, но предварительно сделаем необходимое замечание. Обычное предположение, $V_d \sim L^d$, где L — характерный размер пространства, справедливо для “простых” пространств с положительной кривизной типа d -мерных сфер. Для компактных гиперболических пространств ситуация более интересна. Связь объема и характерного размера определяется асимптотическим соотношением [179]

$$V_{d_1} \simeq \exp[(d_1 - 1)L/b_m], \quad L \gg b_m, \tag{10.28}$$

где L — наибольшее расстояние между точками многообразия M_{d_1} .

Очевидно, при достаточно большой размерности дополнительного пространства d_1 его объем может быть большим, а характерный размер пространства — малым. Таким образом, мы можем варьировать в широких пределах параметр V_{d_1} , а с ним и параметры C_{D_1} , Λ_{D_1} , не вступая в противоречие с экспериментальными ограничениями на размер дополнительного пространства. В качестве иллюстрации в таблицах 10.2.1 и

Таблица 10.1. Зависимость параметров C_{D_1} и Λ_{D_1} от геометрии дополнительного пространства (фактора \mathcal{V}_{d_1}). Значения фиксированных параметров: $D = D_1 + d_1 = 11$, $D_1 = 4$, $\Lambda = -0.6$, $C = -1.9$.

\mathcal{V}_{d_1}	10	1	10^{-1}	10^{-2}	10^{-3}
C_{D_1}	~ 0	$-4 \cdot 10^{-13}$	$-4 \cdot 10^{-6}$	-40	$-4 \cdot 10^8$
Λ_{D_1}	$-2 \cdot 10^{-4}$	$-2 \cdot 10^{-3}$	-0.016	-0.16	-1.6

Таблица 10.2. Зависимость параметров C_{D_1} и Λ_{D_1} от размерности подпространств M_{D_1} и M_{d_1} . В последней строке приведен радиус кривизны b_m дополнительного пространства M_{d_1} . Значения фиксированных параметров: $D = D_1 + d_1 = 40$, $\mathcal{V}_{d_1} = 100$, $\Lambda = -0.6$, $C = -1.9$.

D_1	4	10	20	30	38
C_{D_1}	~ 0	~ 0	~ 0	~ 0	$-3.5 \cdot 10^{-13}$
Λ_{D_1}	$-8.7 \cdot 10^{-10}$	$-8.4 \cdot 10^{-4}$	-0.025	-0.122	-0.343
b_m	50.2	44.5	39.2	32.8	13.0

10.2.1 представлена зависимость новых параметров C_{D_1} и Λ_{D_1} от топологии компактного дополнительного пространства M_{d_1} .

Итак, подпространство M_{D_1} оказывается оснащенной теорией, аналогичной изначальной, действующей в пространстве M_D , но с другими значениями параметров. Важно то, что, несмотря на неизменность исходных параметров (C, Λ) , эффективные, “вторичные” параметры $C_{\text{eff}} = C_{D_1}$ и $\Lambda_{\text{eff}} = \Lambda_{D_1}$ редуцированной теории меняются в широком диапазоне. Конкретные значения эффективных параметров зависят от случайной геометрии и топологии подпространств M_{d_1}, M_{D_1} , сформировавшихся благодаря квантовым флуктуациям. Число различных топологий пространства данной размерности по крайней мере счетно. Следовательно, исходная теория с произвольными, но фиксированными параметрами (в данном случае $C = -1.9$, $\Lambda = -0.6$),

порождает счетное множество редуцированных теорий в пространствах меньшей размерности M_{D_1} , различающихся значениями параметров. Интервал параметров C_{eff} и Λ_{eff} может быть еще более расширен, если принять во внимание последующую редукцию подпространства M_{D_1} в еще меньшее подпространство $M_{D_2} \in M_{D_1}$. Это сделано в следующем разделе.

Каскадное образование вселенных

Как было показано в предыдущем разделе, редукция исходной теории к пространству меньшей размерности генерирует широкий спектр вторичных теорий, различающихся значениями параметров лагранжиана C_{D_1} и Λ_{D_1} . Последние зависят от топологии пространства, возникающего в результате квантовых флуктуаций. Подпространство M_{D_1} , на котором строится эффективная теория (10.27), в свою очередь, так же, как и пространство M_D , подвержено квантовым флуктуациям, которые также приводит к его разбиениям вида

$$M_{D_1} = M_{D_2} \otimes M_{d_2}. \quad (10.29)$$

Последующие шаги образуют каскад

$$M_{D_1} \rightarrow M_{D_2} \otimes M_{d_2}; \quad M_{D_2} \rightarrow M_{D_3} \otimes M_{d_3} \rightarrow \dots \rightarrow M_3 \otimes M_{d_{\text{final}}}. \quad (10.30)$$

В отличие от начального этапа, параметры лагранжиана на промежуточных этапах зависят от предыдущих этапов каскада. Поскольку число топологий на каждом этапе по крайней мере счетно, цепочка (10.30) быстро разветвляется, образуя в конечном итоге бесконечный набор эффективных теорий, отличающихся величиной параметров. Имеется бесконечное число путей “спуска” от начального пространства M_D к конечному. Интересующие нас каскады оканчиваются образованием четырехмерных пространств $R \otimes M_3$ и дополнительных компактных пространств с некоторым числом измерений, равным d_{final} . Очевидно, что

только малая часть образующихся таким образом вселенных по-
хожи на нашу Вселенную.

Обсудим появление планковского масштаба в данной кар-
тине. Поскольку до сих пор мы имели дело с чисто геомет-
рическими свойствами пространств, введение любого масштаба
представлялось искусственным. На последнем же этапе каска-
да (10.30) возникает наше четырехмерное пространство, а ком-
поненты метрического тензора дополнительного пространства
 $M_{d_{\text{final}}}$ воспринимаются будущими наблюдателями как скаляр-
ные и векторные поля. Следовательно, в отличие от предыду-
щего, кинетическим слагаемым K_J в действии (10.17) прене-
брегать нельзя. Используя низкоэнергетический предел (10.7),
пренебрежем высшими степенями скалярной кривизны R_4 , а в
разложении $K_J = K(\varphi)\partial_\mu\varphi\partial\varphi^\mu + \dots$ оставим только первые чле-
ны. Очевидно, что в этом случае наиболее общий вид действия
в картине Эйнштейна есть

$$S \simeq \frac{V_{d_1}}{2} \int d^4x \sqrt{-g} [R_4 + K(\varphi)(\partial\varphi)^2 - 2V(\varphi)], \quad (10.31)$$

где объем дополнительного пространства V_{d_1} выписан явно. Вид
функций $K(\varphi)$ и $V(\varphi)$ зависит от топологии, и геометрии до-
полнительного пространства (см. [94, 98]). Мы же продолжим
анализ в общем виде.

Осцилляции поля φ вокруг положения равновесия $\varphi = \varphi_m$
воспринимаются наблюдателем как кванты скалярного поля.
Вблизи минимума потенциала, $\min V(\varphi) = V(\varphi_m) = V_m$, дей-
ствие (10.31) имеет вид

$$S \simeq \frac{V_{d_1}}{2} \int d^4x [R_4 - K(\varphi_m)(\partial\varphi)^2 - 2V(\varphi_m) - V''(\varphi_m)(\varphi - \varphi_m)^2].$$

Масса m_φ квантов скалярного поля измеряется наблюдате-
лем в размерных единицах. В то же время она непосредственно
связана с формой эффективного потенциала в эйнштейновской
картине и при стандартной форме кинетического члена. Если

кинетический член $K(\phi) > 0$ в области минимума потенциала $V(\phi)$, то замена переменных вида

$$\begin{aligned} x_{\text{phys}} &= x \cdot \frac{\sqrt{V''(\varphi_m)}}{m_\varphi \sqrt{K(\varphi_m)}}, \\ \varphi_{\text{phys}} &= (\varphi - \varphi_m) \cdot m_\varphi \sqrt{V_{d_1}} \frac{K(\varphi_m)}{\sqrt{V''(\varphi_m)}} \end{aligned} \quad (10.32)$$

приводит к теории с обычным кинетическим членом

$$\begin{aligned} S = \frac{1}{2} \int d^4 x_{\text{phys}} \left[V_{d_1} m_\varphi^2 \frac{K(\varphi_m)}{V''(\varphi_m)} \cdot R_4 - \right. \\ \left. - (\partial \varphi_{\text{phys}})^2 - m_\varphi^2 \varphi_{\text{phys}}^2 - 2\Lambda \right], \end{aligned} \quad (10.33)$$

$$\Lambda \equiv V(\varphi_m) V_{d_1} m_\varphi^4 \frac{K^2(\varphi_m)}{V''(\varphi_m)^2}. \quad (10.34)$$

Если ввести обозначение

$$M_{\text{Pl}}^2 = V_{d_1} m_\varphi^2 \frac{K(\varphi_m)}{V''(\varphi_m)}, \quad (10.35)$$

то мы получим стандартный вид действия для некоторого скалярного поля χ в размерных единицах:

$$\begin{aligned} S = \int d^4 x \sqrt{-g} \left[\frac{M_{pl}^2}{2} R_4 + \frac{1}{2} (\partial \chi)^2 - U(\chi) \right], \\ U(\chi) = \frac{1}{2} m_\varphi^2 \chi^2 + \Lambda. \end{aligned} \quad (10.36)$$

При этом наблюдаемая масса Планка связана с параметрами теории посредством (10.35).

Таким образом, на последнем этапе каскада появляется сразу три масштаба — радиус кривизны дополнительного пространства b_m , масса Планка M_{pl} , см. (10.35), и энергия вакуума Λ , см. (10.34), связанные с массой скалярного поля m_φ . Как видно, в

предложенном подходе масса Планка не является фундаментальной постоянной и зависит от конечной конфигурации каскада.

Космологическая постоянная $\Lambda \propto V(\varphi_m)$ зависит от эффективных параметров. Последние могут меняться в широких пределах, как это следует из предыдущего обсуждения, см. таблицы 10.2.1, 10.2.1. Значит, и величина космологической постоянной в пространстве M_4 варьируется в широких пределах в зависимости от свойств каскада (10.30), в результате которого образовалось это пространство. Некоторые каскады могут, видимо, приводить к пространствам M_4 с наблюдаемым значением Λ -члена. Чрезвычайная малость последнего означает лишь что только малая доля каскадов приводит к желаемому результату. “Тонкая настройка” параметров Вселенной происходит при выборе соответствующего каскада.

Определим отношение массы Планка к массе инфлатона

$$\frac{M_{pl}}{m_\varphi} = \sqrt{V_{d_1} \frac{K(\varphi_m)}{V''(\varphi_m)}}. \quad (10.37)$$

Выбрав подпространство M_{10} из набора (10.4) и задав параметры $C = -1.9$ и $\Lambda = -0.6$, т.е. такими же, как и в таблицах 10.2.1, 10.2.1, можно получить

$$M_{pl}/m_\varphi \simeq 2 \cdot 10^6$$

для размерности дополнительного пространства $d_1 = 6$. Полученное численное значение хорошо согласуется с тем, которое обычно и используется в инфляционных моделях.

Вероятность рождения вселенных с дополнительными компактными измерениями

До сих пор мы предполагали, что на структуру пространства наложены свойства (10.5)–(10.8). В данном разделе обсуждается возможность появления подобных топологий в результате квантовых флуктуаций. Проблеме квантового рождения Вселенной

посвящено множество работ. В данном случае ситуация осложняется тем, что рассматривается нелинейный вариант гравитации и, кроме того, имеются дополнительные измерения, стабильность которых должна учитываться особо. Рождение n -мерного пространства с дополнительными измерениями в рамках стандартной гравитации рассматривалось в [112, 174], где изучались также и области стабильности компактного подпространства. В статьях [112, 157] изучалась возможность инфляции при наличии дополнительных измерений. Квадратичная по скаляру Риччи гравитация исследовалась в этом аспекте в [286].

Вероятности рождения Вселенной, получаемые в разных подходах, кардинально отличаются [295], что может указывать как на несовершенство современной теории, так и на сложность предмета. Конечной целью таких вычислений является определение вероятности появления вселенной нашего типа. Вряд ли стоит ожидать, что эта вероятность высока, учитывая факт тонкой настройки параметров Вселенной. В таком случае вычисление вероятности представляет чисто академический интерес, ввиду отсутствия причинных связей между вселенными. По-видимому, на современном этапе необходимым и достаточным является доказательство того, что доля вселенных нашего типа не равна нулю в рамках конкретного подхода. В нашем случае это означает, что вероятность каждого перехода (10.6) в каскаде отлична от нуля.

Основной вклад в амплитуду перехода (10.3) дают классические траектории, на которых действие стационарно. Их форма зависит от граничных условий и, в частности, от свойств многообразия Σ_f . В нашем случае метрика на гиперповерхности Σ_f определяется условиями (10.6), (10.7), (10.8). Поэтому будем искать классические траектории, подчиняющиеся тем же условиям на любом сечении Σ между сечениями Σ_{in} и Σ_f . Начальная гиперповерхность Σ_{in} может либо вовсе отсутствовать (подход Хартла — Хокинга), либо иметь “нулевую геометрию” (интервал между двумя любыми точками этой гиперповерхности равен ну-

лю — подход Виленкина). Ниже показано, что вероятность перехода слабо зависит от свойств гиперповерхности Σ_{in} .

В качестве примера рассмотрим вероятность формирования структуры

$$\Sigma_f = M_3 \otimes M_{d_{\text{final}}}, \quad (10.38)$$

возникающей на последнем этапе каскада. Основной вклад в амплитуду перехода вносят классические траектории. Смена топологии при классическом движении маловероятна. Поэтому для нас будут важны классические траектории, состоящие из гиперповерхностей, также удовлетворяющих условию (10.38). Тогда топология D -мерного риманова пространства между сечениями Σ_{in} и Σ_f есть

$$\mathbb{R} \otimes M_3 \otimes M_{d_{\text{final}}}. \quad (10.39)$$

Как и ранее, см. (10.7), предполагается выполненным неравенство $R_3 \ll R_{d_{\text{final}}}$, которое позволяет воспользоваться результатами предыдущего раздела. Действительно, в использованном приближении действие (10.10) трансформируется в теорию вида (10.31), а затем и в привычное действие Эйнштейна-Гильберта (10.36).

Действие (10.36) использовалось неоднократно для изучения проблемы квантового рождения Вселенной, см., например, [139, 161, 209, 294, 295, 304]. Разница заключается в том, что в работах, посвященных квантовому рождению Вселенной, наличие скалярного поля обычно постулируется, в то время как в нашем подходе оно появляется из компонент метрического тензора дополнительного пространства. Поэтому можно воспользоваться результатами многочисленных исследований, кратко воспроизводя их основные результаты.

Квантовое рождение Вселенной изучается обычно в рамках минисуперпространства, в котором интервал имеет вид, [295]

$$\begin{aligned} ds^2 &= \sigma^2 [N(t)^2 dt^2 - a(t)^2 d\Omega_3^2], \\ \sigma^2 &= \frac{1}{12\pi^2 M_{\text{Pl}}^2}. \end{aligned} \quad (10.40)$$

Волновая функция $\psi(a)$ удовлетворяет уравнению Уилера–ДеВитта

$$\left[\frac{\partial^2}{\partial a^2} - W(a) \right] \psi(a) = 0, \quad (10.41)$$

с потенциалом

$$W(a) = a^2(1 - H^2 a^2), \quad a > 0, \quad H = \frac{\sqrt{U(\chi)}}{6\pi M_{\text{Pl}}^2}.$$

Рождение вселенной описывается подбарьерным переходом с запрещенной областью

$$0 < a < H^{-1}. \quad (10.42)$$

Волновая функция в этой области имеет вид [295]

$$\psi(a) \simeq \exp \left[\int_a^{H^{-1}} \sqrt{-2W(a')} da' \right]. \quad (10.43)$$

Интеграл, входящий в это выражение, плохо определен на нижнем пределе, где $a \rightarrow 0$. В этой области не работает приближение $R_3 \ll R_{d_{\text{final}}}$, поскольку $R_3 = k/a^2 \rightarrow \infty$, и явное выражение для потенциала не определено. Эта же проблема имеет место и в других моделях квантового рождения Вселенной [226]. Тем не менее, вычисленный таким образом интеграл имеет смысл в пределе

$$H \ll M_{\text{Pl}}, \quad (10.44)$$

когда область $a \sim 0$ мала по сравнению со всей областью интегрирования. Наша Вселенная образовывалась при $H \sim 10^{-6} M_{\text{Pl}}$, так что неравенство (10.44) выполняется уже на инфляционной стадии. Вывод о слабой зависимости результата от поведения функции вблизи сингулярности подтверждается и в работах других авторов. Так, начальная волновая функция вида $\delta(a - a_{\text{in}})$ предлагалась в [286]. Распад метастабильного вакуума из состояния с фиксированной энергией изучался в [139, 304]. В обоих

случаях показано, что начальные условия слабо влияют на вероятность перехода. Подробное обсуждение квантового рождения вселенных в многомерной гравитации содержится в [112, 143].

В подходе Виленкина вероятность рождения вселенной есть $dP \propto \exp[+2/(3U(\chi))]$, в то время как приближение Хартла — Хокинга дает $dP \propto \exp[-2/(3U(\chi))]$. Величина скалярного поля χ однозначно связана с размером дополнительного пространства, и формулы дают вероятность рождения дополнительных измерений в зависимости от их характерного размера. При всем различии, главным является то, что вероятность события отлична от нуля, и, следовательно, доля вселенных с заданными свойствами, образующихся в результате каскада редукций, отлична от нуля.

Обсуждение

Существуют проблемы, решение каждой из которых представляет собой серьезную задачу. К ним относятся проблемы:

дополнительных измерений — их число, способ компактификации, возможность экспериментальной проверки;

нелинейности гравитационного действия, неизбежно возникающей в результате квантовых эффектов;

численных значений параметров теории, приводящих к образованию Вселенной со сложной структурой.

В статье [31] показана взаимосвязь этих проблем. При этом оказывается, что совместное их рассмотрение не усложняет, а облегчает понимание путей их решения. Процесс каскадной редукции (10.30) позволяет получить по меньшей мере счетное множество вселенных с различными свойствами. На конечном этапе каскада образуется вселенная, описываемая эффективной низкоэнергетической теорией с параметрами, значения которых определяются конкретным каскадом. Значения параметров варьируются в широких пределах, что указывает на возможность решения проблемы тонкой настройки Вселенной. Сходные идеи разрабатываются в рамках теории струн, см., например, [207],

[178]. В отличие от последней, подход, предложенный нами, базируется лишь на предположении о многомерности пространства, без введения дополнительного постулата о существовании (супер)струн.

Масса Планка, размер дополнительных измерений и космологическая постоянная оказываются зависящими от конкретного каскада и способа определения масс частиц наблюдателем.

Выбор подходящего каскада, образовавшего нашу Вселенную, здесь не обсуждается. Нерешенной остается также проблема выбора начального лагранжиана. В качестве примера нами рассмотрен простейший нелинейный по кривизне лагранжиан. Используя подход, предложенный в [94, 98], нетрудно обобщить рассмотрение на более сложные формы исходного лагранжиана.

В течение многих лет предполагалось, что мы находимся в пространстве Фридмана — Робертсона — Уокера с равной нулю космологической постоянной. С астрономической точки зрения это означает расширение Вселенной с небольшим отрицательным ускорением. На языке квантовой теории поля это означает строгое равенство нулю плотности энергии основного состояния. Причины того, что минимальное значение потенциальной энергии равно нулю, были не очень ясны, но оставалась некоторая надежда на скрытые симметрии (см., напр., обзор [270]).

Наблюдательные данные [257] позволили определить плотность темной энергии $\rho_\Lambda \approx 0.7\rho_m$, оказавшуюся лишь немногим меньше средней плотности материи ρ_m во Вселенной.

С точки зрения квантовой теории поля это неправдоподобно мало, поскольку все квантовые эффекты, которые дают вклад в энергию вакуума, на много порядков превосходят это значение, а механизм взаимного почти полного сокращения вкладов совершенно неочевиден [270]. И в то же время, если бы космологическая постоянная оказалась бы всего в 200 раз больше, то галактики не успели бы сформироваться [301] и жизнь была бы невозможна. Создается впечатление, что физические параметры Вселенной специально подобраны так, чтобы в ней обязательно воз-

никла жизнь. Интересно, что принципиальное решение этой проблемы более или менее ясно, но совершенно непонятен конкретный механизм, лежащий в основе этого решения. По-видимому, у природы имеется набор Вселенных с разными свойствами, которые осуществляются с некоторой вероятностью, и некоторые из них оказываются подходящими для существования жизни. Так, в работах [149, 302] вводится вероятность обнаружения Вселенной с определенной плотностью энергии вакуума и обсуждаются следствия этого с точки зрения антропного принципа.

Описанная выше настройка космологической постоянной не является единственной и даже первой проблемой такого рода. Физика элементарных частиц предоставляет нам большое количество аналогичных проблем. Достаточно напомнить лишь одну из них — малая масса электрона. Если бы масса электрона, отличающаяся от массы нуклона на 3 порядка, была бы всего лишь втрое больше, то нейтроны были бы стабильны, и процесс $p^+ + e^- \rightarrow n + \bar{\nu}$ привел бы к резкому падению числа протонов во Вселенной, что полностью изменило бы ход ее эволюции. Мы видим, что наша Вселенная оказывается выделенной по многим параметрам, один из которых — космологическая постоянная.

Далее рассматривается механизм возникновения большого числа вселенных и их распределение по физическим параметрам. Показано, что в результате квантовых флуктуаций вселенная с любыми наперед заданными параметрами возникает с неизбежностью.

10.2.2. Случайный потенциал

Обычно при создании какой-либо теории постулируется конкретный вид лагранжиана. Константы взаимодействия предполагаются малыми, так что квантовые поправки к первоначально выбранному лагранжиану считаются также малыми. Важно в данном случае то, что за счет квантовых поправок возникает бесконечное число дополнительных слагаемых в первоначально простом лагранжиане. Поправки действительно малы лишь для сла-

бых полей, в случае сильных полей это далеко не так.

В качестве примера рассмотрим лагранжиан скалярного поля с самодействием

$$L = \frac{1}{2}(\partial_\mu\varphi)^2 - \frac{m^2}{2}\varphi^2 - \frac{\lambda}{4}\varphi^4. \quad (10.45)$$

Однопетлевые квантовые поправки к потенциалу имеют вид [210]

$$\delta V = \frac{(3\lambda\varphi^2 + m^2)^2}{64\pi^2} \ln \frac{(3\lambda\varphi^2 + m^2)}{2m^2} - a\varphi^2 - b\varphi^4. \quad (10.46)$$

Последние два слагаемых перенормируют массу и константу связи исходного потенциала и зависят от способа нормировки. Первое же слагаемое изменяет сам вид потенциала. Многопетлевые поправки добавляют новые слагаемые в потенциал.

Нетрудно видеть, сравнивая выражения (10.45) и (10.46), что новые слагаемые малы по сравнению с первоначальными в области $\varphi \ll m \cdot \exp(1/\lambda)$. Если выбрать для оценки значения $m = 100$ ГэВ, $\lambda = 0.1$ (здесь подразумевается, что поле не является инфлатонным), то квантовые поправки к потенциалу становятся большими при $\varphi \sim 10^6$ ГэВ. Это достаточно большая величина при изучении взаимодействий на ускорителях, однако на ранней, инфляционной стадии развития нашей Вселенной поля достигали значений $\varphi > 10^{19}$ ГэВ, и при выбранных значениях параметров необходимо учитывать весь бесконечный ряд дополнительных слагаемых в лагранжиане (10.45). Более того, ограничение на амплитуду скалярного поля еще сильнее. Действительно, логарифм в выражении (10.46) является результатом суммирования бесконечного ряда, который сходится лишь при $\varphi < m/\sqrt{3\lambda}$. Кроме того, уже из вида лагранжиана (10.45) видно, что потенциал взаимодействия становится сравнимым с массовым членом при $\varphi \sim m\sqrt{2/\lambda}$. Две последних оценки хорошо согласуются друг с другом и дают гораздо меньшее значение амплитуды поля, при котором можно говорить о малости квантовых поправок.

Итак, при рассмотрении явлений, связанных с сильными полями, когда $\varphi > m/\sqrt{\lambda}$, необходимо учитывать все дополнительные слагаемые, неизбежно возникающие при учете квантовых поправок. Вид потенциала становится гораздо более сложным, чем можно предполагать, основываясь на теории в ее низкоэнергетическом пределе. Так, в горной местности можно говорить о гладких поверхностях с малой кривизной, лишь находясь в долинах, т.е. в минимумах потенциальной энергии. И только поднявшись на некоторую высоту, становится очевидным, что рельеф гораздо более сложен.

Обычно потенциал взаимодействия скалярного поля постулируется в самой простой форме, подходящей для целей конкретного исследования. При этом требование перенормируемости теории не обязательно, поскольку предполагается, что учет гравитационных эффектов на планковском масштабе приведет к регуляризации интегралов. В обычных условиях поля действительно слабы, и учет квантовых поправок сводится к перенормировке параметров лагранжиана и предположению, что конечные поправки малы. Из приведенных выше аргументов следует, что в момент образования нашей Вселенной, при больших амплитудах поля, квантовые поправки были скорее всего сравнимы с основными слагаемыми, а вид лагранжиана гораздо более сложен, чем предполагается изначально. Поэтому, вместо того чтобы постулировать простой вид лагранжиана и пытаться доказать, что он справедлив и при больших энергиях, сделаем прямо противоположное. Именно, примем как факт, что потенциал является полиномом, содержащим все степени поля φ . Вклад в перенормировку каждого слагаемого является результатом сложной интерференцией вкладов от взаимодействия со всеми сортами полей. Приведенные рассуждения приводят к мысли, что конкретные численные значения полиномиальных коэффициентов слабо связаны между собой. Предлагаемый здесь постулат представляет собой следующее утверждение:

★ *Единственной корреляцией между полиномиальными коэффици-*

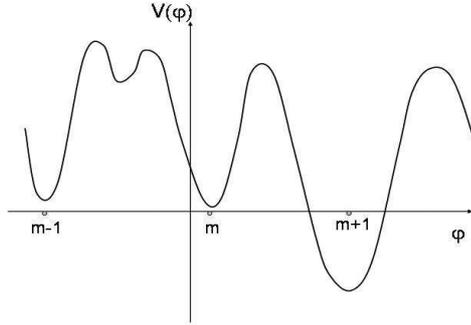


Рис. 10.1. Возможная форма эффективного потенциала

ценностями потенциала является условие $|V(\varphi)| \leq M_{\text{Pl}}^4$.

Последнее неравенство позволяет избежать обсуждения проблем квантования гравитации на планковских масштабах. Далее мы покажем, что этот постулат приводит к интересным последствиям и внутренне непротиворечив. В качестве рабочего примера рассмотрим лагранжиан скалярного поля вида

$$L = \frac{1}{2}(\partial_\mu \varphi)^2 - V(\varphi), \quad V(\varphi) = \sum_{n=0}^{\infty} a_n \frac{\varphi^n}{M_{\text{Pl}}^{n-4}}. \quad (10.47)$$

Потенциалы с произвольной, но конечной степенью полинома используются в рамках модели $\Lambda\lambda$ — инфляции [217]. Поле φ определено в интервале $(-\infty, +\infty)$. Множители при степенях полинома $V(\varphi)$ представляют собой случайные числа. Закон распределения случайных чисел принципиального значения не имеет.

На рис. 10.1 представлен характерный график потенциала на некотором интервале изменения поля φ .

В каждом из минимумов потенциал $V(\varphi)$ аппроксимируется естественным образом: $V(\varphi) \approx V(\varphi_m) + a\phi^2 + b\phi^4$, $\phi = \varphi - \varphi_m$.

Обычно подобный потенциал постулируется изначально с определенными константами a и b , причем константа a связана с массой поля φ , $a = m_\varphi^2/2$, если $a > 0$. Наша Вселенная расположена в одном из таких минимумов. Отметим, что лагранжиан (10.46) является частным случаем более общего лагранжиана, в котором учитывались бы квантовые поправки к кинетическому члену. В следующем разделе изучаются космологические следствия введенных постулатов.

Квантовые флуктуации как генератор вселенных

Все (квази-)стационарные состояния расположены в минимумах потенциала, и наша Вселенная, не являясь исключением, расположена в одном конкретном минимуме. Поскольку имеется счетное множество минимумов, каждый из которых характеризуется своей плотностью энергии, кажется маловероятным, что наша Вселенная оказалась именно в минимуме с малой плотностью энергии. Для оценки этой вероятности предположим, что вероятность обнаружить минимум потенциала с плотностью энергии $\rho_V^{(m)} = V(\varphi_m)$ в интервале $d\rho_V^{(m)}$ есть

$$dP(\rho_V^{(m)}) = d\rho_V^{(m)} / M_{\text{Pl}}^4 \quad (10.48)$$

(т.е. предполагается равномерное распределение $\rho_V^{(m)}$ на всем интервале изменения $(0, M_{\text{Pl}}^4)$). Ожидаемое значение плотности энергии вакуума в нашей Вселенной составляет $\rho_V \approx 10^{-123} M_{\text{Pl}}^4$. Следовательно, доля вселенных с плотностью энергии вакуума, подобной нашей, есть $\sim 10^{-123}$. Тем не менее множество таких Вселенных счетно, поскольку таковым является множество всех минимумов. Квантовые флуктуации, возникающие, в частности, и в области этих минимумов, в ходе инфляции порождают счетное множество вселенных нашего типа.

Возникшая причинно связанная область со временем дробится на все большее число причинно не связанных областей, некоторые из которых за счет полевых флуктуаций движутся к ближайшему максимуму потенциала. Рассмотрим флуктуацию поля

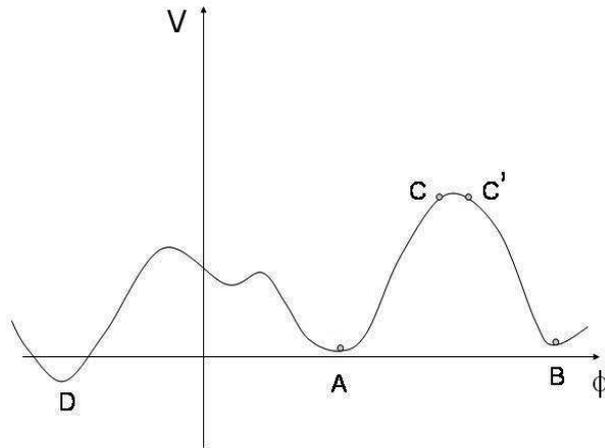


Рис. 10.2. Образование фрактальной структуры при квантовых флуктуациях скалярного поля вблизи максимума потенциала.

вблизи максимума (точка C на рис. 10.2). Спустя время, равное $1/H$, эта пространственная область разделится на e^3 причинно не связанных областей со своим значением поля внутри них. Среднее значение поля φ в некоторых из этих областей окажется по другую сторону максимума (точка C' на рис. 10.2). В свою очередь, каждая из этих областей через время $1/H$ окажется состоящей из e^3 причинно не связанных областей, некоторые из которых пройдут через максимум потенциала в обратную сторону. Подобный процесс непрерывно воспроизводит сам себя, и уже после нескольких шагов в области максимума будет наблюдаться картина фрактального типа.

Известно, что соседние области, расположенные по разные стороны от максимума потенциала, разделены стенкой [251]. Поля в таких областях скатываются к различным (соседним) минимумам, при этом плотность энергии стенки возрастает. Таким образом, соседние Вселенные оказываются разделенными полевыми стенками с большой плотностью энергии.

Итак, из последовательного применения постулата (\star) следует, что наша Вселенная образовалась из одной области, ограниченной замкнутой стенкой. Инфляционный механизм обеспечил экспоненциальное увеличение ее размеров (с точки зрения внутреннего наблюдателя), так что современный размер видимой части Вселенной $\sim 10^{28}$ см оказывается на много порядков меньше характерного масштаба замкнутой области $\sim 10^{10^{12}}$ см [213]. Следовательно, такие стенки ненаблюдаемы.

Рассмотренный механизм позволяет подбирать вселенные и по другим параметрам, например, по массе частиц. Обычно взаимодействие скалярного поля с фермионами постулируется в виде юкавской связи

$$V_F = g\varphi\bar{\psi}\psi. \quad (10.49)$$

При этом мы сталкиваемся с серьезной трудностью. Дело в том, что подходящий минимум потенциала может оказаться достаточно далеко от значения $\varphi = 0$. Следовательно, вклад в массу фермиона $\mu_F = g\varphi_m$, взаимодействующего с инфлатоном, окажется аномально большим. Даже если прямое взаимодействие вида (10.49) отсутствует изначально, оно появляется за счет квантовых поправок. Проблема разрешается, если учесть, что квантовые поправки, в полной аналогии с предыдущим, модернизируют потенциал взаимодействия и приводят его к виду

$$V_F = G(\varphi)\bar{\psi}\psi, \quad (10.50)$$

который является естественным обобщением выражения (10.49). Функция $G(\varphi)$ есть полином со случайными коэффициентами, по аналогии с основным потенциалом $V(\varphi)$. При таком выборе масса фермиона μ_F и константа g его взаимодействия с полем $\phi = \varphi - \varphi_m$ оказываются зависящими от номера Вселенной (или, точнее, от положения m -го минимума потенциала φ_m):

$$\mu_F = G(\varphi_m); \quad g = G'_\varphi(\varphi_m)(\varphi - \varphi_m) + \dots \quad (10.51)$$

Очевидно, для любого заданного значения интервала масс фермиона $(\mu_F, \mu_F + \delta)$ и функции $G(\varphi)$ из счетного множества

вселенных можно подобрать вселенную с таким потенциалом в минимуме $V(\varphi_m)$, что выполняется равенство $\mu_F \cong G(\varphi_m)$. Это позволяет использовать данный механизм также и для тонкой настройки Вселенной по другим параметрам, а не только по плотности энергии вакуума. Действительно, пусть, например, необходимо, чтобы масса данного фермиона находилась в интервале $(\mu_F, \mu_F + \delta\mu_F)$. Тогда из счетного множества вселенных с подходящей плотностью энергии вакуума всегда найдется вселенная с подходящим значением $G(\varphi_m)$, таким, что масса фермиона попадет в заданный интервал. Более того, таких вселенных — счетное множество; из них выбираются вселенные, подходящие для жизни по другим параметрам. Происходит автоматический “поиск” вселенных, подходящих для жизни по всем параметрам. Процесс образования каждой из них уникален, поскольку уникальна форма потенциала в области каждого минимума. Следовательно, процессы образования различных вселенных должны описываться разными инфляционными теориями. На данный момент известно достаточно большое число реалистичных инфляционных моделей. По-видимому, любая из них описывает некоторое подмножество вселенных нашего типа.

Введем конечный набор физических параметров ℓ_k , $k = 1, 2, \dots, N_{\text{lif}}e$, необходимых для рождения жизни нашего типа во Вселенной. Пусть $\mathfrak{R}(\{\ell\}_n)$ обозначает множество вселенных, обладающих ограниченным набором из n таких параметров $\ell_1, \ell_2, \dots, \ell_n$. Тогда процесс отбора Вселенной, подходящей для рождения жизни, может быть представлен в виде

$$\mathfrak{R}(\{\ell\}_0) \mapsto \mathfrak{R}(\{\ell\}_1) \mapsto \mathfrak{R}(\{\ell\}_2) \mapsto \dots \mapsto \mathfrak{R}(\{\ell\}_{N_{\text{lif}}e}). \quad (10.52)$$

Каждое последующее множество есть подмножество предыдущего, но тем не менее остается счетным множеством. Если число параметров, численные значения которых важны для появления жизни, конечно, $N_{\text{lif}}e < \infty$, то процесс (10.52) приведет нас к счетному множеству вселенных, в которых возможно появление жизни.

Обсудим, наконец, механизм Хиггса появления масс легких фермионов. Поскольку хиггсовское поле взаимодействует с инфлатонным полем (не обязательно в первом порядке по \hbar), то очевидно, что потенциал поля Хиггса также должен содержать бесконечный ряд степеней этого поля, свой для каждой вселенной. Очевидно, существует счетное множество вселенных, где потенциал поля Хиггса имеет два близлежащих минимума, отделенных максимумом подходящей высоты. Очевидно также, что массы фермионов, возникшие благодаря их взаимодействию с полем Хиггса в любых двух вселенных, являются некоррелированными величинами. В свою очередь, это означает, что квантовые поправки к исходному потенциалу (10.47) от взаимодействия с этими фермионами, также некоррелированы. Таким образом, постулат (*) непротиворечив.

На основе предложенного подхода можно также делать и некоторые количественные предсказания. Основным моментом здесь является то, что мера множества вселенных с точным значением любого из параметров равна нулю. Поэтому, например, строгое равенство нулю космологического члена в данной схеме исключено. Аналогично этому, не должно существовать скалярных вещественных полей, имеющих потенциал с вырожденными минимумами. Последнее утверждение выполняется и для других сортов полей в отсутствие изначально заданных симметрий.

Заключение

Спор между сторонниками множественности вселенных (М) и Единой теории (Е)

Е: Все развитие науки направлено на объединение постулатов. Успехи налицо (например, Стандартная модель). Вы же хотите полностью отказаться от этого?

М: А почему Вы считаете, что подобный путь будет успешным всегда? Стремление создать “теорию всего” — не похоже ли на стремление создать теорию массы планеты Земля? Да и Стандартная модель была создана десятки лет назад. С тех пор продвижения нет.

Е: Ну, разработка инфляционной идеи — тоже существенный прогресс. По крайней мере мы движемся в позитивном направлении — объединяя теории и проверяя выводы экспериментально. А как Вы можете проверить теорию множественности миров? Если, как Вы сами признаете, попасть в иную Вселенную невозможно, то никаких доказательств правоты не существует. А значит, это не лучше, чем создавать теорию массы Земли.

М: Надо признать, это действительно кажется слабым местом идеи о множественности Вселенных. Правда, в таком случае и, например, теория фотоэффекта также не имеет права на существование — ведь Вы не можете попасть внутрь твердого тела и убедиться, что фотон действительно поглощается электроном.

В обоих случаях приходится принимать опосредствованные экспериментальные доказательства. С другой стороны, Ваш путь построения единой теории никогда не приведет к конечному результату. Следующим вопросом будет, почему Ваша Теория Всего содержит именно такие параметры, а не другие. Например, если М-теория объяснит все существующие явления, то следующим вопросом будет — почему дополнительных измерений 11 или 10?

Е: А кто сказал, что процесс познания конечен? Асимптотическое приближение к Единой теории — тоже не самый плохой вариант. По крайней мере по сравнению с тем, что предлагаете Вы. Если все теории возможны, то зачем вообще изучать что-либо?

М: Не совсем так. Ведь существование множества звезд не мешает детальному исследованию нашего Солнца. Много чего еще осталось понять и в нашей собственной Вселенной. Ну а асимптотическое стремление к Единой теории — действительно достойное занятие. Только в этом направлении не видно путей ответа на экспериментальный факт: тончайший подбор параметров Единой теории, как будто специально для того, чтобы возник разум.

Е: Единая теория должна будет ответить и на этот вопрос, который в рамках множественности вселенных выглядит тривиальным. И все-таки, напоследок: где же расположены все эти вселенные?

М: Это действительно сложный, почти философский вопрос. Ответить на него мы сможем, когда поймем, где расположена наша Вселенная.

Почему правильные теории кажутся красивыми?

По-видимому, критерием красоты служит степень экономности средств, которыми достигается цель — объяснение некоторого

круга явлений и предсказание новых. Выразим это в виде простой формулы

$$R = N/P,$$

где R — коэффициент качества теории: чем он выше, тем красивее нам кажется теория, N — число объясненных природных явлений, P — количество информации (в битах), использованной для начальных постулатов теории.

Теперь вопрос можно сформулировать по-другому: почему теории с большим значением R оказываются верными? Приняв идею множественности вселенных, проблему красоты теорий можно свести к следующему вопросу: почему доля вселенных с большим коэффициентом R велика по сравнению с долей других вселенных? В этом случае было бы неудивительно, что наша цивилизация возникла именно в такой вселенной. Итак, надо обосновать, почему доля вселенных

$$N_{\text{Universe}} \sim R^a, \quad a > 0.$$

Разумно считать, что чем сложнее устроена конкретная вселенная, тем реже вселенные такого типа встречаются. Конечно, разумная жизнь не возникает в совсем простых вселенных, каких большинство. Есть некий порог сложности, ниже которого разум не зарождается. Их число много меньше, чем простых вселенных, но только они и интересны. Другими словами, если все множество вселенных разбить по степени их сложности, то получим набор

$$K_1, K_2, K_3, \dots, K_U, K_{U+1}, \dots$$

Индекс означает степень сложности вселенной, K_1 — число самых простых вселенных, K_2 — число вселенных, устроенных чуть сложнее и т.д. Число самых простых вселенных, где возможно появление разума, обозначено K_U . K_{U+1} — число более

сложных вселенных, которые могут быть населены разумом. Выше было предположено, что чем сложнее вселенные, тем их доля меньше. Это означает, что

$$K_1 > K_2 > K_3 > \dots > K_U > K_{U+1} > \dots$$

Значит, вероятность разумной цивилизации обнаружить себя в относительно простой вселенной выше (а скорее много выше), чем в более сложной. Термин “простая” в данном случае означает, что свойства данной вселенной определяются наименьшим числом постулатов. А это как раз и означает, что наша вселенная, населенная разумной жизнью, должна описываться красивой теорией.

Предположение о существовании дополнительных измерений оказывается удивительно плодотворным. На его основе находит естественное объяснение идея множественности миров и важнейший факт тонкой настройки параметров Вселенной. Дополнительные измерения служат источником физических полей. Представляется вероятным, что источником всех явлений и теорий, изложенных в книге, служат именно дополнительные измерения. И даже такой необъяснимый факт как красота правильных физических теорий становится если не доказанным, то более понятным. К сожалению, дополнительные измерения могут оказаться весьма трудно наблюдаемыми, например, из-за своих малых размеров, так что их убедительное экспериментальное обнаружение может оказаться делом будущего.

Литература

- [1] Н.М. Бочарова, К.А. Бронников, В.Н. Мельников, Вестн. МГУ, физ., астрон., 1970, е 6, 706.
- [2] В.Б. Брагинский, В.И. Панов, ЖЭТФ **61**, 873 (1971).
- [3] К.А. Бронников, М.А. Ковальчук, в сб. “Проблемы теории гравитации и элементарных частиц” (под ред. К.П. Станюковича), вып. 10, Атомиздат, М., 1979.
К.А. Bronnikov and M.A. Kovalchuk, J. Phys. A: Math. Gen. **13**, 187 (1980).
- [4] Ю.С. Владимиров, Системы отсчета в теории гравитации. Энергоиздат, М., 1982.
- [5] Ю.С. Владимиров, Геометрофизика. БИНОМ, М., 2005.
- [6] Э.Б. Глинер, ДАН СССР **192**, 771 (1970).
- [7] А.А. Гриб, С.Г. Мамаев, В.М. Мостепаненко, Квантовые эффекты в интенсивных внешних полях. Атомиздат, М., 1980; А.А. Grib, S.G. Mamayev and V.M. Mostepanenko, Vacuum Quantum Effects in Strong Fields. Friedmann Laboratory Publ., SPb, 1994.
- [8] А.К. Гуц, Элементы теории времени. Наследие, Омск, 2004.
- [9] В.И. Докучаев, УФН 161, 1 (1991); Phys. Usp. **34**, 447 (1991).
- [10] Б.А. Дубровин, С.П. Новиков, А.Т. Фоменко, Современная геометрия. Методы и приложения. УРСС, Москва, 1988.
- [11] В. Захаров, В. Манаков, В. Новиков, Д. Питаевский, Теория солитонов: метод обратной задачи. Наука, М., 1980.

- [12] А.Л. Зельманов, Хронометрические инварианты. American Research Press, Rehoboth, New Mexico, 2006.
- [13] А.Л. Зельманов, В.Г. Агаков, Элементы общей теории относительности. Наука, М., 1989.
- [14] Д.Д. Иваненко, П.И. Пронин, Г.А. Сарданашвили, Калибровочная теория гравитации. Изд. МГУ, М., 1985.
- [15] С.В. Кетов, Введение в квантовую теорию струн и суперструн. Наука, Новосибирск, 1990.
- [16] С.А. Кононогов, Метрология и фундаментальные физические константы. Стандартиформ, М., 2008.
- [17] С.А. Кононогов, В.Н. Мельников, Измер. техника **6**, 1 (2005).
- [18] Н.П. Коноплева, В.Н. Попов, Калибровочные поля. Атомиздат, М., 1980.
- [19] Р.В. Коноплич, Теор. Мат. Физ. **73**, 1286 (1987).
- [20] Р.В. Коноплич, С.Г. Рубин, ЯФ **37**, 1330 (1983); ЯФ **42**, 1282 (1985).
- [21] Р.В. Коноплич, С.Г. Рубин, ЯФ **44**, 558 (1986).
- [22] Л.Д. Ландау, Е.М. Лифшиц, Квантовая механика. Изд. физ.-мат. лит., М., 1963.
- [23] Л.Д. Ландау, Е.М. Лифшиц, Теория поля. Наука, М., 1973.
- [24] Н.В. Мицкевич, Физические поля в ОТО. Наука, М., 1969.
- [25] Н.В. Мицкевич, А.П. Ефремов, А.И. Нестеров, Динамика полей в общей теории относительности. Энергоатомиздат, М., 1985.
- [26] И.Д. Новиков, В.П. Фролов, Физика черных дыр, Наука, М., 1986. V.P. Frolov and I.D. Novikov, Black Hole Physics: Basic Concepts and New Developments, 1997.
- [27] А. Полнарев, М. Хлопов, УФН **145**, 369 (1985).
- [28] В.Н. Пономарев, А.О. Барвинский, Ю.Н. Обухов, Геометродинамические методы и калибровочный подход к теории гравитационных взаимодействий. Энергоатомиздат, М., 1985.

- [29] П.К. Рашевский, Риманова геометрия и тензорный анализ. Наука, М., 1967.
- [30] В.И. Родичев, Теория тяготения в ортогональном репере. Наука, М., 1974.
- [31] С.Г. Рубин, *ЖЭТФ* **133**, 820 (2008); *Sov. Phys. JETP* **106**, 714 (2008).
- [32] К.П. Станюкович, В.Н. Мельников, Гидродинамика, поля и константы в теории гравитации. Энергоатомиздат, М., 1983.
- [33] И.З. Фишер, *ЖЭТФ* **18**, 636 (1948); gr-qc/9911008.
- [34] Б.Н. Фролов, Пуанкаре-калибровочная теория гравитации. МГПУ, Прометей, М., 2003.
- [35] М.Ю. Хлопов, Основы космомикрoфизики, УРСС, 2004.
- [36] L.F. Abbott, E. Fahri and M. Wise, *Phys. Lett. B* **117**, 29 (1982).
- [37] E. Abdalla, B. Cuadros-Melgar, A.B. Pavan and C. Molina, *Nucl. Phys. B* **752**, 40-59 (2006).
- [38] S.T. Abdyrakhmanov, K.A. Bronnikov and B.E. Meierovich, *Grav. Cosmol.* **11**, 75 (2005).
- [39] F.C. Adams, *Phys. Rev. D* **47**, 426 (1993).
- [40] S. Adler and R.B. Pearson, *Phys. Rev. D* **18**, 2798 (1978).
- [41] I. Affleck and M. Dine, *Nucl. Phys. B* **249**, 361 (1985).
- [42] A. Aguirre, in: *Universe or Multiverse?*, ed. B.J. Carr (Cambridge University Press, 2005).
- [43] P. Aguiar and P. Crawford, *Phys. Rev. D* **62**, 123511 (2000).
- [44] K. Akama, *Lect. Notes Phys.* **176**, 267 (1982).
- [45] K. Akama, *Progr. Theor. Phys.* **78**, 184 (1987).
- [46] I. Antoniadis, *Phys. Lett. B* **246**, 377 (1985).
- [47] O. Arias, R. Cardenas and I. Quiros, *Nucl. Phys. B* **643**, 187 (2002).
- [48] C. Armendáriz-Picon, *Phys. Rev. D* **65**, 104010 (2002).
- [49] R. Arnowitt and J. Dent, *Phys. Rev. D* **71**, 124024 (2005).

- [50] E. Ayon-Beato and A. Garcia, Phys. Rev. Lett. **80**, 5056 (1998).
- [51] T. Banks and M. O’Loughlin, Phys. Rev. D **47**, 540 (1993).
- [52] N. Barbosa-Cendejas and A. Herrera-Aguilar, JHEP **0510**, 101 (2005); Phys. Rev. D **73**, 084022 (2006).
- [53] C. Barceló and M. Visser, Phys. Lett. **466B**, 127 (1999).
- [54] C. Barceló and M. Visser, Class. Quantum Grav. **17**, 3843 (2000).
- [55] J.M. Bardeen, Non-singular general-relativistic gravitational collapse. Proc. Int. Conf. GR5, Tbilisi, 1968, p. 174.
- [56] J.M. Bardeen, B. Carter and S.W. Hawking, Commun. Math. Phys. **31**, 161 (1973).
- [57] J.D. Barrow and S. Cotsakis, Phys. Lett. B **214** 515 (1988).
- [58] J.D. Barrow and A.C. Ottewill, J. Phys. A **10**, 2757 (1983).
- [59] J.D. Barrow and F.J. Tipler, The Anthropic Cosmological Principle (Clarendon Press, Oxford, 1986).
- [60] D. Bazeia, F.A. Brito and J.R. Nascimento, Phys. Rev. D **68**, 085007 (2003).
- [61] R. Bean and J. Magueijo, Phys. Rev. D **66**, 063505 (2002).
- [62] J. Bekenstein, Phys. Rev. D **7**, 949 (1973).
- [63] J.D. Bekenstein, in: Cosmology and Gravitation, ed. M. Novello, Atlantisciences, France, 2000, pp. 1-85.
- [64] C.L. Bennett, Astrophys. J. Lett. **464** L1 (1996).
- [65] W. Berej, J. Matyjasek, D. Tryniecki and M. Woronowicz, Gen. Rel. Grav. **38**, 885 (2006).
- [66] P.G. Bergmann, Int. J. Theor. Phys. **1**, 25 (1968).
- [67] G.D. Birkhoff and R. Langer, Relativity and Modern Physics. Harvard Univ. Press, Harvard, 1923.
- [68] N.D. Birrell and P.C.W. Davies, Quantum Fields in Curved Space. Cambridge UP, Cambridge et al., 1982. Русский перевод: Н. Биррелл, П. Девис, Кватованные поля в искривленном пространстве-времени. Мир, М., 1984.

- [69] U. Bleyer, M. Mohazzab and M. Rainer, gr-qc/9508035.
- [70] A. Bogojevic and D. Stojkovic, Phys. Rev. D **61**, 084011 (2000); gr-qc/9804070.
- [71] N. Bostrom, Anthropic Bias: Observation Selection Effects in Science and Philosophy (Routledge, New York, 2002).
- [72] R.H. Boyer and R.W. Lindquist, J. Math. Phys. **8**, 265 (1967).
- [73] C. Brans and R.H. Dicke, Phys. Rev. **124**, 925 (1961).
- [74] K.A. Bronnikov, Acta Phys. Pol. **B4**, 251 (1973).
- [75] K.A. Bronnikov, Phys. Rev. D **63**, 044005 (2001).
- [76] K.A. Bronnikov, Phys. Rev. D **64**, 064013 (2001).
- [77] K.A. Bronnikov, J. Math. Phys. **43**, 6096 (2002).
- [78] K.A. Bronnikov and M.S. Chernakova, Izv. Vuzov, Fiz. No.9, 46 (2005); Russ. Phys. J. **48**, 940 (2005); gr-qc/0503025.
- [79] K.A. Bronnikov and M.S. Chernakova, Grav. Cosmol. **13**, 1 (49), 51 (2007).
- [80] K.A. Bronnikov, M.S. Chernakova, J.C. Fabris, N. Pinto-Neto and M.E. Rodrigues, "Cold black holes and conformal continuations", gr-qc/0609084.
- [81] K.A. Bronnikov, G. Clément, C.P. Constantindis and J.C. Fabris, Grav. Cosmol. **4**, 128 (1998); Phys. Lett. A **243**, 121 (1998).
- [82] K.A. Bronnikov, C.P. Constantinidis, R.L. Evangelista and J.C. Fabris, Int. J. Mod. Phys. **D 8**, 481 (1999).
- [83] K.A. Bronnikov, H. Dehnen and V.N. Melnikov, Phys. Rev. D **68**, 024025 (2003).
- [84] K.A. Bronnikov, H. Dehnen and V.N. Melnikov, Gen. Rel. Grav. **39**, 973 (2007).
- [85] K.A. Bronnikov, A. Dobosz and I.G. Dymnikova, Class. Quantum Grav. **20**, 3797 (2003).
- [86] K.A. Bronnikov and J.C. Fabris, Phys. Rev. Lett. **96**, 251101 (2006).
- [87] K.A. Bronnikov, S.B. Fadeev, and A.V. Michtchenko, Gen. Rel. Grav. **35**, 505 (2003).

- [88] K.A. Bronnikov and S.V. Grinyok, *Grav. Cosmol.* **10**, 237 (2004).
- [89] K.A. Bronnikov and S.V. Grinyok, *Grav. Cosmol.* **11**, 75 (2005).
- [90] K.A. Bronnikov and Yu.N. Kireyev, *Phys. Lett. A* **67**, 95 (1978).
- [91] K.A. Bronnikov and S.A. Kononogov, *Metrologia* **43**, R1 (2006).
- [92] K.A. Bronnikov, S.A. Kononogov and V.N. Melnikov, *Gen. Rel. Grav.* **38**, 1215 (2006).
- [93] K.A. Bronnikov, S.A. Kononogov, V.N. Melnikov and S.G. Rubin, *Grav. Cosmol.* **14**, 230 (2008).
- [94] K.A. Bronnikov, R.V. Konoplich and S.G. Rubin, *Class. Quantum Grav.* **24**, 1261 (2007).
- [95] K.A. Bronnikov and B.E. Meierovich, *Grav. Cosmol.* **9**, 313 (2003).
- [96] K.A. Bronnikov and V.N. Melnikov, *Gen. Rel. Grav.* **27**, 465 (1995).
- [97] K.A. Bronnikov and V.N. Melnikov, in: *Proceedings of the 18th Course of the School on Cosmology and Gravitation: The Gravitational Constant, Generalized Gravitational Theories and Experiments (30 April-10 May 2003, Erice)*. Ed. G.T. Gillies, V.N. Melnikov and V. de Sabbata, Kluwer, Dordrecht/Boston/London, 2004, pp. 39–64; gr-qc/0310112.
- [98] K.A. Bronnikov and S.G. Rubin, *Phys. Rev. D* **73**, 124019 (2006).
- [99] K.A. Bronnikov and S.G. Rubin, *Grav. Cosmol.* **13**, 191 (2007).
- [100] K.A. Bronnikov and G.N. Shikin, *Grav. Cosmol.* **8**, 107 (2002), gr-qc/0109027.
- [101] K.A. Bronnikov and A.A. Starobinsky, *Pis'ma v ZhETF* **85**, 1, 3-8 (2007); *JETP Lett.* **85**, 1, 1-5 (2007).
- [102] K.A. Bronnikov and Sung-Won Kim, *Phys. Rev. D* **67**, 064027 (2003).
- [103] T. Bunch and P. Davies, *Proc. R. Soc. A* **360**, 117 (1978).
- [104] R.V. Buniy and S.D.H. Hsu, *Phys. Lett. B* **632**, 127 (2006); hep-th/0504003.
- [105] A. Burinskii and S.R. Hildebrandt, *Phys. Rev. D* **65**, 104017 (2002); *Grav. Cosmol.* **9**, 20 (2003); *Czech. J. Phys.* **53** B283 (2003).

- [106] Rong-Gen Cai and N. Ohta, Phys. Rev. D **74**, 064001 (2006); hep-th/0604088.
- [107] P. Callin and F. Ravndall, Phys. Rev. D **70**, 104009 (2004).
- [108] M. Campanelli and C.O. Lousto, Int. J. Mod. Phys. D **2**, 451 (1993).
- [109] B.J. Carr, Astrophys. J. **201**, 1 (1975).
- [110] S.M. Carroll et al., Phys. Rev. D **66**, 024036 (2002); hep-th/0110149.
- [111] S.M. Carroll, M. Hoffman and M. Trodden, Phys. Rev. D **68**, 023509 (2003).
- [112] E. Carugno et. al, Phys. Rev. D **53**, 6863 (1996).
- [113] R. Casadio, A. Fabbri and L. Mazzacurati, Phys. Rev. D **65**, 084040 (2002).
- [114] R. Casadio and L. Mazzacurati, Mod. Phys. Lett. A **18**, 651 (2003).
- [115] C. Castro, A. Granik and M.S. El Naschie, hep-th/0004152.
- [116] K.C.K. Chan, J.H. Horne and R.B. Mann, Nucl. Phys. B **447**, 441 (1995).
- [117] V.M. Chechetkin, M.Yu. Khlopov, M.G. Sapozhnikov and Ya.B. Zeldovich, Phys. Lett. B **118**, 329 (1982).
- [118] A.G. Cohen and D.B. Kaplan, Phys. Lett. B **199**, 251 (1987).
- [119] A.G. Cohen and D.B. Kaplan, Nucl. Phys. B **308**, 913 (1988).
- [120] S. Coleman, Phys. Rev. D **15**, 2929 (1977).
- [121] S. Coleman, Phys. Rev. D **16**, 1762 (1977).
- [122] S. Coleman, Nucl. Phys. B **307**, 867 (1988).
- [123] E.J. Copeland, M. Sami and S. Tsujikawa, Int. J. Mod. Phys. D **15**, 1753 (2006); hep-th/0603057.
- [124] C. Csaki, J. Erlich, T.J. Hollowood and Y. Shirman, Nucl. Phys. B **581**, 309 (2000).
- [125] P. Danies, Mod. Phys. Lett. A **19**, 727 (2004).
- [126] N. Deruelle and M. Sasaki, Progr. Theor. Phys. **110**, 441 (2003).

- [127] B.S. DeWitt, *Dynamical Theory of Groups and Fields*, Gordon and Breach, NY, 1965. Русский перевод: Б.С. ДеВитт, *Динамическая теория групп и полей*. Наука, М., 1987.
- [128] B.S. DeWitt and N. Graham, eds., *The Many-Worlds Interpretation of Quantum Mechanics*. Princeton UP, Princeton, 1973.
- [129] V. Dokuchaev, Yu. Eroshenko and S. Rubin, *Grav. Cosmol.* **11**, 99 (2005).
- [130] A.D. Dolgov and A.D. Linde, *Phys. Lett. B* **116**, 329 (1982).
- [131] A. Dolgov, *Phys. Rep.* **222**, 309 (1992).
- [132] A. Dolgov, K. Freese, R. Rangarajan and M. Srednicki, *Phys. Rev. D* **56**, 6155 (1997).
- [133] A. Dolgov and L. Silk, *Phys. Rev. D* **47**, 4244 (1993).
- [134] A.D. Dolgov and S.H. Hansen, *Nucl. Phys. B* **54**, 408 (1999).
- [135] J.F. Donoghue, *Phys. Rev. D* **50**, 3874 (1994).
- [136] J.P. Druisseu and R. Kerner, *Gen. Rel. and Grav.* 15 797 (1983).
- [137] I. Dymnikova, *Gen. Rel. Grav.* **24**, 235 (1992).
- [138] I.G. Dymnikova, A. Dobożh, M.L. Filchenkov and A. Gromov, *Phys. Lett. B* **506**, 351 (2001).
- [139] I. Dymnikova and M. Fil'chenkov, gr-qc/0209065.
- [140] D.A. Easson and R.H. Brandenberger, *JHEP* **0106**, 024 (2001).
- [141] H. Ellis, *J. Math. Phys.* **14**, 104 (1973).
- [142] J. Ellis, N. Kaloper, K.A. Olive and J. Yokoyama, *Phys. Rev. D* **59**, 103503 (1999).
- [143] H. van Elst, J.E. Lidsey and R. Tavakol, *Class. Quantum Grav.* **11**, 2483 (1994).
- [144] H. Everett, *Rev. Mod. Phys.* **29**, 454 (1957).
- [145] V. Faraoni, *Class. Quantum Grav.* **22**, 3235 (2005).
- [146] R.P. Feynman and A.R. Hibbs, *Quantum Mechanics and Path Integrals*. MacGraw-Hill, NY, 1965. Русский перевод: Р. Фейнман, А. Хибс, *Квантовая механика и интегралы по траекториям*. Мир, М., 1968.

- [147] L. Flamm, Phys. Z. **17**, 48 (1916).
- [148] V.P. Frolov and I.D. Novikov, Phys. Rev. D **42**, 1057 (1990).
- [149] J. Garriga and A. Vilenkin, Phys. Rev. D **64**, 023517 (2001).
- [150] G.W. Gibbons and S.W. Hawking, Phys. Rev. D **15**, 2738 (1977).
- [151] K. Gödel, Rev. Mod. Phys. **21**, 447 (1949).
- [152] J.A. González, F.S. Guzmán and O. Sarbach, Arxiv: 0806.0608.
- [153] M. Greene, J. Schwarz and E. Witten, Superstring Theory. Cambridge UP, Cambridge, 1987. Русский перевод: М. Грин, Дж. Шварц, Э. Виттен, Теория суперструн. Мир, М., 1990.
- [154] Ø. Grøn and S. Hervik, Einstein's General Theory of Relativity, Springer, NY, 2007.
- [155] N. Günther, D. Nicole, and D. Wallace, J. Phys. A **13**, 1755 (1980).
- [156] U. Günther and A. Zhuk, Remarks on dimensional reduction in multidimensional cosmological models, gr-qc/0401003.
- [157] U. Günther, P. Moniz and A. Zhuk, Astrophys. Space Sci. **283**, 679 (2003).
- [158] A.H. Guth, Phys. Rev. D **23**, 347 (1981).
- [159] J.C. Hafele and R.E. Keating, Science **177**, 166 and 168 (1972).
- [160] M. Hamermesh, Group theory and its application to physical problems. Addison-Wesley, Reading, Massachusetts — Palo Alto — London, 1964. Русский перевод: М. Хамермеш, Теория групп. Мир, Москва, 1966.
- [161] J.B. Hartle and S.W. Hawking, Phys. Rev. D **28**, 2960 (1983).
- [162] S.W. Hawking, Nature **248**, 30 (1974); Commun. Math. Phys. **43**, 199 (1975).
- [163] S.W. Hawking and G.F.R. Ellis, The Large Scale Structure of Space-Time, Cambridge UP, Cambridge, 1973. Русский перевод: С. Хокинг, Дж. Эллис, Крупномасштабная структура пространства-времени. Мир, М., 1977.

- [164] S.W. Hawking and W. Israel, ed. *General Relativity, Einstein centenary survey*. Cambridge University Press, 2003. Русский перевод: С. Хокинг, В. Израэль (ред.), *Общая теория относительности*. Мир, М., 1983.
- [165] S.A. Hayward, *Phys. Rev. Lett.* **96**, 031103 (2006).
- [166] A. Hewish et al., *Nature* **217**, No. 5139, 709 (1968).
- [167] D. Hochberg, A.A. Popov and S.V. Sushkov, *Phys. Rev. Lett.* **78**, 2050 (1997).
- [168] D. Hochberg and M. Visser, *Phys. Rev. D* **56**, 4745 (1997).
- [169] D. Hochberg and M. Visser, *Phys. Rev. D* **58**, 044021 (1998).
- [170] C.J. Hogan, in: *Universe or Multiverse?*, ed. B.J. Carr, Cambridge University Press, 2005.
- [171] R. Holman, E.W. Kolb, S.I. Vadas and Y. Wang, *Phys. Rev. D* **43**, 995 (1991).
- [172] P. Hořava and E. Witten, *Nucl. Phys. B* **460**, 506 (1996); *ibid.* **475**, 94 (1996).
- [173] S.D.H. Hsu, *Phys. Lett. B* **644**, 67-71 (2007); hep-th/0608175.
- [174] X.M. Hu and Z.C. Wu, *Phys. Lett. B* **149**, 87 (1984).
- [175] C. Itzykson and J.B. Zuber, *Quantum Field Theory*. McGraw-Hill, NY, 1984.
- [176] P. Jordan, *Schwerekraft und Weltall*. Vieweg, Braunschweig, 1955.
- [177] M.Yu. Khlopov, R.V. Konoplich, R. Mignani, S.G. Rubin, and A.S. Sakharov, *Astroparticle Physics* **12** 367 (2000).
- [178] S. Kachru, R. Kallosh, A. Linde and S. P. Trivedi, *Phys. Rev. D* **68**, 046005 (2003).
- [179] N. Kaloper et. al, *Phys. Rev. Lett.* **85**, 928 (2000).
- [180] T. Kaluza, *Sitzungsber. d. Berl. Akad.*, 1921, S. 966. Русский перевод: в кн. “Альберт Эйнштейн и теория гравитации”, Мир, М., 1979, с. 529.
- [181] G.L. Kane, M.J. Perry, and A. N. Zytchow, *New Astron.* **7**, 45 (2002).

- [182] P. Kanti, *Int. J. Mod. Phys. A* **19**, 4899 (2004).
- [183] N.S. Kardashev, I.D. Novikov and A.A. Shatskiy, *Int. J. Mod. Phys. D* **16**, 909 (2007); astro-ph/0610441.
- [184] R. Kerner, *Gen. Rel. Grav.* **14**, 453 (1982).
- [185] R.P. Kerr, *Phys. Rev. Lett.* **11**, 237 (1963).
- [186] S. Khalil and C. Muños, *Contemp. Phys.* **43**, 51 (2002).
- [187] M. Khlopov, S. Rubin, and A. Sakharov, *Phys. Rev. D* **62**, 083505 (2000).
- [188] M.Yu. Khlopov and S.G. Rubin, *Cosmological Pattern of Microphysics in the Inflationary Universe*, vol. 144 (2004). Kluwer Academic Publishers, Dordrecht-Boston-London, series: *Fundamental Theories of Physics*.
- [189] N. Khviengia, Z. Khviengia, H. Lü and C.N. Pope, *Class. Quantum Grav.* **15**, 759 (1998).
- [190] C. Kiefer, in: *Canonical Gravity – from Classical to Quantum*, ed. by J. Ehlers and H. Friedrich, Springer, Berlin, 1994.
- [191] A.A. Kirillov, *Phys. Lett. B* **632**, 453 (2006).
- [192] A.A. Kirillov and E. P. Savelova, *Phys. Lett. B* **660**, 93 (2008).
- [193] A.A. Kirillov and E. P. Savelova, *Grav. Cosmol.* **14**, 256 (2008).
- [194] O. Klein, *Z. Phys.* **37**, 895 (1926); **46**, 188 (1927).
- [195] S. Kobayashi, K. Koyama and J. Soda, *Phys. Rev. D* **65**, 064014 (2002).
- [196] L. Kofman, A. Linde, and A. Starobinsky, *Phys. Rev. Lett.* **73**, 3195 (1994).
- [197] E.W. Kolb, *Physica Scripta* **36**, 199 (1991).
- [198] R.V. Konoplich, *Sov. Nucl. Phys.* **32**, 1132 (1980).
- [199] R.V. Konoplich, M.Yu. Khlopov, S.G. Rubin, and A.S. Sakharov, *Formation of black holes in first order phase transitions*, Preprint 1203, Rome1, Rome, 1998.
- [200] M.Yu. Konstantinov, *Int. J. Mod. Phys. D* **7**, 1 (1998).

- [201] S. Krasnikov, *Phys. Rev. D* **62**, 084028 (2000).
- [202] A.B. Krebs and S.G. Rubin, *Phys. Rev. B* **49**, 11808 (1994).
- [203] H. Kroger, G. Melkonian and S.G. Rubin, *Gen. Rel. Grav.* **36**, 1649 (2004).
- [204] M.D. Kruskal, *Phys. Rev.* **119**, 1743 (1960).
- [205] G. Lazarides, hep-ph/0204294.
- [206] J.P.S. Lemos, F.S.N. Lobo and S.Q. de Oliveira, *Phys. Rev. D* **68**, 064004 (2003).
- [207] W. Lerche, D. Lust and A. N. Schellekens, *Nucl. Phys. B* **287**, 477 (1987).
- [208] A. Lichnerovicz, *Theorie globale des connexions et des groupes d'holonomie*. Edizioni Cremonese, Roma, 1955. Русский перевод: А. Лихнерович, *Теория связностей в целом и группы голономий*, ИЛ, М., 1960.
- [209] A.D. Linde, *Lett. Nuovo Cim.* **39**, 401 (1984).
- [210] A.D. Linde, *The Large-Scale Structure of the Universe*. Harwood Academic Publishers, London, 1990.
- [211] A.D. Linde, *Particle Physics and Inflationary Cosmology*, Harvard Acad. Press, Geneva, 1990.
- [212] A. Linde, *Phys. Lett. B* **259**, 38 (1991).
- [213] A.D. Linde, *Physica Scripta* **36**, 35 (1991).
- [214] A. Linde, *Phys. Rev. D* **49**, 1783 (1994).
- [215] B. Liu, L. McLerran and N. Turok, *Phys. Rev. D* **46**, 2668 (1992).
- [216] M. Livio and M. J. Rees, *Science* **309**, 1022 (2005).
- [217] V. Lukash, *Sov. Phys. JETP* **79**, 1601 (1980).
- [218] D.H. Lyth and E.D. Stewart, *Phys. Rev. D* **54**, 7186 (1996).
- [219] R. Maartens, *Living Rev. Rel.* **7**, 7 (2004); gr-qc/0312059.
- [220] K. Maeda, *Phys. Rev. D* **39**, 3159 (1989).
- [221] K. Maeda, J.A. Stein-Schabes and T. Futamase, *Phys. Rev. D* **39**, 2848 (1989).

- [222] G. Magnano and L.M. Sokolowski, *Phys. Rev. D* **50**, 5039 (1994).
- [223] M. Malquarti, E.J. Copeland, A.R. Liddle and M. Trodden, *Phys. Rev. D* **67**, 123503 (2003).
- [224] A.S. Mazumdar and S.K. Sethi, *Phys. Rev. D* **46**, 5315-5320 (1992).
- [225] A. Melfo, N. Pantoja and A. Skirzewski, *Phys. Rev. D* **67**, 105003 (2003).
- [226] M.B. Mijić, M.S. Morris, and W.M. Suen, *Phys. Rev. D* **39**, 1496 (1989).
- [227] C.W. Misner, K.S. Thorne and J.A. Wheeler, *Gravitation*. Freeman, San Francisco, 1973. Русский перевод: Ч. Мизнер, К. Торн, Дж. Уилер, Гравитация, в 3 томах. Мир, М., 1977.
- [228] M. Morris, K.S. Thorne and U. Yurtsever, *Phys. Rev. Lett.* **61**, 1446 (1988).
- [229] H. Mouri and Y. Taniguchi, *Astroph. J.* **566**, L17 (2002); *astro-ph/0201102*.
- [230] D. Müller, H.V. Fagundes and R. Opher, *Phys. Rev. D* **66**, 083507 (2002).
- [231] S. Nasri, P.J. Silva, G.D. Starkman and M. Trodden, *Phys. Rev. D* **66**, 045029 (2002).
- [232] H.P. Nilles, *Phys. Rep.* **110**, 1 (1984).
- [233] S. Nojiri and S.D. Odintsov, *Phys. Lett. B* **562** 147 (2003).
- [234] S. Nojiri, S.D. Odintsov and S. Ogushi, *Int. J. Mod. Phys. A* **17**, 4809 (2002).
- [235] K. Nordtvedt, *Astroph. J.* **161**, 1059 (1970).
- [236] I. Novikov, A.G. Polnarev, A. Starobinsky and Ya.B. Zel'dovich, *Astron. Astrophys.* **80**, 104 (1979).
- [237] I.D. Novikov, The time machine and self-consistent evolutions in problems with self-interaction. Preprint Nordita-90/38A. Nordisk Inst. f. Teoretisk Fysik, Kobenhavn, 1990.
- [238] U. Nucamendi and M. Salgado, *Phys. Rev. D* **68**, 044026 (2003).
- [239] H. Ochiai and K. Sato, *Prog. Theor. Phys.* **103**, 893 (2000).

- [240] I. Olasagasti and A. Vilenkin, *Phys. Rev. D* **62**, 044014 (2000).
- [241] J.R. Oppenheimer and G. Volkoff, *Phys. Rev.* **55**, 374 (1939).
- [242] J.P. Ostriker and P.J.E. Peebles, *Astroph. J.* **186**, 467 (1973).
- [243] M. Pavšič, *Phys. Lett. A* **116**, 1 (1986).
- [244] R. Penney, *Phys. Rev.* **182**, 1383 (1969).
- [245] R. Penrose, *Rev. del Nuovo Cim.* **1**, 252 (1969).
- [246] A.M. Perelomov, *Ann. Inst. Henri Poincaré* **24**, 161 (1976).
- [247] J. Polchinski, *String Theory*, Cambridge UP, Cambridge, 1998.
- [248] R.V. Pound and G.A. Rebka, Jr., *Phys. Rev. Lett.* **4**, 337 (1960).
- [249] R.V. Pound and J.L. Snider, *Phys. Rev.* **140**, B788 (1960).
- [250] F. Quevedo, *Class. Quantum Grav.* **19**, 5721 (2002).
- [251] R. Rajaraman, *Solitons and Instantons. An Introduction to Solitons And Instantons in Quantum Field Theory*. North-Holland, Amsterdam — New-York — Oxford, 1982.
- [252] L. Randall and R. Sundrum. *Phys. Rev. Lett.* **83**, 3370 (1999).
- [253] L. Randall and R. Sundrum, *Phys. Rev. Lett.* **83**, 3690 (1999).
- [254] M.J. Rees. *Our Cosmic Habitat*. Princeton Univ. Press, Princeton, 2002.
- [255] M. Rees, R. Ruffini and J.A. Wheeler. *Black Holes, Gravitational Waves and Cosmology: an Introduction to Current Problems*. Gordon and Breach, NY — London — Paris, 1974. Русский перевод: М. Рис, Р. Руффини, Дж. Уилер. Черные дыры, гравитационные волны и космология. Мир, М.. 1977.
- [256] S.-J. Rey, *Nucl. Phys. B* **284**, 706 (1987).
- [257] A.G. Riess, *Astron. J.* **116**, 1009 (1998).
- [258] W. Rindler, *Essential relativity*. Springer, NY, 1979.
- [259] T.G. Rizzo, *JHEP* **0506**, 079 (2005); hep-ph/0503163.
- [260] E. Roessl and M. Shaposhnikov, *Phys. Rev. D* **66**, 084008 (2002); hep-th/0205320.

- [261] V.A. Rubakov. *Physics-Uspekhi* **171**, 9, 913 (2001).
- [262] V.A. Rubakov and M.E. Shaposhnikov, *Phys. Lett. B* **125**, 136 (1983).
- [263] S. Rubin, *JETP Lett.* **74**, 275 (2001).
- [264] S.G. Rubin, *Chaos Solitons Fractals* **14**, 891 (2002).
- [265] S.G. Rubin, in: I.Ya. Pomeranchuk and *Physics at the Turn of the Century*, (Moscow, 2003), p. 413; astro-ph/0511181.
- [266] S.G. Rubin, *Grav. Cosmol.* **9**, 243 (2003).
- [267] S. Rubin, M. Khlopov and A. Sakharov, *Grav. Cosmol.* **6**, Suppl., 51–58 (2000).
- [268] S.G. Rubin, M.Yu. Khlopov and A.S. Sakharov, *Sov. Phys. JETP* **92**, 921 (2001).
- [269] S. Rubin, A. Sakharov, and M.Yu. Khlopov, *Sov. Phys. JETP* **92**, 921 (2001).
- [270] V. Sahni and A. Starobinsky, *Int. J. Mod. Phys. D* **9**, 373 (2000).
- [271] V. Sahni and A. Starobinsky, *Int. J. Mod. Phys. D* **15**, 2105 (2006); astro-ph/0610026.
- [272] K. Schwarzschild, *Sitzber. Deut. Akad. Wiss. Berlin, Kl. Math.-Phys. Tech.*, 424 (1916).
- [273] A. Sen, *JHEP* **0204**, 048 (2002); **0207**, 065 (2002).
- [274] S. Shankaranarayanan and N. Dadhich, *Int. J. Mod. Phys. D* **13**, 1095 (2004); gr-qc/0306111.
- [275] I.I. Shapiro et al., *Astroph. J.* **234**, L219 (1979).
- [276] T. Shiromizu, K. Maeda and M. Sasaki, *Phys. Rev. D* **62**, 024012 (2000).
- [277] Y. Shtanov, J. Traschen and R. Brandenberger, *Phys. Rev. D* **51**, 5438 (1995); hep-ph/9407247.
- [278] P. Sikivie, *Phys. Rev. Lett.* **48**, 1156 (1982).
- [279] W. de Sitter, *Proc. Kon. Ned. Akad. Wet.* **20**, 229 (1917).

- [280] M.N. Smolyakov and I.P. Volobuev, Linearized gravity, Newtonian limit and light deflection in RS1 model, hep-th/0208025.
- [281] D.N. Spergel et al., ApJS **170**, 377 (2007); astro-ph/0603449.
- [282] A. Starobinsky, JETP Lett. **30**, 682 (1979).
- [283] A. Starobinsky, Phys. Lett. B **91**, 99–102 (1980).
- [284] A.A. Starobinsky, Pis'ma v Astron. Zh. **7**, 67 (1981); Sov. Astron. Lett. **7**, 361 (1981).
- [285] A. Strominger, Phys. Rev. D **24**, 3082 (1981).
- [286] W.-M. Suen and K. Young, Phys. Rev. D **39**, 2201 (1989).
- [287] J.L. Synge, Proc. Roy. Irish Acad. A **53**, 83 (1950).
- [288] J.L. Synge, Relativity: The General Theory. North Holland, Amsterdam, 1960. Русский перевод: Дж.Л. Синг, Общая теория относительности. ИЛ, М., 1963.
- [289] P. Szekeres, Publ. Mat. Debrecen **7**, 285 (1960).
- [290] M. Tegmark et al, Phys. Rev. D **73**, 023505 (2006).
- [291] R.C. Tolman, Proc. Nat. Acad. Sci. U.S.A. **20**, 410 (1934).
- [292] A.A. Tseytlin and C. Vafa, Nucl. Phys. B **372**, 443 (1992).
- [293] J.-Ph. Uzan, Rev. Mod. Phys. **75**, 403 (2003).
- [294] A. Vilenkin, Phys. Lett. B **117**, 25 (1982).
- [295] A. Vilenkin, Phys. Rev. D **37**, 888 (1988).
- [296] A. Vilenkin and L. Ford, Phys. Rev. D **26**, 1231 (1982).
- [297] M. Visser, Lorentzian wormholes. Springer, NY–Berlin–Heidelberg, 1996.
- [298] R. Wagoner, Phys. Rev. D **1**, 3209 (1970).
- [299] R. Wald, General Relativity. Univ. of Chicago Press, Chicago, 1984.
- [300] S. Weinberg, Gravitation and Cosmology: Principles and Applications of GR. Wiley, NY, 1972. Русский перевод: С. Вайнберг, Гравитация и космология. Мир, М., 1975.
- [301] S. Weinberg, Phys. Rev. Lett. **59**, 2607 (1987).

- [302] S. Weinberg, The cosmological constant problem, astro-ph/0005265.
- [303] S. Weinstein, *Class. Quantum Grav.* **23**, 4231 (2006).
- [304] U. Weiss and W. Haeffner, *Phys. Rev. D* **27**, 2916 (1983).
- [305] J.A. Wheeler, *Phys. Rev.* **97**, 511 (1955).
- [306] J.A. Wheeler, in: *Groups and Topology*, eds. B.S. and C.M. DeWitt (Gordon and Breach, NY, 1963).
- [307] J.A. Wheeler, *Am. Sci.* **59** 1 (1968).
- [308] C. Will, *Theory and Experiment in Gravitational Physics*. Cambridge UP, Cambridge et al., 981. Русский перевод: К. Уилл, *Теория и эксперимент в гравитационной физике*. Энергоатомиздат, М., 1985.
- [309] D.L. Wiltshire, in: *Cosmology: the Physics of the Universe*, eds B. Robson, N. Visvanathan and W.S. Woolcock (World Scientific, Singapore, 1996), pp. 473–531.
- [310] B. Whitt, *Phys. Lett. B* 145 176 (1984).
- [311] E. Witten, *Nucl. Phys. B* **471**, 135 (1996).
- [312] P. Worden, *Acta Astronautica* **5**, 27 (1978).
- [313] J. Yokoyama, *Phys. Rev. D* **58**, 083510 (1998).
- [314] Ya.B. Zel'dovich and I.D. Novikov, *Sov. Astron.* **10**, 602 (1967).
- [315] Ya.B. Zel'dovich and M.Yu. Khlopov, *Phys. Lett. B* **79**, 239 (1978).

Кирилл Александрович Бронников
Сергей Георгиевич Рубин

Лекции по гравитации и космологии
Учебное пособие

Редактор Н.Н. Антонова

Верстка книги полностью соответствует
предоставленному авторами оригиналу-макету

Подписано в печать 11.11.08. Формат 60 × 841/16
Печ. л. 28,5 Уч.-изд.л. 28 Изд. N 1/4 Тираж 150 экз.
Заказ N

Московский инженерно-физический институт
(государственный университет)
Москва, 115409, Каширское шоссе, 31

Типография издательства “Тривант”
г. Троицк Московской обл