

В. Н. ЛУКАШ
Е. В. МИХЕЕВА

ФИЗИЧЕСКАЯ КОСМОЛОГИЯ

Введение

Слово «космология» в буквальном переводе означает «наука о Вселенной». Из этого определения следует, что все естественные науки, которые изучают строение окружающего мира, по сути являются разделами космологии. На практике это не совсем так. Космология изучает структуру и эволюцию Вселенной, то есть устройство мира в больших масштабах. Экспериментальной базой для нее служат наблюдения (на телескопах) во многих информационных каналах: не только электромагнитном — от радио- до гамма-диапазона, но также и тех, что дают нам приемники гравитационных волн, нейтрино, космических лучей и пр. Широко применяется численное моделирование. Таким образом, космология — это физическая наука, опирающаяся на астрономические данные.

Устройство и происхождение окружающего мира всегда волновало человека. Осознание своего места в меняющемся мире превратило его из члена стада в исследователя, раздвигающего границы познания. Космология — самая мировоззренческая из всех естественных наук. Долгое время являясь лишь проявлением общих философских представлений о природе, она получила строгое научное обоснование и возможность последовательной физической экстраполяции только в начале XX века — с появлением общей теории относительности (ОТО). К началу XXI века на основе ОТО была создана космологическая стандартная модель (КСМ), объясняющая всю совокупность имеющихся наблюдательных и экспериментальных данных.

Любая физическая наука опирается на эксперимент. Очевидно, все эксперименты по проверке ОТО тестируют и фундамент космологии. Однако, основываясь на ОТО, космология к ней не сводится. Громадный прогресс в этой науке был бы немыслим без освоения передовых наблюдательных технологий, развития космических исследований, создания новых методов обработки данных и численного моделирования. Эти два мощных рычага — дееспособная теория и своевременная и достаточная экспериментальная база — выдвинули космологию в ряд передовых наук и обеспечили прорыв в наших знаниях как о физике

современного макромира, так и о ранней Вселенной, где космология стыкуется с физикой микромира и высоких энергий.

О плодотворном взаимодействии космологической теории и наблюдений можно судить уже по тому, как часто экспериментальные открытия подтверждали известные ранее теоретические результаты и предсказания. Так, исследование уравнений Эйнштейна привело А. А. Фридмана в 1922 году к созданию космологической модели, которую сегодня мы называем фридмановской моделью (в англоязычной литературе за ней закрепилось более пространное имя — модель Фридмана–Робертсона–Уокера); в ней гравитирующая материя распределена однородно и изотропно и динамика ее разлета описывается эволюцией масштабного фактора $a(t)$. Спустя два года после публикации статьи А. А. Фридмана Э. Хаббл, исследуя движение ближайших галактик, получил соотношение $V = Hr$, которое связывает скорость удаления галактики V с расстоянием до нее r . Существование хаббловского потока материи с коэффициентом H , не зависящим от r , является экспериментальным подтверждением фридмановской модели с масштабным фактором $a = \exp(\int H dt)$.

Подобных примеров множество. Достаточно вспомнить «горячую» модель Вселенной, предложенную Г. А. Гамовым, и открытие электромагнитного реликтового излучения (РИ), предсказание «сахаровских осцилляций» и открытие пиков в спектре анизотропии РИ, предсказание квантово-гравитационного эффекта рождения космологических затравочных неоднородностей плотности и открытие крупномасштабной структуры Вселенной и др. Конечно, были и неподтвердившиеся предсказания — в теории всегда есть свободные параметры, которые может уточнить только эксперимент. Здесь следует подчеркнуть решающую роль наблюдений анизотропии РИ и крупномасштабной структуры Вселенной в создании КСМ и определении космологических параметров.

Вплоть до начала 90-х годов XX века наблюдательная космология развивалась в рамках, традиционных для астрономии. Средством исследования долгое время оставались галактики и связанные с ними структуры (активные галактические ядра, скопления и пр.). Выявление статистических свойств распределения галактик во Вселенной — как и насколько однородно они распределены в пространстве, какие иерархические структуры образуют, как эволюционируют во времени и др. — служило единственным источником наблюдательной информации о свойствах Вселенной.

Качественно новая эра в развитии космологии началась с долгожданного открытия анизотропии реликтового излучения. В отличие от дипольной анизотропии, связанной с движением Земли относительно микроволнового фона, космологическая анизотропия содержит бесценную информацию о процессах и свойствах ранней Вселенной, когда еще не было никаких галактик, а существовали лишь их «зародыши» — начальные неоднородности плотности, которые и возмущали реликтовый фон. Открытие космологической анизотропии РИ нарушило информационную монополию галактик. Впервые стало возможным увидеть развитие структуры Вселенной во времени. Подтвердились ожидания теоретиков о космологической эволюции и формировании галактик из малых возмущений хаббловского потока в ходе развития динамического процесса гравитационной неустойчивости. В результате удалось снять проблему вырождения параметров, неизбежную при анализе данных какого-то одного типа, найти состав и уточнить геометрию Вселенной, значительно повысить точность определения космологических параметров и построить КСМ, подвергнув ее всесторонней проверке независимыми наблюдательными данными.

Космология не стоит на месте — она бурно развивается. То, что десять лет назад было предметом ожесточенных споров и дискуссий, сегодня либо стало твердо установленным фактом, либо отброшено как ошибочная гипотеза. Скептику может показаться, что писать книгу о космологии — неблагодарная задача. Завтра мы будем знать больше, чем сегодня. Однако уже сегодня имеются отшлифованные многочисленными тестами бесценные крупницы знаний, которые составляют физическую основу КСМ и не могут измениться в будущем с появлением новой наблюдательной информации. При увеличении точности эксперимента значения известных нам космологических параметров будут лишь уточнены, а модель сможет быть доработана и приобретет дополнительные очертания только в том секторе, который сегодня спрятан в области ошибок наблюдений.

О каких знаниях мы говорим и о чем эта книга? Если совсем коротко — о крупномасштабной структуре Вселенной и истории ее образования — от зарождения хаббловского потока и начальных космологических неоднородностей до завершения процесса генерации нелинейных гравитационно-связанных систем (гало галактик, групп и скоплений) через хаббловский период времени в недалеком по космологическим меркам будущем. Строго фридмановская геометрия математически проста и малоинформативна. Изучение отклонений от нее дает богатую информацию

о природе. Очень интересен вопрос о том, почему и как возник наблюдаемый хаббловский поток материи, отчего он разрушается в процессе космологического расширения. Создание такого потока — это создание Вселенной. А разрушение изначального хаббловского потока мы называем образованием крупномасштабной структуры Вселенной.

Вселенная — это нестационарная гравитирующая материальная система. Оба взаимоположных процесса — восстановление и нарушение хаббловского течения — управляются гравитацией и зависят от физического состояния вещества (соотношения между эффективным давлением и плотностью энергии). В отличие от других типов взаимодействий, которые можно локализовать (сильное, слабое) или экранировать (электромагнитное), гравитационное взаимодействие универсально и присуще любым формам материи. Благодаря своим свойствам гравитация способна при определенных условиях создавать из хаотических начальных состояний материи высокоорганизованные упорядоченные структуры. Примером может служить коллапс сложной газодинамической системы (звезды) с образованием черной дыры, характеризуемой только двумя внешними параметрами — массой и угловым моментом. Другим примером является создание из достаточно общего распределения материи при ее локально-инфляционном расширении высокосимметричной причинно-несвязанной области вселенной с однородным и изотропным распределением вещества и неоднородной кривизной.

Контрапунктом этой книги является гравитация. Сначала мы напомним основные положения теории и выведем уравнение Эйнштейна «на пальцах», переписав динамику Ньютона в лоренц-инвариантном виде. Затем в качестве примеров мы рассмотрим ряд гравитирующих систем, которые актуальны в свете современных и будущих астрофизических исследований и где главная роль принадлежит гравитации. Далее мы подробно обсудим один из таких примеров — Вселенную, акцентируясь на физических основах космологической модели и ее структуре.

Основные обозначения

$\alpha, \beta, \gamma, \dots$	координатные индексы $(0, i)$
i, j, \dots, n	пространственные индексы $(1, 2, 3)$
a, b, c	индексы тетрады $(0, p)$
p, q, r	реперные индексы триады $(1, 2, 3)$
+ - - -	сигнатура метрики
$\delta_{\alpha\beta} = \text{diag}(1, 1, 1, 1)$	единичный тензор
$\eta_{\alpha\beta} = \text{diag}(1, -1, -1, -1)$	тензор Минковского
t, \mathbf{x}	декартовы координаты
$d\mathbf{x} \equiv d^3x$	пространственный объем
$\Delta \equiv \partial^2 / \partial \mathbf{x}^2$	пространственный лапласиан
$g_{\mu\nu}$	метрический тензор
$(.)$	координатная производная $(\eta_{\alpha\beta, \gamma} = 0)$
$(:)$	ковариантная производная $(g_{\alpha\beta; \gamma} = 0)$
e_{α}^a	тетрадные векторы $(g_{\alpha\beta} = e_{\alpha}^a e_{\beta}^b \eta_{ab})$
p^{μ}	волновой вектор света $(p_{\mu} p^{\mu} = 0)$
$\mathbf{e} = e_i = \frac{p_i}{\sqrt{p_1^2 + p_2^2 + p_3^2}}$	направление на небесной сфере
u^{μ}	скорость материи $(u_{\mu} u^{\mu} = 1)$
$p_{\mu\nu} = u_{\mu} u_{\nu} - g_{\mu\nu}$	проекционный тензор
$a_{\mu} = u^{\nu} u_{\mu; \nu}$	ускорение
$T_{\mu\nu}$	тензор энергии-импульса материи
$\varepsilon = \rho c^2 = u^{\mu} u^{\nu} T_{\mu\nu}$	плотность энергии
$p_{\text{tot}} = \frac{1}{3} p^{\mu\nu} T_{\mu\nu}$	полное давление
S, S_i, S_i^j	потенциалы анизотропного давления
$R_{\alpha\beta\gamma\delta}$	приливной тензор Римана

$R_{\mu\nu} = g^{\alpha\beta} R_{\mu\alpha\nu\beta}$	тензор Риччи
$R = R_{\mu}^{\mu}$	скаляр Риччи
$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R$	тензор Эйнштейна
φ	потенциал скорости ($u_{\mu} = \varphi_{,\mu}/w$)
$w^2 = \varphi_{,\mu} \varphi^{,\mu}$	кинетический скаляр
N, N^i	функции хода и сдвига
$g_{\mu\nu}^{(F)} = \text{diag}(1, -a^2, -a^2, -a^2)$	тензор Фридмана
$a = a(t) = (1+z)^{-1}$	фоновый масштабный фактор
$\eta = \int \frac{dt}{a} = \eta_0 - \int_0^z \frac{dz}{H(z)}$	конформное (минковское) время
$(\dot{}) \equiv \partial/\partial t$	производная по фридмановскому времени
$(') \equiv \partial/\partial \eta$	производная по времени η
$\square \equiv \frac{\partial^2}{\partial \eta^2} - \Delta$	световой даламбертиан
$H = \frac{\dot{a}}{a}$	хаббловская функция
$\bar{H} \equiv aH = \dot{a} = \frac{a'}{a}$	хаббловская частота
$k = \mathbf{k} $	пространственная частота
$\mathbf{k} = k_i = \text{const}$	пространственный волновой вектор
$\mathbf{V}_H = \bar{H} \mathbf{x}$	хаббловская скорость материи
$\bar{\gamma} \equiv -\frac{\bar{H}'}{\bar{H}^2} = -\frac{d \ln \bar{H}}{d \ln a}$	производная хаббловской частоты
$\gamma \equiv -\frac{H'}{H^2} = -\frac{d \ln H}{d \ln a}$	производная хаббловской функции
S, V, T	скалярная, векторная и тензорная моды космологических возмущений
$\mathbf{a} = \mathbf{a}(t, \mathbf{x}) = a \cdot (1 - A)$	скалярный масштабный фактор
$q \equiv \ln \left \frac{a(t_c)}{a} \right = A + H v$	S-поле (потенциал кривизны или нормированный потенциал скорости)
$\mathbf{v}_{\text{pec}} = -\frac{\nabla v}{a} \equiv -\frac{\nabla q}{H}$	пекулярная скорость материи
$v \equiv \frac{q}{H}, \quad \mathbf{q} = q - \Phi$	потенциал пекулярной скорости

$\Phi = \frac{H}{a} \int (\gamma q - S) \frac{da}{H}$	гравитационный потенциал
$t_c = t + v$	сопутствующее время
$v = \frac{\delta\varphi}{\dot{\varphi}}$	потенциал 3-скорости, $u_\mu = (1 + D, v, i)$
A, B, C, D	метрические функции S-моды
$v - C + aB' = v,$	
$\dot{v} - D = \frac{\dot{q}}{H}$	соотношения между потенциалами
$v = C = 0$	лагранжево представление
$B = C = 0$	эйлерово представление
$v = B = 0$	«лагранжево-эйлерово» представление
$\delta_\varepsilon \equiv \frac{\delta\varepsilon - \dot{\varepsilon}v}{\varepsilon + p} = \frac{\Delta\Phi}{\gamma\bar{H}^2}$	сопутствующее возмущение плотности
$\delta_p \equiv \frac{\delta p - \dot{p}v}{\varepsilon + p} =$ $= \frac{\delta w - \dot{w}v}{w} = \frac{\dot{q}}{H}$	сопутствующее возмущение давления
$S = S(k) \equiv q_k^2$	S-спектр мощности, $\langle q^2 \rangle = \int S \frac{dk}{k}$
$\langle \dots \rangle$	усреднение по состоянию поля
$n_S = \frac{d \ln S}{d \ln k}$	индекс наклона S-спектра
$\tau_S = \frac{dn_S}{d \ln k}$	индекс изменения наклона
$q_k (\bar{H} \gg k)$	начальный S-спектр
$q_0(k) = q_k (z \ll 10^3)$	послерекombинационный S-спектр
$T(k) = \frac{q_0(k)}{q_k (\bar{H} \gg k)}$	переходная функция S-моды
$P(k) = g^2(1) \cdot k q_0^2(k)$	спектр мощности возмущений
	плотности, $\langle \delta_\varepsilon^2 \rangle = g^2(a) \int q_0^2(k) k^3 dk$
$g(a) = \frac{1}{\gamma\bar{H}^2} \left(1 - \frac{H}{a} \int_0^a \frac{da}{H} \right)$	фактор роста возмущений плотности
$\nu(a) = \frac{1}{a^2} \int_0^a \frac{da}{H}$	фактор роста пекулярной скорости,
	$\langle \mathbf{v}_{pec}^2 \rangle = \nu^2(a) \int q_0^2(k) k dk$
\mathfrak{S}_i	V-поле (вихревая мода)
\mathfrak{S}_{ij}	T-поле (гравитационные волны)
$T = T(k) \equiv q_{k\oplus}^2 + q_{k\otimes}^2$	T-спектр мощности, $\langle \mathfrak{S}_{ij} \mathfrak{S}^{ij} \rangle = \int T \frac{dk}{k}$

$q_{k\lambda}$	Т-спектр в поляризации $\lambda = \oplus, \otimes$
$n_T = \frac{d \ln T}{d \ln k}$	индекс наклона Т-спектра
$r(k) = \frac{T}{S}$	отношение спектров мощности
$H_0 \simeq 70 \text{ км} \cdot \text{с}^{-1} \cdot \text{Мпк}^{-1}$	современная постоянная Хаббла
$\rho_c = \frac{3H_0^2}{8\pi G} \simeq 10^{-29} \text{ г} \cdot \text{см}^{-3}$	критическая плотность Вселенной
$\Omega = \frac{\rho}{\rho_c}$	параметр космологической плотности
$\Omega_r \simeq 9 \cdot 10^{-5}$	плотность радиации
$\Omega_m = \Omega_M + \Omega_b \simeq 0,3$	плотность нерелятивистской материи
$\Omega_M \simeq 0,25$	плотность темной материи
$\Omega_b \simeq 0,05$	плотность барионов
$\Omega_E = 1 - \Omega_m \simeq 0,7$	плотность темной энергии
$M_{10} = \frac{4\pi}{3} \rho_m (10 \text{ Мпк})^3$ $\simeq 1,6 \cdot 10^{14} M_\odot$	масса материи во Вселенной в сфере радиусом 10 Мпк

ГРАВИТАЦИЯ

В этой главе мы расскажем о пространстве-времени и покажем, что законы движения материи и законы гравитации являются геометрическими свойствами пространства-времени. Мы построим релятивистскую теорию гравитации — уравнения Эйнштейна — прямым обобщением нерелятивистской динамики Ньютона, предполагая, что законы движения одинаковы для всех наблюдателей.

1.1. Принцип относительности

Современная физика опирается на *принцип относительности*, который означает независимость физических законов от наблюдателя. Исторически построение физики началось с лабораторной системы отсчета, связанной с Землей. Одним из первых относительность движения понял Г. Галилей. Позднее была построена нерелятивистская ньютоновская динамика, инвариантная относительно преобразований Галилея: они связывают координаты неускоренных наблюдателей, но оставляют абсолютным время.

Когда экспериментально была установлена конечность скорости света c и ее независимость от системы отсчета, возникло понимание того, что постоянная c является *геометрической* константой пространства-времени, входящей в преобразования пространственных координат и времени при переходе от одной инерциальной системы отсчета к другой (преобразования Лоренца)¹⁾ и был сформулирован принцип *специальной относительности*, определяющий свойства пространства-времени в отсутствие гравитации.

Специальная теория относительности оперирует понятием инерциальных (неускоренных) систем отсчета, для которых скорость любого свободно движущегося тела постоянна. Движение с постоянной скоростью возможно в идеализированном случае

¹⁾ В пределе движений с малыми скоростями преобразования Лоренца переходят в преобразования Галилея.

без учета гравитационного взаимодействия, являющегося по своей природе дальнедействующим и универсальным¹⁾). Принцип специальной относительности опирается на весь опыт экспериментальной и теоретической физики и означает, что законы физики одни те же для любых инерциальных систем отсчета.

В присутствии гравитационного поля роль инерциальных наблюдателей играют свободно падающие системы отсчета, и измеренная скорость любой свободной (от негравитационного воздействия) частицы остается постоянной лишь в ограниченной области пространства-времени. А. Эйнштейн одним из первых осознал локальную применимость специальной относительности. Основываясь на двух положениях локального соответствия:

- специальной теории относительности,
- ньютоновской теории гравитации,

он построил релятивистскую теорию гравитации, названную общей теорией относительности.

Релятивистские уравнения гравитации являются прямым обобщением уравнения Пуассона, связывающего гравитационный потенциал²⁾ и плотность материи в ньютоновской механике. Как и уравнение Пуассона, уравнения Эйнштейна линейны по высшим (вторым) производным гравитационного поля и содержат единственную константу ньютоновской гравитации G , постоянство которой постулируется и подлежит экспериментальной проверке. В этом переходе от уравнения Пуассона к уравнениям Эйнштейна мы имеем дело с однозначной и прямолинейной экстраполяцией (без введения дополнительных параметров) известных нам земных законов физики в область сильных и слабых гравитационных полей, где прямых экспериментов нет. Гидом в этой экстраполяции является принцип *общей относительности*, который постулирует инвариантность физических законов относительно любых пространственно-временных преобразований, связывающих всевозможные системы отсчета.

¹⁾ Остальные взаимодействия можно «отключить» от воздействия на тело или частицу, поскольку они либо сверхслабые (например, темная энергия), либо короткодействующие (сильное и слабое взаимодействия), либо могут быть экранированы (электромагнитное взаимодействие). Отсюда словосочетание *специальная относительность*, означающее локальную применимость физики в отсутствие гравитации.

²⁾ Роль потенциалов гравитационного поля в теории относительности играют компоненты метрического тензора. Употребление таких терминов, как «поле» и «потенциал» объясняется консерватизмом языка

Удивительным оказалось то, что такая *минимальная* экстраполяция теории привела к осознанию целого ряда новых и необычных явлений и эффектов, которые никак не могли вместиться в рамки ньютоновской физики. Это и хаббловский поток, и гравитационные волны, и черные дыры, и кротовые норы, и равная способность гравитации к притяжению и отталкиванию, и многое другое. Часть этих *теоретически* предсказанных эффектов уже обнаружена экспериментально, часть еще ожидает своего подтверждения. Все это говорит об эффективности, работоспособности и «правильности» принципа относительности, лежащего в основе физического понимания окружающего нас мира.

Пока неизвестен ни один экспериментальный или наблюдательный факт, свидетельствующий о неприменимости ОТО в известных нам границах. Если мы столкнемся с такими фактами в будущем, то это будет означать либо нарушение в определенных условиях локальной относительности (лоренц-инвариантности), либо ее дальнейшее обобщение (на многомерные, биметрические, скалярно-тензорные и другие теории гравитации), либо модернизацию самой ньютоновской динамики (например, МОНД — модифицированная ньютоновская динамика, введение дополнительных фундаментальных констант в теорию, учет дискретности геометрии и пр.).

Принцип относительности, отделивший физику от систем отсчета, впервые позволил говорить о физике как о геометрической науке. Дальнейшее обобщение принципа относительности связано с гипотезой многомерности. Направление геометризации физики очень актуально и быстро развивается. Мы лишь в начале этого пути. Согласно геометрической гипотезе, реалии окружающего нас мира — четырехмерность, общая и специальная относительность, материя и взаимодействия между элементарными частицами, и т. д. — все это следствия спонтанно нарушенной изначальной неизвестной нам многомерной симметрии (глобальной относительности). Традиционные для физики полевые концепции и теории в четырехмерном пространстве-времени — это лишь артефакты нарушенной прагеометрии. В данной книге мы тоже будем пользоваться исторически привычными и «удобными» понятиями гравитационных «полей», «потенциалов», «взаимодействий» и др., понимая под этим определенные компоненты метрического тензора и, где требуется, оговаривая и уточняя то, что мы имеем в виду.

О том, как произошло нарушение симметрии и возникло гравитационное взаимодействие, мы можем пока только дога-

дываться. Важно, что это *универсальное далекодействующее* взаимодействие присуще всей материи (фермионам) и переносчикам взаимодействий (бозонам, включая собственно гравитон). Эти свойства гравитации определяют ее особую важность для макроскопических систем.

Всего нам известно четыре типа взаимодействий. Каждому из них соответствует своя стабильная (долгоживущая) материальная частица (фермионная система), поддерживаемая квантовыми обменными силами, связанными со статистикой (полуцелый спин), и данным типом взаимодействия. Это:

- электроны, $m_e \simeq 0,5 \text{ МэВ}$ (электромагнитное взаимодействие);
- нейтрино, $m_\nu \sim 0,1 \text{ эВ}$ (слабое взаимодействие);
- протоны, $m_p \simeq 1 \text{ ГэВ}$ (сильное взаимодействие);
- нейтронные звезды, $M_\star \simeq M_p^3/m_p^2 \simeq 1,8 M_\odot$ (гравитационное взаимодействие).

Если для структур микроскопических частиц гравитация малосущественна, то для макроскопических объектов с $m > M_p \simeq 10^{19} \text{ ГэВ}$ ее необходимо учитывать, а для эволюции массивных структур (планет, звезд, галактик и т.д.) гравитация играет первостепенную роль. Столь высокий масштаб гравитационного воздействия объясняется малостью гравитационной постоянной G или, эквивалентно, большим безразмерным энергетическим числом:

$$\mathcal{N}_p \equiv \frac{M_p}{m_p} = \left(\frac{Gm_p^2}{\hbar c} \right)^{-1/2} = \left(\frac{\ell_C}{\ell_G} \right)^{1/2} = 1,3 \cdot 10^{19}, \quad (1.1.1)$$

которое связано с отношением комптоновского и гравитационного радиусов протона $\left(\ell_C = \frac{\hbar}{m_p c}, \ell_G = \frac{Gm_p}{c^2} \right)$ и, по сути, определяет иерархию масштабов и масс нашего мира.

1.2. Как измерить геометрию

Описание геометрии начинается с определения *события* — точки пространства-времени. Измерение геометрии — это измерение пространственно-временных расстояний между событиями (так называемая *метрическая геометрия*). С практической точки зрения геометрию пространства-времени удобно исследовать с помощью пробных частиц, не оказывающих обратного воздействия на пространство-время. Изучая геодезические — мировые линии свободных от негравитационных воздействий пробных

частиц как функции их собственного времени или аффинного параметра, — мы судим о свойствах пространства-времени, в котором они существуют.

В нашей книге мы будем пользоваться для измерения геометрии этим методом траекторий. Зная все геодезические на протяжении всего интервала изменения их аффинных параметров, мы имеем исчерпывающую информацию о *геодезически полном пространстве-времени*, в котором они существуют.

Здесь мы сталкиваемся с двумя важнейшими характеристиками геометрии. Мы должны уметь описывать любые окрестности геодезических, то есть определять пространственно-временной интервал между близкими событиями и представлять его в разных системах отсчета, а также знать границы мировых линий свободных частиц, где они начинаются и где заканчиваются. В первом случае мы говорим о локальных метрических свойствах пространства-времени. Во втором — о глобальных топологических свойствах геометрии и пространственно-временных границах.

1.3. Система единиц

В обыденной жизни люди используют международную систему единиц СИ, в которой за единицы измерения расстояния, времени и массы приняты, соответственно, метр, секунда и килограмм. Все остальные размерные величины (скорость, энергия и т. д.) получаются комбинацией этих трех единиц.

В физике характерные размеры и величины иные. Они определяются четырьмя фундаментальными константами физических взаимодействий или, в пересчете на константы физических теорий, — скоростью света $c \simeq 3 \cdot 10^5$ км·с⁻¹ (лоренцева симметрия), постоянной Планка $\hbar \simeq 10^{-34}$ Дж·с (квантовая теория), гравитационной постоянной $G \simeq 6,67 \cdot 10^{-8}$ дин·см⁻²·г⁻² (ОТО) и большим безразмерным числом $\mathcal{N}_p \simeq 10^{19}$ (или массой протона, см. (1.1.1)). На трех размерных константах основана планковская система единиц (аналог СИ), с их помощью можно получить размерности любых физических переменных.

Приравнивая гравитационный размер частицы массы m

$$\ell_G = \frac{Gm}{c^2}$$

к ее квантовому (комптоновскому) размеру

$$\ell_C = \frac{\hbar}{mc},$$

получаем планковские единицы измерения массы и длины:

$$M_P = \sqrt{\hbar c/G} \simeq 2 \cdot 10^{-5} \text{ г} \simeq 1,2 \cdot 10^{19} \text{ ГэВ}, \quad (1.3.1)$$

$$\ell_P = \sqrt{\hbar G/c^3} \simeq 1,6 \cdot 10^{-33} \text{ см.} \quad (1.3.2)$$

Планковская единица измерения времени получается делением ℓ_P на скорость света:

$$t_P = \sqrt{\hbar G/c^5} \simeq 5 \cdot 10^{-44} \text{ с.} \quad (1.3.3)$$

Классическими представлениями об интервале между событиями, на которых базируется ОТО, можно пользоваться для пространственно-временных промежутков больше планковских. На меньших расстояниях (временах) важны эффекты квантовой теории гравитации и говорить о непрерывном пространстве-времени нельзя.

В этой книге мы будем пользоваться по возможности безразмерными переменными. Чтобы каждый раз не писать размерные константы, по умолчанию будем считать, что скорость света и приведенная постоянная Планка равны единице:

$$c = \hbar = 1, \quad (1.3.4)$$

оставив в явном написании лишь константу G . Условия (1.3.4) позволяют выразить гравитационную постоянную через любую из планковских единиц (1.3.1), (1.3.2) или (1.3.3), что можно использовать при восстановлении истинной размерности физических переменных. Например, скорости частицы v и гравитационному потенциалу Φ соответствуют безразмерные соотношения

$$\frac{v}{c} \quad \text{и} \quad \frac{\Phi}{c^2}, \quad (1.3.5)$$

а размерности плотности ρ или постоянной Хаббла H , выраженных в единицах планковской массы, восстанавливаются следующим образом:

$$\frac{\rho \ell_P^3}{M_P} = \frac{\rho}{M_P^4}, \quad H t_P = \frac{H}{M_P}. \quad (1.3.6)$$

Размерность любой переменной можно получить введением тождественных множителей $1 = t_P M_P$ или $1 = \ell_P M_P$ и др.

1.4. Физика без гравитации

Мир, в котором мы живем, четырехмерен: он имеет три пространственных измерения и одно временное. Это опытный факт. Поэтому каждое событие в нашем пространстве-времени задается четырьмя координатами.

Сделав это утверждение, мы пока еще не определили геометрию, мы ввели только картографирование пространства-времени — «натянули» на него координатную сетку (систему отсчета). Ее можно строить произвольным образом в любой области пространства-времени, и покрытие геометрии не обязано быть полным.

В пренебрежении гравитацией ($G = 0$) наблюдатель, свободный от воздействия каких-либо полей или сил, может пользоваться декартовыми координатами $x^\alpha = (t, \mathbf{x})$ для маркировки событий. Нам нужно описать с позиций этого инерциального наблюдателя, то есть в терминах его карты, траектории движения пробных частиц в пространстве-времени.

Введем аффинный параметр λ , обозначающий метку положения частицы на ее мировой линии¹⁾. Тогда траектория любой частицы в координатах x^α задается четырьмя функциями:

$$x^\alpha = x^\alpha(\lambda). \quad (1.4.1)$$

Определим *собственное время* $\tau(\lambda)$ на мировой линии частицы (1.4.1), отсчитываемое от точки $\lambda = 0$, следующим образом:

$$\tau(\lambda) = \int_0^\lambda \left[\left(\frac{dt}{d\lambda} \right)^2 - \left(\frac{d\mathbf{x}}{d\lambda} \right)^2 \right]^{1/2} d\lambda = \int_0^\lambda \left(\eta_{\alpha\beta} \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda} \right)^{1/2} d\lambda, \quad (1.4.2)$$

где тензор $\eta_{\alpha\beta} = \text{diag}(1, -1, -1, -1)$ называется метрическим тензором Минковского. Он совпадает со своим обратным тензором $\eta^{\alpha\beta} = \eta_{\alpha\beta}$, где $\eta_{\alpha\beta}\eta^{\beta\gamma} = \delta_\alpha^\gamma = \text{diag}(1, 1, 1, 1)$.

Мы предполагаем, что $\tau(\lambda)$ есть время, измеряемое наблюдателем, движущимся вдоль мировой линии $x^\alpha(\lambda)$. Это положение, известное как *хронометрическая гипотеза*, подсказано всем опытом экспериментальной и теоретической физики. В частности, из хронометрической гипотезы следует, что для нашего инерциального наблюдателя, находящегося в точке $\mathbf{x} = 0$, время t является его собственным временем.

Уравнения (1.4.2), справедливые для любых траекторий (1.4.1) в конечной области пространства-времени, опреде-

¹⁾ Аффинный параметр определен с точностью до преобразований $\lambda = \lambda(\lambda')$. В частности, значение $\lambda = 0$ может быть выбрано для произвольной точки на мировой линии. По умолчанию мы предполагаем монотонное изменение λ вдоль траектории движения частицы — одно значение λ соответствует одной точке мировой линии частицы.

ляют плоскую геометрию Минковского в этой области¹⁾. Это и есть геометрия нашего пространства-времени в отсутствие гравитации: квадрат интервала между близкими событиями, разделенными координатами dx^α , равен:

$$ds^2 = \eta_{\alpha\beta} dx^\alpha dx^\beta, \quad (1.4.3)$$

где интервал между точками с $ds^2 > 0$ совпадает с промежутком собственного времени между ними по часам наблюдателя, мировая линия которого пересекает обе точки ($ds = d\tau$).

В зависимости от знака ds^2 вводятся следующие определения:

- времениподобный интервал: $ds^2 > 0$;
- пространственноподобный интервал: $ds^2 < 0$;
- нулевой интервал: $ds = 0$.

Собственное расстояние между точками с $ds^2 < 0$ совпадает с расстоянием между ними по линейке наблюдателя, для которого обе точки (события) одновременны (так называемая *релятивистская одновременность*): $dl = \sqrt{-ds^2}$.

Массивные тела движутся вдоль времениподобных мировых линий, на которых всюду $ds^2 > 0$, и их собственное время можно использовать в качестве аффинного параметра. Траектория свободно движущейся массивной частицы — *времениподобная геодезическая* — определена условием постоянства скорости ее движения (ускорение равно нулю):

$$x^\alpha = x^\alpha(\tau), \quad \frac{d^2 x^\alpha}{d\tau^2} = 0. \quad (1.4.4)$$

Если вместо массивного тела мы рассмотрим свет, то его мировая линия является *нулевой геодезической* (всюду $ds = 0$):

$$\frac{d^2 x^\alpha}{d\lambda^2} = 0, \quad \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda} \eta_{\alpha\beta} = 0, \quad (1.4.5)$$

где λ — аффинный параметр, определяемый из совместного решения уравнений (1.4.5). Как видно из (1.4.2), собственное время фотона равно нулю. Таким образом, если для тела с некоторой (ненулевой) массой покоя его собственное время можно рассматривать как аффинный параметр, то для света это не так.

¹⁾ Также надо не забыть определить границы геометрии Минковского. В реальном мире ($G \neq 0$) о ней можно говорить лишь локально в масштабах много меньше радиусов кривизны. В идеализированном случае вдали от гравитирующих масс область изменения собственного времени зависит от многообразия M . Например, для неускоренных наблюдателей в $M = R^4$: $\tau \in (-\infty, \infty)$.

До сих пор мы описывали траектории частиц в пространстве-времени Минковского с точки зрения одного произвольно взятого инерциального наблюдателя x^α , расположенного в центре своих пространственных координат \mathbf{x} и отсчитывающего время t по собственным часам. Хронометрическая гипотеза позволяет построить подобную декартову систему отсчета $x'^\alpha = (t', \mathbf{x}')$ вокруг любого другого неускоренного наблюдателя $\mathbf{x}' = 0$, движущегося вдоль ненулевой геодезической. Из условий (1.4.4) получаем линейные соотношения между инерциальными системами отсчета:

$$x'^\alpha = \Lambda^\alpha_\beta x^\beta + T^\alpha, \quad (1.4.6)$$

где постоянными коэффициентами являются вектор трансляций $T^\alpha = (t_0, \mathbf{x}_0)$ (четыре трансляции) и матрица поворотов Λ^α_β (преобразования Лоренца).

В четырехмерном пространстве-времени возможны 6 поворотов: 3 чисто пространственных (эйлеровы повороты) и еще 3 в плоскостях (\mathbf{x}, t) (однородные лоренцевы повороты или *лоренцевы бусты*). На матрицу поворотов можно наложить следующие ограничения, не влияющие на траектории наблюдателей:

- $\eta_{\alpha\beta} \Lambda^\alpha_\mu \Lambda^\beta_\nu = \eta_{\mu\nu}$ — условие ортогональности;
- $\Lambda^0_0 \geq 1$ — запрет обращения времени;
- $\det \Lambda^\alpha_\beta = 1$ — запрет зеркального отражения.

С учетом трансляций преобразования (1.4.6) содержат всего десять свободных параметров, а соответствующая группа движений называется *группой Пуанкаре*. Системы отсчета, связанные только трансляциями, измеряют один и тот же промежуток времени и одно расстояние между двумя близкими событиями. При лоренцевых преобразованиях систем отсчета временные и пространственные расстояния между событиями изменяются (инвариантен лишь интервал). *Лоренцева группа* — подгруппа группы Пуанкаре — имеет шесть параметров. Если взять только однородные пространственно-временные повороты, то группы не получится: результат нескольких поворотов будет зависеть от их последовательности. Для реализации группы Лоренца нужно добавить к трем бустам еще три пространственных поворота.

Матрица однородных преобразований Лоренца выглядит следующим образом:

$$\Lambda^0_\alpha = \Lambda^\alpha_0 = (\Gamma, \mathbf{v}\Gamma), \quad \Lambda^i_j = \delta_{ij} + \frac{\Gamma^2}{1 + \Gamma} v_i v_j, \quad (1.4.7)$$

где $\Gamma \equiv (1 - \mathbf{v}^2)^{-1/2}$, $\mathbf{v} = v_i$. Три параметра лоренцевых поворотов в нашей геометрии мы воспринимаем как три компоненты

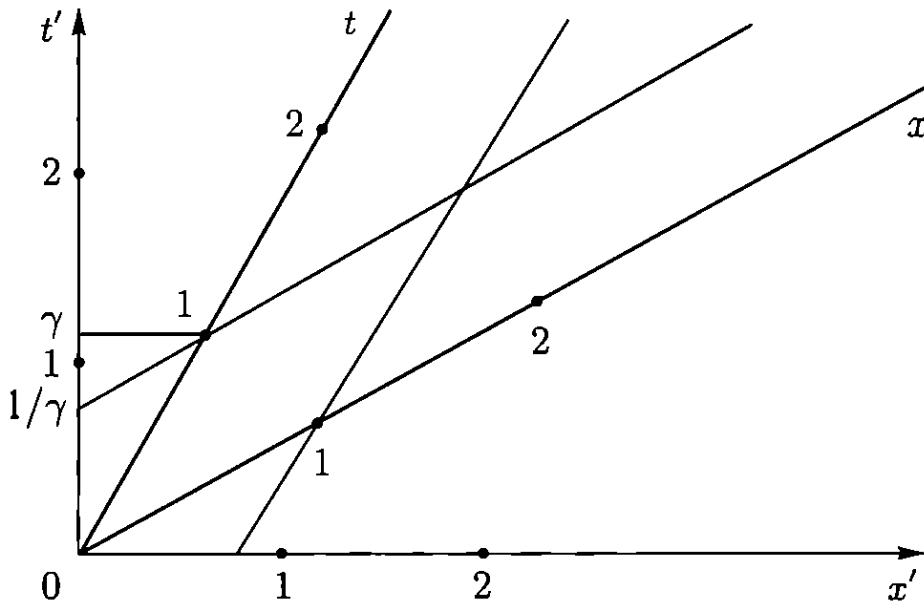


Рис. 1.1. Иллюстрация преобразований Лоренца в плоскости $y = z = 0$, содержащей мировые линии близнецов: $x' = 0$ (черная линия времени t') и $x = 0$ (красная линия времени t); $t' = \gamma t$ — возраст «красного» близнеца t по часам «черного» t' ; $t = \gamma t'$ — возраст «черного» близнеца t' по часам «красного» t

скорости v наблюдателя $x = 0$ (системы отсчета x^α) относительно системы отсчета x'^α наблюдателя $x' = 0$. В пределе малых скоростей $v^2 \ll 1$ (или $c \rightarrow \infty$) получаем преобразования Галилея:

$$t' = t + t_0, \quad x' = x + vt + x_0. \quad (1.4.8)$$

Здесь и в дальнейшем при экстраполяции лабораторной физики в области, не охватываемые нашим опытом, мы сталкиваемся с поучительной интерпретацией движения. То, что мы называем движением, переносом, законами сохранения, оказываются простыми геометрическими свойствами пространства-времени, в котором мы существуем. На этом пути обнаруживаются «парадоксы», не являющиеся парадоксами природы, а возникающие в нашем представлении в силу ограниченного опыта и неадекватности наших моделей отображения мира.

В качестве примера рассмотрим один из таких «парадоксов», известный как парадокс близнецов.

Предположим, что у нас есть два наблюдателя-близнеца, каждый из которых движется со своей скоростью от события рождения. В системе отсчета любого из близнецов движущийся относительно него другой близнец всегда моложе (при приближении к скорости света собственное время останавливается). Получается, что каждый из близнецов по своим собственным часам старше другого. Здесь однако нет никакого парадокса, это свойство нашей геометрии — относительность времени. Мы име-

ем дело с разными синхронизациями одних и тех же событий (собственных возрастов каждого из близнецов) сначала по часам одного, потом по часам другого близнеца (рис. 1.1). Для действительного сравнения им надо снова встретиться: тогда моложе окажется тот, кто испытал большее ускорение.

Преобразования (1.4.6) сохраняют тензор Минковского, собственное время, интервал $ds'^2 = ds^2$ и уравнение геодезических. Равенство

$$\frac{d^2 x'^{\alpha}}{ds'^2} = 0$$

выполняется тогда и только тогда, когда

$$\frac{d^2 x^{\alpha}}{ds^2} = 0.$$

Таким образом, уравнения движения в плоском пространстве-времени инвариантны относительно преобразований группы Пуанкаре (принцип специальной относительности). Это относится ко всем известным законам физики.

Подводя итог, мы можем сказать, что специальная теория относительности, основанная на преобразованиях (1.4.6), есть теория плоского пространства-времени. В идеализированном случае бесконечно продолженного плоского пространства-времени ($\mathcal{M} = R^4$) мы имеем однозначное картографирование событий: уравнения движения выглядят идентично во всех инерциальных сетках.

1.5. Физика с гравитацией

Релятивистская теория гравитации опирается на два основания: принцип относительности и ньютоновский предел.

Общая относительность локально сохраняет принципы специальной относительности. Можно сказать, что общая теория относительности есть теория искривленного пространства-времени, в каждой точке которого можно построить касательное плоское пространство-время. Это значит, что в малой окрестности любого события в сколь угодно сложной геометрии законы физики для свободно падающих тел ¹⁾ такие же, что и в специальной теории относительности. Именно это положение позволит нам найти общие законы динамики пространства-времени,

¹⁾ Такие тела (частицы, наблюдатели, поля), свободные от всех взаимодействий кроме гравитационного, мы для краткости будем называть просто свободными.

исходя из известных нам на опыте локальных законов движения и предполагая, что они идентичны для любых свободных наблюдателей.

В своей точной формулировке принцип относительности утверждает, что законы движения пробных частиц, свободно падающих в постоянном гравитационном поле или движущихся неускоренно вдали от гравитационных масс, тождественны. Иными словами, локальное поле и его градиенты неизмеримы (равны нулю) для наблюдателей, движущихся по геодезическим. Этот принцип локального соответствия, постулирующий неразличимость измерений в свободно падающих и инерциальных системах отсчета, был подсказан Эйнштейну принципом эквивалентности, устанавливающим равенство между гравитационной массой, участвующей в законе всемирного тяготения, и инертной массой, входящей в уравнения движения (второй закон Ньютона). В силу этого соответствия мы будем называть свободно падающие системы отсчета с локально декартовыми координатами вблизи события p локально инерциальными системами отсчета в точке p .

Принцип относительности можно проиллюстрировать на примере нашей земной лаборатории. Местная группа галактик, в которой находится наша Галактика, движется в пространстве относительно реликтового излучения со скоростью около $600 \text{ км} \cdot \text{с}^{-1}$, Солнечная система вращается вокруг галактического центра со скоростью $\sim 200 \text{ км} \cdot \text{с}^{-1}$, а Земля вокруг Солнца — со скоростью $30 \text{ км} \cdot \text{с}^{-1}$. Все это не мешает нам находиться в состоянии покоя на Земле и исследовать физические законы локального мира, применяя координаты Минковского.

Принцип соответствия означает, что в пределе малых скоростей и слабого гравитационного поля релятивистская теория гравитации переходит в ньютоновскую динамику. Мы воспользуемся ньютоновским пределом в следующем разделе для вывода уравнений Эйнштейна из уравнения Пуассона. А сейчас вернемся к принципу относительности.

Теперь мы можем пользоваться декартовыми координатами только локально. Для измерения конечной области искривленного пространства-времени декартова сетка неприменима и нам придется использовать искривленные системы отсчета $x^\alpha = (x^0, x^1, x^2, x^3)$. Заметим, что криволинейными координатами x^α можно покрыть конечные области геометрии, а для других областей могут понадобиться другие системы отсчета.

Согласно принципу соответствия, в окрестности любого события p сколь угодно сложной геометрии можно ввести локально

инерциальные системы отсчета $y^\alpha = y^\alpha(x^\beta)$, в которых траектория $y^\alpha(s)$ любой свободно падающей частицы, проходящей через p , имеет вид (1.4.4):

$$y^\alpha = y^\alpha(x^\beta), \quad \left. \frac{d^2 y^\alpha}{ds^2} \right|_p = 0, \quad (1.5.1)$$

где $s = \tau$ — собственное время свободной частицы (см. (1.4.2)). Тогда интервал пространства-времени вблизи p в координатах y^α выглядит следующим образом:

$$ds^2 = \eta_{\alpha\beta} dy^\alpha dy^\beta. \quad (1.5.2)$$

Переходя обратно к координатам x^α , получаем уравнение геодезической в конечной области пространства-времени, покрываемой системой отсчета x^α , для всех событий на мировой линии частицы:

$$\frac{d^2 x^\alpha}{ds^2} + \Gamma_{\beta\gamma}^\alpha \frac{dx^\beta}{ds} \frac{dx^\gamma}{ds} = 0, \quad (1.5.3)$$

где связь между координатами x^α и y^α спрятана в коэффициенты $\Gamma_{\beta\gamma}^\alpha$, которые называются символами Кристоффеля или *связностью*:

$$\Gamma_{\beta\gamma}^\alpha = \frac{\partial x^\alpha}{\partial y^\delta} \frac{\partial^2 y^\delta}{\partial x^\beta \partial x^\gamma}. \quad (1.5.4)$$

При замене y^α на x^α во всех точках p геометрический интервал приобретает вид:

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu, \quad (1.5.5)$$

где тензор

$$g_{\mu\nu} = \frac{\partial y^\alpha}{\partial x^\mu} \frac{\partial y^\beta}{\partial x^\nu} \eta_{\alpha\beta} \quad (1.5.6)$$

называется *метрическим тензором*. Отметим, что коэффициенты связности $\Gamma_{\beta\gamma}^\alpha$ зависят от системы отсчета и не образуют тензора, тогда как метрический тензор по построению является истинным симметричным тензором — он преобразуется по тензорным правилам при преобразованиях систем отсчета и, следовательно, является характеристикой самого пространства-времени.

Аналогично получаем уравнения для нулевой геодезической $x^\alpha = x^\alpha(\lambda)$:

$$\frac{d^2 x^\alpha}{d\lambda^2} + \Gamma_{\beta\gamma}^\alpha \frac{dx^\beta}{d\lambda} \frac{dx^\gamma}{d\lambda} = 0 \quad (1.5.7)$$

и

$$\frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda} g_{\alpha\beta} = 0, \quad (1.5.8)$$

где λ — аффинный параметр (см. (1.4.5)).

Теперь можно определить и обратный тензор:

$$g^{\mu\nu} = g_{\mu\nu}^{-1} = \frac{\partial x^\mu}{\partial y^\alpha} \frac{\partial x^\nu}{\partial y^\beta} \eta^{\alpha\beta}. \quad (1.5.9)$$

Однако тут следует остановиться и подумать: ведь обратный метрический тензор можно построить и без знания координат y^α , используя лишь условия $g^{\alpha\beta} g_{\beta\gamma} = \delta_\gamma^\alpha$.

Здесь есть тонкость. Все, что было сделано до сих пор, это только преобразования координат. Можно, конечно, использовать преобразования между y^α и x^α и вычислять в каждой точке p компоненты связности и метрического тензора, но это неудобно. Оказывается, и это легко проверить прямым вычислением, можно выразить коэффициенты связности через первые производные метрического тензора $g_{\alpha\beta}(x^\gamma)$ по координатам x^γ :

$$\Gamma_{\beta\gamma}^\alpha = \frac{1}{2} g^{\alpha\delta} (g_{\delta\beta,\gamma} + g_{\delta\gamma,\beta} - g_{\beta\gamma,\delta}), \quad (1.5.10)$$

где $,_\alpha$ означает частную производную по координате x^α . После этого можно забыть о декартовых системах отсчета y^α , поскольку в уравнения геодезических эти координаты непосредственно не входят, и основным инструментом геометрии становится метрический тензор $g_{\mu\nu}$.

Итак, применив принцип соответствия и используя декартовы системы отсчета, мы пришли к выводу, что синонимом геометрии является метрический тензор $g_{\mu\nu}$ ¹⁾. Вместо поиска локально-инерциальных систем отсчета в каждой точке пространства-времени мы возьмем метрический тензор как *определяющий* элемент пространства-времени. Нашим результатом стала метрическая теория гравитации²⁾.

1.6. Лоренцева ковариантность

Там, где нет гравитации, все коэффициенты связности могут быть выбраны равными нулю. В присутствии гравитации всегда можно выбрать такую систему отсчета, в которой $\Gamma_{\beta\gamma}^\alpha = 0$ в любой наперед заданной точке пространства-времени (локально инерциальная система отсчета), однако в других местах возникнет ненулевая связность. Это означает, что в искривленном пространстве-времени нет однозначного картографирования

¹⁾ Напомним, что надо еще дополнить информацию о геометрии определением ее границ (многообразия, топологии).

²⁾ Общая теория относительности — один из примеров метрической теории.

событий, поскольку в разных сетках коэффициенты связности оказываются разными.

Тем не менее, уравнения динамики сохраняют свой вид в любых системах отсчета, если для описания движения пользоваться *геометрическими объектами* или переменными (векторами, тензорами и т. д.), а не координатами карт.

Действительно, связности, как мы видели, возникают только при операциях дифференцирования метрического тензора. Поэтому, для того чтобы связности не появлялись в явном виде, в уравнениях движения вместо простого дифференцирования необходимо использовать операцию *ковариантного дифференцирования*¹⁾, по отношению к которой метрический тензор ведет себя как константа (в том же смысле как тензор Минковского при простом дифференцировании):

$$g_{\mu\nu;\alpha} = 0. \quad (1.6.1)$$

Ковариантное дифференцирование геометрической переменной также приводит к геометрической переменной, то есть эта операция не выводит нас из класса геометрических объектов. Тензор $g_{\mu\nu}$ (и обратный к нему $g^{\mu\nu}$) можно «пронести» через любые операторы ковариантного дифференцирования. Тем самым все манипуляции с четырехмерными греческими индексами (поднятие, опускание, свертка и т. д.) производятся однозначным образом с помощью метрического тензора.

Итак, уравнение геодезической (1.5.3) можно переписать следующим образом:

$$u^\beta u_{\alpha;\beta} = 0, \quad (1.6.2)$$

где $u^\alpha = \frac{dx^\alpha}{d\lambda}$ — вектор, касательный к мировой линии частицы $x^\alpha = x^\alpha(\lambda)$, $u_\alpha = g_{\alpha\beta} u^\beta$ — ковариантный вектор. Очевидно, скалярное произведение $u^\alpha u_\alpha = \text{const}$ является интегралом движения частицы, что можно использовать для нормировки касательного вектора: $u^\alpha u_\alpha = 1$ для массивной частицы (тогда u^α есть скорость частицы и $\lambda = s = \tau$) и $u^\alpha u_\alpha = 0$ для света (в этом случае $u^\alpha = p^\alpha$ — волновой вектор).

Аналогично, законы сохранения

$$T_{\alpha;\beta} = 0 \quad (1.6.3)$$

¹⁾ Обозначается точкой с запятой. Соответственно, сам дифференциал d заменяется на *ковариантный дифференциал* D .

и все другие уравнения движения также сохраняют свой вид при любых преобразованиях координат. Это свойство движения, называемое *лоренцевой ковариантностью*, имеет глубокий смысл: оно связано с независимостью геометрии и законов динамики от наблюдателя (системы отсчета).

Из сравнения с уравнениями предыдущего раздела получаем связь между ковариантной и обычной производными при дифференцировании тензоров:

$$X_{\alpha;\gamma}^{\beta} = X_{\alpha,\gamma}^{\beta} - \Gamma_{\alpha\gamma}^{\delta} X_{\delta}^{\beta} + \Gamma_{\delta\gamma}^{\beta} X_{\alpha}^{\delta},$$

то есть удлинение производной линейно по каждому из индексов. Для удобства дальнейших вычислений перепишем уравнения (1.6.2) и (1.6.3) в терминах обычных производных касательного вектора u_{α} и симметричного тензора $T_{\alpha\beta}$:

$$u^{\beta} u_{\alpha,\beta} = \frac{1}{2} g_{\beta\gamma,\alpha} u^{\beta} u^{\gamma}, \quad (1.6.4)$$

$$\frac{(\sqrt{-g} T_{\alpha}^{\beta})_{,\beta}}{\sqrt{-g}} = \frac{1}{2} g_{\beta\gamma,\alpha} T^{\beta\gamma}, \quad (1.6.5)$$

где $g = \det(g_{\alpha\beta})$ и $g_{,\alpha} = g g^{\beta\gamma} g_{\beta\gamma,\alpha}$.

Лоренцева ковариантность позволяет ввести *тетрадное* представление геометрии, в явном виде демонстрирующее взаимосвязь искривленного пространства-времени с локальной лоренцевой симметрией:

$$g_{\alpha\beta} = e_{\alpha}^a e_{\beta}^b \eta_{ab}, \quad (1.6.6)$$

где тетрада e_{α}^a — это четверка векторов, для обозначения которых мы используем буквы a, b или c , пробегающие значения $0, 1, 2, 3$. Обратные к ним векторы строятся стандартным образом:

$$e_{\alpha}^a e_{\alpha}^b = \delta_a^b, \quad e_{\alpha}^a e_a^{\beta} = \delta_{\alpha}^{\beta},$$

что эквивалентно поднятию и опусканию каждой из групп индексов с помощью соответствующих тензоров $g_{\alpha\beta}$ и η_{ab} .

Дифференциальные формы $e_{\alpha}^a dx^{\alpha}$ не сводятся к полным дифференциалам, однако это всегда можно сделать в малой окрестности любого события (принцип локального соответствия), что и означает переход к локально-декартовым координатам или нулевым связностям (см. (1.5.6)).

Матрица e_{α}^a является, по сути, квадратным корнем метрического тензора. Вместо десяти компонент $g_{\alpha\beta}$ мы теперь можем описывать гравитационное поле с помощью шестнадцати компонент тетрады e_{α}^a . Шесть «лишних» компонент (или их комби-

наций) являются внутренними степенями свободы выбора самой тетрады (в рамках заданной геометрии) по числу параметров локальной группы Лоренца.

Из (1.6.6) следует общее правило для нахождения собственных расстояний между близкими событиями, а также для определения физических компонент геометрических объектов в локально-лоренцевых лабораторных системах отсчета:

$$\delta l^a = e_a^\alpha dx^\alpha, \quad u^a = e_a^\alpha u^\alpha, \quad T_a^b = e_a^\alpha e_\beta^b T_\alpha^\beta$$

и т.д.

1.7. Красное смещение

Как известно, эффектом Доплера называется изменение частоты света или звука из-за движения источника. В специальной теории относительности смещение частоты светового сигнала возможно только при относительном движении источника из-за разности хода времени в *движущихся* системах отсчета (см. (1.4.6)). Однако в присутствии гравитации время течет по-разному в разных местах пространства и в *статическом* случае.

Явление гравитационного красного смещения — прямое следствие того, что теория гравитации является метрической теорией. Рассмотрим движение нерелятивистской частицы в слабом гравитационном поле. В этом случае метрический тензор можно представить в виде:

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}, \quad |h_{\alpha\beta}| \ll 1. \quad (1.7.1)$$

С учетом малости скорости пробной частицы

$$\left| \frac{d\mathbf{x}}{d\tau} \right| \ll \frac{dt}{d\tau} \quad (1.7.2)$$

уравнение геодезической имеет вид:

$$\frac{d^2 x^\alpha}{d\tau^2} + \Gamma_{00}^\alpha \left(\frac{dt}{d\tau} \right)^2 = 0, \quad (1.7.3)$$

где коэффициенты связности определяются выражением:

$$\Gamma_{00}^\alpha = -\frac{1}{2} \eta^{\alpha\beta} h_{00,\beta}. \quad (1.7.4)$$

Перейдем теперь к ньютоновскому пределу, предполагая, что источником гравитационного поля является нерелятивистская материя. Тогда в формуле (1.7.4) мы можем пренебречь произ-

водной по времени по сравнению с пространственными градиентами (квазистатическое приближение), и уравнение геодезической (1.7.3) в ньютоновском пределе приобретает вид:

$$\frac{d^2 t}{d\tau^2} = 0, \quad \frac{d^2 \mathbf{x}}{d\tau^2} = -\frac{1}{2} \nabla h_{00}. \quad (1.7.5)$$

Поскольку нам известно, что ускорение тела в ньютоновской динамике определяется воздействующей на него силой, что в нашем случае есть градиент гравитационного потенциала со знаком «минус»:

$$\frac{d^2 \mathbf{x}}{d\tau^2} = -\nabla \Phi, \quad (1.7.6)$$

мы получаем отсюда следующее соотношение для нулевой компоненты метрического тензора:

$$g_{00} = 1 + 2\Phi. \quad (1.7.7)$$

Таким образом, компоненты метрического тензора имеют смысл гравитационных потенциалов. Разница с ньютоновской гравитацией состоит лишь в том, что в последней имеется только один потенциал Φ , тогда как в релятивистской гравитации их десять¹⁾ по числу независимых компонент метрического тензора (напомним, что $g_{\alpha\beta} = g_{\beta\alpha}$).

Формула (1.7.7) проясняет физический смысл гравитационного красного смещения.

Поскольку собственное время течет в разных областях пространства-времени по-разному,

$$d\tau = (g_{\alpha\beta} dx^\alpha dx^\beta)^{1/2} = g_{00}^{1/2} dt = (1 + \Phi)dt, \quad (1.7.8)$$

то отношение частот принятого (R) и излученного (E) сигналов определяется выражением:

$$\frac{\omega_R}{\omega_E} = \frac{\Delta\tau_E}{\Delta\tau_R} = 1 + \Phi_E - \Phi_R,$$

то есть изменение частоты излученного сигнала составляет величину

$$\frac{\delta\omega}{\omega} = \Phi_E - \Phi_R. \quad (1.7.9)$$

¹⁾ Следует уточнить, что в неинерциальных системах отсчета не все компоненты метрического тензора можно интерпретировать как гравитационное поле: согласно принципу эквивалентности некоторые компоненты или части $g_{\mu\nu}$ могут описывать неинерциальные деформации.

Для Солнца

$$\frac{\delta\omega}{\omega} = \Phi_{\odot} = -\frac{GM_{\odot}}{r_{\odot}} \simeq -2 \cdot 10^{-6}. \quad (1.7.10)$$

Таким образом, вблизи фотосферы Солнца время идет на указанную величину медленнее. Эта величина мала, значит, гравитационное поле Солнца слабое ($|\Phi_{\odot}| \ll c^2$) и релятивистские эффекты гравитации в Солнечной системе малы.

Аналогичным образом можно убедиться в том, что величина гравитационного поля нашей Галактики, а также других галактик и большей части видимой Вселенной также очень мала. Гравитационный потенциал нарастает и приближается к c^2 либо на больших масштабах, сравнимых с размером космологического горизонта, либо на малых расстояниях — вблизи релятивистских объектов (например, черных дыр).

Тем не менее, иногда ньютоновская теория оказывается несостоятельной даже при слабом гравитационном поле и необходимо пользоваться релятивистской теорией гравитации. Это происходит в том случае, если источниками гравитационных полей являются релятивистские формы материи (например, излучение, темная энергия и др.), либо если мы имеем дело с гравитационными волнами.

1.8. От уравнения Пуассона к уравнениям Эйнштейна

В нерелятивистской теории тяготения источником гравитационного поля Φ является плотность материи ρ :

$$\Delta\Phi = 4\pi G\rho, \quad (1.8.1)$$

где G — гравитационная постоянная (константа связи); $\Delta = \delta^{ij} \frac{\partial^2}{\partial x^i \partial x^j}$ — оператор Лапласа в плоском пространстве. Уравнение Пуассона (1.8.1) представлено в декартовых координатах (t, \mathbf{x}) и инвариантно относительно преобразований Галилея (1.4.8).

Обобщим это уравнение для произвольной геометрии.

Для этого вместо ньютоновского потенциала Φ необходимо перейти к метрическому тензору $g_{\alpha\beta}$, содержащему полную информацию о гравитационном поле, а вместо плотности материи ρ использовать хорошо известный в физике тензор энергии-импульса материи $T_{\alpha\beta}$, одной из компонент которого является плотность ρ .

Симметричный тензор $T_{\alpha\beta}$ содержит информацию о динамических законах движения материи, позволяющих описать перенос ее энергии и импульса в пространстве-времени, не вдаваясь в детали внутреннего строения вещества. В геометрии Минковского уравнения движения

$$T_{\alpha,\beta}^{\beta} = 0 \quad (1.8.2)$$

сохраняют свой вид при преобразованиях Лоренца (1.4.6), где $T_{\alpha}^{\beta} = T_{\alpha\gamma}\eta^{\gamma\beta}$. Нулевая компонента (1.8.2) описывает перенос плотности энергии T_0^0 (уравнение непрерывности, или закон сохранения энергии), а пространственные компоненты — перенос плотности импульса T_0^i (уравнение Эйлера, или закон сохранения импульса). Компоненты T_{ij} описывают эффективный тензор давлений среды. Уравнения динамики (1.8.2) справедливы для всех известных нам форм материи. Здесь мы сталкиваемся с еще одним свидетельством того, что законы движения (сохранение энергии-импульса) являются проявлением геометрических свойств пространства-времени.

Перепишем уравнение Пуассона в более удобных для нас обозначениях:

$$\Delta g_{00} = 8\pi G T_{00}. \quad (1.8.3)$$

В ньютоновском пределе $T_{00} = \rho$ и $g_{00} = 1 + 2\Phi$ (см. (1.7.7)), поэтому фактор $4\pi G$ превращается в $8\pi G$ в правой части (1.8.3). Для релятивистского обобщения этого уравнения нам осталось модифицировать оператор Лапласа для локально инерциальных систем отсчета с учетом лоренцевых преобразований.

Не вводя дополнительных гипотез, сохраним главное свойство лапласиана — его линейность по вторым производным координат. Тогда релятивистский аналог уравнения Пуассона представим в виде

$$G_{\alpha\beta} = 8\pi G T_{\alpha\beta}, \quad (1.8.4)$$

где левая часть (1.8.4) в любой локально инерциальной системе отсчета y^{α} есть линейная комбинация всех вторых производных метрического тензора $g_{\alpha\beta}$:

$$G_{\alpha\beta} = A_1 \square g_{\alpha\beta} + A_2 \left(g_{\alpha,\beta\gamma}^{\gamma} + g_{\beta,\alpha\gamma}^{\gamma} \right) + A_3 \eta_{\alpha\beta} g_{\gamma\delta}^{\gamma\delta} + \\ + A_4 g_{\gamma,\alpha\beta}^{\gamma} + A_5 \eta_{\alpha\beta} \square g_{\gamma}^{\gamma}, \quad (1.8.5)$$

где $\square \equiv \eta^{\alpha\beta} \frac{\partial^2}{\partial y^{\alpha} \partial y^{\beta}}$ и $g_{\alpha}^{\beta} = \eta^{\beta\gamma} g_{\alpha\gamma}$. Мы выписали все возможные комбинации вторых производных, и у нас есть пять постоянных

коэффициентов A_1, \dots, A_5 , значения которых следует определить из требований лоренцевой ковариантности и соответствия уравнению (1.8.3) в ньютоновском пределе.

Из закона сохранения (1.8.2) следует

$$G_{\alpha, \beta}^{\beta} = 0, \quad (1.8.6)$$

и, значит,

$$(A_1 + A_2) \square g_{\alpha, \beta}^{\beta} + (A_2 + A_3) g_{, \beta \gamma \alpha}^{\beta \gamma} + (A_4 + A_5) \square g_{\beta, \alpha}^{\beta} = 0. \quad (1.8.7)$$

Поскольку все три комбинации производных метрического тензора несводимы друг к другу, то для выполнения равенства (1.8.7) каждая из входящих в него скобок должна быть равна нулю:

$$A_1 = -A_2 = A_3, \quad A_4 = -A_5.$$

Вместо пяти неизвестных коэффициентов у нас осталось только два, следовательно, тензор $G_{\alpha\beta}$ можно переписать в виде:

$$G_{\alpha\beta} = A_1 \left(\square g_{\alpha\beta} - g_{\alpha, \beta \gamma}^{\gamma} - g_{\beta, \alpha \gamma}^{\gamma} + \eta_{\alpha\beta} g_{, \gamma \delta}^{\gamma \delta} \right) + A_4 \left(g_{\gamma, \alpha\beta}^{\gamma} - \eta_{\alpha\beta} \square g_{\gamma}^{\gamma} \right). \quad (1.8.8)$$

Теперь мы можем использовать ньютоновский предел для определения оставшихся двух коэффициентов. Чтобы избавиться от косых производных, составим из равенств (1.8.8) следующую комбинацию с пространственной сверткой: $G_{00} - G_i^i$. Заметим, что она не сводится к пространственно-временной свертке G_{α}^{α} (напомним, что $G_{\alpha}^{\beta} = G_{\alpha\gamma} \eta^{\gamma\beta}$). Переходя к пределу слабого поля ($|\partial/\partial t| \ll |\partial/\partial x|$ и $\square \simeq -\Delta$), получаем следующее уравнение:

$$\begin{aligned} 8\pi G (T_{00} - T_i^i) &= G_{00} - G_i^i = \\ &= -(A_1 + A_4) \Delta g_{00} + (A_1 - A_4) \Delta g_i^i. \end{aligned} \quad (1.8.9)$$

В ньютоновском пределе источник гравитационного поля нерелятивистский, то есть $|T_i^i| \ll T_{00}$ и вторым членом в левой части (1.8.9) можно пренебречь. Сравнивая уравнения (1.8.3) и (1.8.9), находим значения коэффициентов

$$A_1 = A_4 = -\frac{1}{2},$$

и, следовательно,

$$G_{\alpha\beta} = \frac{1}{2} \left(g_{\alpha, \beta\gamma}^{\gamma} + g_{\beta, \alpha\gamma}^{\gamma} - \square g_{\alpha\beta} - g_{\gamma, \alpha\beta}^{\gamma} \right) - \frac{1}{2} \eta_{\alpha\beta} \left(g_{, \gamma\delta}^{\gamma\delta} - \square g_{\gamma}^{\gamma} \right). \quad (1.8.10)$$

Полученный тензор $G_{\alpha\beta}$ называется тензором Эйнштейна. Он записан в локально-инерциальной системе отсчета, где связности равны нулю. Используя принцип соответствия и переходя к произвольным координатам x^{α} , получим общее выражение для $G_{\alpha\beta}$:

$$G_{\alpha\beta} \equiv R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R, \quad (1.8.11)$$

где тензор

$$R_{\alpha\beta} = \frac{1}{2} g^{\gamma\delta} (g_{\alpha\gamma, \beta\delta} + g_{\beta\gamma, \alpha\delta} - g_{\alpha\beta, \gamma\delta} - g_{\gamma\delta, \alpha\beta}) + g^{\gamma\delta} g_{\mu\nu} \left(\Gamma_{\alpha\gamma}^{\mu} \Gamma_{\beta\delta}^{\nu} - \Gamma_{\alpha\beta}^{\mu} \Gamma_{\gamma\delta}^{\nu} \right) \quad (1.8.12)$$

называется тензором Риччи, а $R = R_{\alpha}^{\alpha}$ — его след (опускание и поднятие индексов производится с помощью тензора $g_{\alpha\beta}$ и обратного к нему тензора $g^{\alpha\beta}$).

Наконец, мы можем выписать уравнения гравитации в знакомом виде:

$$R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R = 8\pi G T_{\alpha\beta}, \quad (1.8.13)$$

а также в других представлениях:

$$R = -8\pi G T, \quad (1.8.14)$$

$$R_{\alpha\beta} = 8\pi G \left(T_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} T \right). \quad (1.8.15)$$

Уравнения Эйнштейна $G_{\alpha\beta} = 8\pi G T_{\alpha\beta}$ имеют интересную структуру: вторые производные по времени входят только в пространственные компоненты G_{ik} . Соответственно пространственные уравнения $G_{ik} = 8\pi G T_{ik}$ являются динамической системой уравнений поля при заданном источнике.

Что же касается уравнений

$$G_{0\alpha} = 8\pi G T_{0\alpha}, \quad (1.8.16)$$

то они, являясь первыми интегралами динамической системы, задают связи для отбора возможных решений уравнений движения. Эти ограничения на гравитационные поля возникают из начальных и граничных условий задачи, определяемых структурой энергии-импульса источника $T_{0\alpha}$.

Уравнения переноса энергии и импульса материи следуют из геометрических тождеств

$$G^{\beta}_{\alpha;\beta} = 0, \quad (1.8.17)$$

называемых *тождествами Бианки*. Мы можем заключить, что именно это свойство геометрии — тождества Бианки — реализует динамические законы движения любых форм материи, существующих в нашем мире.

1.9. Гравитационное действие

Отметим еще одно важное свойство уравнений (1.8.13). Их можно получить варьированием по метрике действия Гильберта (с требованием его неизменности), в котором гравитационная лагранжева плотность пропорциональна скаляру Риччи:

$$S[g^{\alpha\beta}, m] = \int \left(-\frac{R}{16\pi G} + \mathcal{L}_m \right) \sqrt{-g} d^4x, \quad (1.9.1)$$

где «m» означает материальные поля. Универсальность гравитационного взаимодействия обеспечена тем, что константа G содержится в первом (гравитационном) члене. Какие бы виды материи не присутствовали в пространстве-времени, им будет присуще универсальное гравитационное взаимодействие (и самодействие), зависящее от энергии и импульса материи:

$$T_{\alpha\beta} = \frac{2\partial(\sqrt{-g} \mathcal{L}_m)}{\sqrt{-g} \partial g^{\alpha\beta}}. \quad (1.9.2)$$

Для обеспечения аддитивности действия в рамках вариационного принципа необходимо исключить из (1.9.1) члены со вторыми производными по времени.

Это можно сделать с помощью тетрадного представления гравитации. Дело в том, что геометрической переменной первого порядка по координатным производным от метрики не существует (см. (1.6.1)), однако можно построить вектор, составленный из первых производных тетрады ¹⁾:

$$K_{\alpha} = e^a_{\alpha} e^{\beta}_{a;\beta},$$

¹⁾ Напомним, что ковариантные производные по греческим индексам выполняются с помощью полной метрики ($g_{\alpha\beta;\gamma} = 0$), а манипуляции с тетрадными индексами — с помощью тензора Минковского η_{ab} .

и через этот вектор выразить скаляр Риччи (см. (1.6.6)):

$$-R = e_{a\alpha;\beta} e^{a\beta;\alpha} - K_\alpha K^\alpha + 2K^\alpha_{;\alpha}. \quad (1.9.3)$$

Нетрудно заметить, что высшие (вторые) производные содержатся лишь в последнем члене (1.9.3), который является дивергенцией вектора K^α . Этот член можно исключить при определении лагранжевой плотности гравитации:

$$\bar{\mathcal{L}}_g = -\frac{R + 2K^\alpha_{;\alpha}}{16\pi G} = \frac{1}{16\pi G} (e_{a\alpha;\beta} e^{a\beta;\alpha} - K_\alpha K^\alpha), \quad (1.9.4)$$

что эквивалентно добавлению к (1.9.1) поверхностного члена, не влияющего на локальные уравнения движения.

Еще более простой способ избавиться от высших производных состоит в требовании выполнения условия $K^\alpha_{;\alpha} = 0$ (аналог лоренцевой калибровки). Любая из этих процедур — будь то модификация лагранжиана или выбор калибровки — приводит к нарушению лоренцевой инвариантности, поскольку лагранжиан приобретает зависимость от тетрады.

Действительно, выбор конкретного вектора тетрады K^α накладывает ограничение на локальные лоренцевы преобразования, что, вообще говоря, противоречит опыту (напомним, что преобразования Лоренца используются при конструировании ускорителей). Такая «странность» теории гравитации по отношению к лоренцевой инвариантности говорит о непоследовательности самой релятивистской теории и стимулирует дальнейшее развитие наших представлений о природе гравитации.

Мы отсылаем интересующихся этими проблемами читателей к специальной исследовательской литературе, а сейчас вернемся к практическим вопросам построения и применения лагранжевой плотности гравитационного поля.

Выбор тетрады означает выбор определенного расщепления геометрии. Наиболее часто используются пространственно-временные расщепления, для чего необходимо определить в явном виде (обозначить) лишь *один* из четырех векторов тетрады. Мотивация таких расслоений геометрии на *пространство* и *время* связана с сигнатурой нашего мира и с векторным видом уравнений связи (1.8.16), которые являются проекцией тензора G^β_α на времениподобный вектор.

Итак, речь идет о времениподобном тетрадном векторе. Выберем ковектор в простейшем виде:

$$e^0_\alpha \equiv n_\alpha = (N, 0, 0, 0), \quad (1.9.5)$$

а соответствующий ему контравектор — в самой общей форме, сохраняющей нормировку $n_\alpha n^\alpha = 1$:

$$e_0^\alpha \equiv n^\alpha = \left(\frac{1}{N}, -\frac{N^i}{N} \right). \quad (1.9.6)$$

Такой выбор тетрадного вектора называется расслоением Арновитта–Дезера–Мизнера (АДМ). Функции *хода* N и *сдвига* N^i — произвольные функции координат (t, x^i) .

Построим тензор, ортогональный вектору n^α :

$$\gamma_{\alpha\beta} = n_\alpha n_\beta - g_{\alpha\beta}, \quad \gamma^{\alpha\beta} n_\beta = 0. \quad (1.9.7)$$

Из всех контравариантных компонент проекционного тензора ненулевые только пространственные (см. (1.9.5)):

$$\gamma^{\alpha\beta} = \begin{pmatrix} 0 & 0 \\ 0 & \gamma^{ij} \end{pmatrix}. \quad (1.9.8)$$

Считая матрицу γ^{ij} невырожденной и определяя обратную к ней γ_{ij} , получаем интервал в представлении АДМ:

$$ds^2 = N^2 dt^2 - \gamma_{ij} (N^i dt + dx^i) (N^j dt + dx^j), \quad (1.9.9)$$

где γ_{ij} — *пространственный* метрический тензор¹⁾. Для определителя и производной вдоль вектора n^α имеем:

$$g = \det(g_{\mu\nu}) = -N^2 \gamma_n, \quad \partial_n \equiv n^\alpha \frac{\partial}{\partial x^\alpha},$$

где $\gamma_n \equiv \det(\gamma_{ij})$.

С помощью тензора *внешней* кривизны

$$\mathcal{K}_{ij} \equiv N^{-1} \kappa_{ij} = -n_{i;j} = \frac{1}{2} \partial_n \gamma_{ij} - n^\alpha \gamma_{\alpha(i,j)}, \quad (1.9.10)$$

$$\kappa_{ij} = \frac{1}{2} (\dot{\gamma}_{ij} - N_{i|j} - N_{j|i}), \quad (1.9.11)$$

приходим к искомому выражению для скаляра Риччи:

$$-R = R^{(3)} + \mathcal{K}_i^j \mathcal{K}_j^i - \mathcal{K}^2 + 2E^\alpha_{;\alpha}, \quad (1.9.12)$$

¹⁾ С его помощью выполняются манипуляции с пространственными индексами, например $N_i = \gamma_{ij} N^j$, $\varphi^i = \gamma^{ij} \varphi_j$, $\kappa_i^j = \gamma^{jk} \kappa_{ik}$. Внутренней и внешней кривизной пространства $t = \text{const}$ являются соответственно построенный по метрике γ_{ij} 3-скаляр Риччи $R^{(3)}$ и пространственная часть тензора производных $\mathcal{K}_{\alpha\beta} = \gamma_\beta^\gamma n_{\alpha;\gamma}$, ортогонального вектору n^α (см. (1.9.10)).

где

$$E_\alpha = \mathcal{K} n_\alpha - n^\beta n_{\alpha;\beta}, \quad (1.9.13)$$

$$\mathcal{K} \equiv N^{-1} \kappa = n^\alpha_{; \alpha}, \quad (1.9.14)$$

$$\kappa = \frac{(\sqrt{\gamma_n})'}{\sqrt{\gamma_n}} - N^i_{|i}. \quad (1.9.15)$$

Точка над буквой или скобкой и вертикальная прямая в индексе означают соответственно производную по времени t и пространственную ковариантную производную в метрике γ_{ij} .

Опуская член с полной производной (1.9.13), получаем действие (1.9.1) в следующей форме:

$$S = S[n^\mu, \gamma^{ij}, m] = \int (\mathcal{L}_g + N\mathcal{L}_m) \sqrt{\gamma_n} dt d\mathbf{x}, \quad (1.9.16)$$

где

$$\mathcal{L}_g = \frac{N}{16\pi G} \left(\mathcal{K}_i^j \mathcal{K}_j^i - \mathcal{K}^2 + R^{(3)} \right) \quad (1.9.17)$$

— лагранжева плотность гравитационного поля в пространстве γ_{ij} (ср. (1.9.4)). Заметим, что в (1.9.17) не входят производные по времени от вектора n^μ . Поэтому вариация (1.9.16) по n^μ дает в экстремуме не динамические уравнения движения, а уравнения связи ($\partial g^{\alpha\beta} / \partial n^\mu = \delta_\mu^\alpha n^\beta + \delta_\mu^\beta n^\alpha$):

$$n^\nu G_{\mu\nu} = 8\pi G n^\nu T_{\mu\nu}. \quad (1.9.18)$$

Переходя от варьирования по n^α к прямому варьированию действия (1.9.16) по N и N^i и используя формулы перехода

$$N \frac{\partial}{\partial N} = -n^\mu \frac{\partial}{\partial n^\mu}, \quad N \frac{\partial}{\partial N^i} = -\frac{\partial}{\partial n^i},$$

получаем соответственно действие и лагранжиан:

$$S = S[N, N^i, \gamma^{ij}, m],$$

$$\mathcal{L}_g = \frac{1}{16\pi G} \left(\frac{\kappa_i^j \kappa_j^i - \kappa^2}{N} + NR^{(3)} \right), \quad (1.9.19)$$

а также явный вид уравнений связей:

$$\mathcal{K}^2 - \mathcal{K}_i^j \mathcal{K}_j^i + R^{(3)} = 16\pi G n^\mu n^\nu T_{\mu\nu}, \quad (1.9.20)$$

$$\mathcal{K}_{i|j}^j - \mathcal{K}_{,i} = 8\pi G n^\mu T_{\mu i}. \quad (1.9.21)$$

При выполнении связей (1.9.20) и (1.9.21) действие (1.9.16), (1.9.19) оказывается функцией пространственной метрики γ^{ij} и материальных полей m . Условия экстремума действия по этим

переменным дают соответственно динамические уравнения гравитационного и материальных полей.

Лагранжианы (1.9.4), (1.9.17), (1.9.19) и их модификации успешно применяются на практике при исследовании эволюции и структуры гравитационного поля. В случае *выделенной* симметрии, например в астрофизических задачах (компактный источник) или в космологии (изначально нарушена лоренцева симметрия), выбор тетрады или ее части (см., напр., (1.9.5)) физически мотивирован и позволяет непротиворечиво ввести «тензор» энергии-импульса гравитационных волн.

1.10. Ньютоновский предел

Ньютоновский предел означает введение некоторых ограничений на гравитационное поле и его источник: поле предполагается слабым, а его источник — состоящим из нерелятивистской материи (скорости частиц, определяющих тензор энергии-импульса, много меньше скорости света).

Если в слабом гравитационном поле частица движется с малой скоростью, то она остается нерелятивистской и в процессе движения. В этом приближении лоренцев буст сводится к преобразованиям Галилея и связь компоненты метрического тензора g_{00} с ньютоновским гравитационным потенциалом Φ имеет вид (1.7.7).

Однако в ньютоновском поле могут перемещаться и релятивистские частицы (например, свет). В этом случае лоренцевым бустом пренебрегать нельзя, и на движение таких частиц будут влиять и другие компоненты метрического тензора $g_{\alpha i}$. Нам предстоит определить, каким образом они связаны с ньютоновским потенциалом.

Итак, в ньютоновском пределе имеем:

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}, \quad |h_{\alpha\beta}| \ll 1, \quad |p| \ll \rho, \quad |\mathbf{v}| \ll 1,$$

где \mathbf{v} — скорость гравитирующей материи. Пренебрегая давлением p и скоростью \mathbf{v} , получаем единственную ненулевую компоненту тензора энергии-импульса $T_{00} = \rho$. Тогда из уравнений (1.8.15) следует, что

$$R_{\alpha\beta} = 8\pi G \left(T_{\alpha\beta} - \frac{1}{2} \eta_{\alpha\beta} T \right) = 4\pi G \rho \delta_{\alpha\beta}, \quad (1.10.1)$$

где

$$R_{\alpha\beta} = \frac{1}{2} \left(h_{\alpha\gamma,\beta}{}^{,\gamma} + h_{\beta\gamma,\alpha}{}^{,\gamma} - h_{\alpha\beta,\gamma}{}^{,\gamma} - h_{\gamma,\alpha\beta}{}^{,\gamma} \right). \quad (1.10.2)$$

Исходя из вида правой части уравнения (1.10.1), решение для возмущений метрики ищем в виде:

$$h_{\alpha\beta} = 2\Phi \delta_{\alpha\beta}, \quad (1.10.3)$$

и после подстановки в (1.10.2) получаем:

$$R_{\alpha\beta} = \Phi_{,\alpha\gamma} \eta_{\beta}^{\gamma} + \Phi_{,\beta\gamma} \eta_{\alpha}^{\gamma} - \Phi_{,\gamma}^{\gamma} \delta_{\alpha\beta} + 2\Phi_{,\alpha\beta} \simeq \Delta\Phi \delta_{\alpha\beta}$$

(здесь мы пренебрегли временными производными по сравнению с пространственными). Тождественность уравнений (1.10.1) и (1.8.1) подтверждает правильность решения (1.10.3), где Φ — ньютоновский потенциал.

Таким образом, метрический интервал в ньютоновском пределе имеет вид ($|\Phi| \ll 1$):

$$ds^2 = (1 + 2\Phi) dt^2 - (1 - 2\Phi) dx^2, \quad (1.10.4)$$

$$\Delta\Phi = 4\pi G\rho. \quad (1.10.5)$$

Метрика (1.10.4) описывает слабое гравитационное поле на фоне плоского пространства-времени.

Мы живем в другой геометрии, которая на больших масштабах описывается пространственно-плоской метрикой Фридмана:

$$g_{\mu\nu}^{(F)} = \text{diag} (1, -a^2, -a^2, -a^2), \quad (1.10.6)$$

где масштабный фактор $a = a(t) \equiv (1 + z)^{-1}$ есть растущая функция времени, z — красное смещение. Фридмановская система отсчета (t, \mathbf{x}) сопутствует свободно разлетающейся гравитирующей материи, плотность которой постоянна на сечениях $t = \text{const}$ и меняется вдоль линий тока:

$$\rho^{(F)} = \rho(t).$$

Поскольку $g_{\mu\nu}^{(F)}$ не зависит от \mathbf{x} , ковариантная компонента волнового вектора света является интегралом движения: $k = \text{const} = a\omega$ (см. (1.6.4)), откуда следует связь масштабного фактора с красным смещением:

$$\frac{\omega_E}{\omega_R} = a^{-1} \equiv 1 + z.$$

Конечно, слабые гравитационные поля в окрестности любого фридмановского наблюдателя можно рассматривать в его локально инерциальной системе отсчета, но делать это неудобно, поскольку каждый наблюдатель будет иметь свои координаты. Для решения астрофизических задач нам необходимо знать ньютоновский предел во фридмановской системе отсчета.

Рассмотрим слабое неоднородное поле в модели Фридмана, заполненной нерелятивистской материей (также добавим для

общности космологическую постоянную Λ , которая не может быть источником неоднородной части поля):

$$g_{\mu\nu} = g_{\mu\nu}^{(F)} + h_{\mu\nu}, \quad |h_{\mu}^{\nu}| \ll 1, \quad |p| \ll \rho, \quad |\mathbf{v}| \ll 1,$$

где \mathbf{v} — пекулярная скорость, описывающая отклонение скорости материи от хаббловской. Действуя по аналогии с предыдущим выводом, получаем интересующую нас формулу для метрического интервала в ньютоновском пределе:

$$ds^2 = (1 + 2\Phi) dt^2 - a^2 (1 - 2\Phi) dx^2, \quad (1.10.7)$$

где потенциал Φ удовлетворяет модифицированному уравнению Пуассона:

$$\frac{\Delta\Phi}{a^2} = 4\pi G (\rho - \rho^{(F)}), \quad (1.10.8)$$

а функция $a(t)$ определяется из уравнения Фридмана:

$$3H^2 = 8\pi G \rho^{(F)} + \Lambda, \quad (1.10.9)$$

где $\rho^{(F)} = \rho_m/a^3$, $H = \dot{a}/a$, ρ_m и Λ — постоянные, точка над буквой означает производную по времени t .

Условия ньютоновского предела ограничивают массу вещества $M \sim \rho \ell^3$ в рассматриваемой области с размером ℓ :

$$\frac{GM}{\ell} \ll c^2: \quad r_g \ll \ell \ll \frac{c}{\sqrt{G\rho}} \sim \frac{c}{H}, \quad (1.10.10)$$

где r_g — гравитационный радиус массы M . Подчеркнем, что величина плотности в (1.10.5) и (1.10.8) практически не ограничена, то есть пространственное распределение $\rho(t, \mathbf{x})$ может быть нелинейным. При сохранении условия ньютоновского предела $|\Phi| \ll 1$ нелинейные вариации плотности эволюционно *впервые* возникают в масштабах, много меньших хаббловского ¹⁾.

Отдельно следует сказать о системах отсчета (1.10.4) и (1.10.7). Пространственные координаты в геометриях Минковского и Фридмана мы выбирали декартовыми. С учетом релятивистских поправок Φ/c^2 пространственные части метрик (1.10.4) и (1.10.7) приобретают зависимость от \mathbf{x} , и координаты \mathbf{x} перестают быть декартовыми. Однако они остаются *локально*

¹⁾ Скорость света c не входит в уравнение Пуассона, однако при нарушении условий (1.10.10) в уравнениях (1.10.5), (1.10.8) надо учитывать релятивистские поправки. Для простоты записи мы используем одни и те же буквы Φ (и Ψ , см. далее (4.1.12), (5.1.12)) для обозначения *разных* гравитационных потенциалов локально-инерциальных (1.10.4) и фридмановских (1.10.7) наблюдателей. Какой из потенциалов имеется в виду, ясно из контекста, в котором они употребляются.

декартовыми для любых наблюдателей, покоящихся в сетке \mathbf{x} . Условия ортогональности ($g_{0i} = 0$) и локальной декартовости однозначно фиксируют выбор координат (t, \mathbf{x}) по отношению к преобразованиям Лоренца. Такие координаты называются *координатами Эйлера*. Оставшиеся степени свободы в выборе эйлеровых координат (растяжения и пространственные повороты) позволяют определить Φ с точностью до аддитивной постоянной.

1.11. Приливной тензор Римана

Мы исследуем геометрию с помощью геодезических — траекторий свободных частиц. Гравитационное поле — это искривление пространства-времени, которое мы ощущаем и измеряем как приливные силы, действующие на соседние частицы. В силу относительности движения нам надо научиться описывать динамику близких частиц с точки зрения свободного наблюдателя. Это делается с помощью *приливногo тензора Римана*, имеющего прямое отношение к явлению *отклонения геодезических*.

Рассмотрим две близкие геодезические. Одну частицу примем за наблюдателя $x^\alpha(\tau)$, а другую — за объект наблюдения $x^\alpha(\tau) + \delta x^\alpha(\tau)$. Как влияют приливные силы на относительное отклонение свободных частиц $\delta x^\alpha(\tau)$?

Для ответа на этот вопрос нам потребуется понятие *параллельного переноса*. В локально инерциальной системе отсчета наблюдателя на траектории $y^\alpha(\tau)$, проходящей вблизи события p , вектор v^α переносится параллельно вдоль элемента траектории $d\tau$, если

$$\left. \frac{dv^\alpha}{d\tau} \right|_p = 0. \quad (1.11.1)$$

Переходя к общим координатам, получаем уравнение параллельного переноса вектора v^α вдоль траектории $x^\alpha(\tau)$:

$$\frac{Dv^\alpha}{d\tau} \equiv \frac{dv^\alpha}{d\tau} + \Gamma_{\beta\gamma}^\alpha \frac{dx^\beta}{d\tau} v^\gamma = 0. \quad (1.11.2)$$

Тогда для двух близких геодезических $x^\alpha(\tau)$ и $x^\alpha(\tau) + \delta x^\alpha(\tau)$ имеем:

$$\begin{aligned} 0 &= \frac{D}{d\tau} \left[\frac{d}{d\tau} (x^\alpha + \delta x^\alpha) \right] - \frac{D}{d\tau} \left(\frac{dx^\alpha}{d\tau} \right) = \\ &= \frac{d^2 \delta x^\alpha}{d\tau^2} + \Gamma_{\beta\gamma}^\alpha \frac{dx^\beta}{d\tau} \frac{d\delta x^\gamma}{d\tau} + \Gamma_{\beta\gamma, \epsilon}^\alpha \frac{dx^\beta}{d\tau} \frac{dx^\gamma}{d\tau} \delta x^\epsilon, \end{aligned} \quad (1.11.3)$$

где $\delta^\alpha \equiv \delta x^\alpha$. Здесь мы учли зависимость связности от координат траекторий. Собирая члены, составляющие ковариантную производную, получаем относительное ускорение двух частиц:

$$\frac{D^2}{d\tau^2} (\delta x^\mu) = -g^{\mu\nu} R_{\nu\alpha\beta\gamma} \frac{dx^\alpha}{d\tau} \frac{dx^\beta}{d\tau} \delta x^\gamma, \quad (1.11.4)$$

где

$$R_{\alpha\beta\gamma\delta} = \frac{1}{2} (g_{\alpha\delta, \beta\gamma} + g_{\beta\gamma, \alpha\delta} - g_{\alpha\gamma, \beta\delta} - g_{\beta\delta, \alpha\gamma}) + g_{\mu\nu} \left(\Gamma_{\alpha\delta}^\mu \Gamma_{\beta\gamma}^\nu - \Gamma_{\alpha\beta}^\mu \Gamma_{\delta\gamma}^\nu \right) \quad (1.11.5)$$

— тензор Римана. Обратим внимание на то, что ускорение между частицами пропорционально расстоянию между ними — это типичное свойство приливных гравитационных сил, с которым мы часто сталкиваемся (вспомним хаббловские потоки или лунные приливы).

Уравнение (1.11.4) было выведено для вариации скоростей δu^α , каждая из которых является касательным вектором к своей геодезической ($u^\alpha = dx^\alpha/d\tau$). Если параллельно переносится любой другой (не касательный) вектор v^α , то для его изменения мы получаем похожее уравнение:

$$\frac{D}{d\tau} (\delta v^\nu) = -R_{\alpha\beta\gamma}^\nu u^\alpha v^\beta \delta x^\gamma. \quad (1.11.6)$$

В любом случае тензор Римана полностью описывает локальные свойства геометрии при движении частиц.

Приливной тензор получен чисто метрическим образом из уравнений геодезических. Он содержит всю информацию о кривизнах пространства-времени и обладает замечательным свойством — удовлетворяет тождествам Бианки:

$$R_{\mu\alpha\beta;\gamma}^\nu + R_{\mu\beta\gamma;\alpha}^\nu + R_{\mu\gamma\alpha;\beta}^\nu = 0. \quad (1.11.7)$$

Эти тождества дифференциальной геометрии обеспечивают сохранение энергии-импульса материи, находящейся в пространстве-времени, при условии, что $T_{\alpha\beta} = \text{const} \cdot G_{\alpha\beta}$.

По построению тензор Римана (1.11.5) антисимметричен по каждой из пар индексов $\alpha\beta$ и $\gamma\delta$ и симметричен по отношению к перестановкам этих пар. Кроме того, равна нулю сумма компонент при циклической перестановке индексов:

$$R_{\nu\alpha\beta\gamma} + R_{\nu\beta\gamma\alpha} + R_{\nu\gamma\alpha\beta} = 0.$$

Всего, таким образом, у тензора кривизны 20 независимых компонент, причем 10 из них нам известны — это компоненты тензора Риччи, получаемые простой сверткой:

$$R_{\alpha\beta} = g^{\gamma\delta} R_{\alpha\gamma\beta\delta}.$$

Таким образом, тензор Римана не определяется целиком через уравнения Эйнштейна, то есть у него есть компоненты, которые не зависят от источника гравитации.

Это означает, что уравнения Эйнштейна, в которых источник гравитации равен нулю, являются нетривиальными уравнениями. Одно из возможных решений — это пространство Минковского, для которого все компоненты тензоров Римана и Риччи равны нулю. Однако существуют и такие решения, для которых тензор Римана ненулевой, а тензор Риччи нулевой — это свободные гравитационные волны. Соответствующие гравитационно-волновые симметрии можно выделить из тензора Римана, чтобы в остатке не было компонент, сводимых к тензору Риччи:

$$C_{\alpha\beta\gamma\delta} = R_{\alpha\beta\gamma\delta} + \frac{1}{2} (g_{\alpha\delta} R_{\beta\gamma} + g_{\beta\gamma} R_{\alpha\delta} - g_{\alpha\gamma} R_{\beta\delta} - g_{\beta\delta} R_{\alpha\gamma}).$$

Такая конструкция называется тензором Вейля. Если $C_{\alpha\beta\gamma\delta} \neq 0$, то в системе присутствуют гравитационно-волновые степени свободы.

Таким образом, можно условно разделить гравитационные поля на свободные (волны) и связанные с материей («кулоновская» часть). Однако корректно это можно сделать лишь в линейном порядке по возмущениям метрики. В общем виде полей с источниками такое разделение невозможно в силу нелинейности уравнений по первым производным метрического тензора: кулоновская и волновые части взаимодействуют друг с другом. Исключения могут быть сделаны для волновых зон при наличии специальных симметрий.

В следующем разделе мы исследуем свободные волны кривизны.

1.12. Гравитационные волны

Рассмотрим предел слабого поля в метрике Минковского:

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}, \quad |h_{\alpha\beta}| \ll 1. \quad (1.12.1)$$

Тензор Риччи в этом приближении легко получить, убрав все связности как члены второго порядка малости (1.8.12), (1.10.2):

$$R_{\alpha\beta} = -\frac{1}{2} \left(\square h_{\alpha\beta} + h_{,\alpha\beta} - h_{\alpha,\gamma}^{\gamma} - h_{\beta,\gamma}^{\gamma} \right), \quad (1.12.2)$$

где $h \equiv h_{\gamma}^{\gamma}$. Тогда

$$R_{\alpha\beta} = 8\pi G T_{\alpha\beta}, \quad T_{\alpha\beta} \equiv T_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} T. \quad (1.12.3)$$

Нас интересуют решения этого уравнения, причем уже любого типа, а не только скалярного, как было ранее (см. раздел 1.10). В частности, это могут быть решения тензорного типа, из которых нельзя составить скаляр.

Прежде всего, уравнения (1.12.3) можно привести к стандартному волновому виду. Для этого необходимо вспомнить, что у нас имеется геометрия и координатная сетка, которой эту геометрию можно покрыть. Систему координат мы можем выбирать произвольно, какую хотим, геометрия от этого не меняется. Например, если от одной системы координат перейти к другой, которая мало отличается от первой, то требование малости полей сохранится.

Итак, рассмотрим преобразование координат

$$\tilde{x}^{\alpha} = x^{\alpha} + \xi^{\alpha}, \quad (1.12.4)$$

где ξ^{α} — четыре произвольные, но малые функции. При переходе от одних координат к другим сохраняется интервал ($g_{\alpha\beta} dx^{\alpha} dx^{\beta} = \text{inv}$), а возмущенная часть метрического тензора как добавка к тензору Минковского в новых координатах изменяется:

$$\tilde{h}_{\alpha\beta} = h_{\alpha\beta} - \xi_{\alpha,\beta} - \xi_{\beta,\alpha}, \quad (1.12.5)$$

где $\xi_{\alpha} = \eta_{\alpha\beta} \xi^{\beta}$. Эти преобразования называются *калибровочными*. Поясним их смысл.

Геометрия, которую мы рассматриваем, представлена как пространство-время Минковского плюс малое возмущение. Тензор Минковского — это единица и минус единицы по диагонали. При этом молчаливо предполагалось, что имеются уже координаты времени и пространства: только тогда в данных координатах можно написать тензор Минковского. Перейдем к другим координатам. Получим другой тензор Минковского, тоже диагональный, но уже в других координатах. Это означает, что расщепление геометрии на два слагаемых — фон и возмущение — неоднозначно, надо также указать в какой системе отсчета оно сделано. Можно сказать и по-другому: малые члены $h_{\alpha\beta}$

определены с точностью до преобразований (1.12.5). Бесспорно, геометрические степени свободы не зависят от системы отсчета и должны оставаться инвариантными по отношению к калибровочным преобразованиям.

Выберем такой вектор ξ^α , чтобы выполнялось следующее уравнение (так называемая *гармоническая калибровка*):

$$\square \xi_\alpha = h_{\alpha, \beta}^\beta - \frac{1}{2} h_{, \alpha}. \quad (1.12.6)$$

Тогда уравнения (1.12.3) приобретают вид:

$$\square \tilde{h}_{\alpha\beta} = -16\pi G T_{\alpha\beta}. \quad (1.12.7)$$

Здесь все уже привычно: слева строит даламбертиан поля, а справа — источник. Мы получили волновое уравнение с источником, который порождает гравитационные волны.

Пока мы только упражнялись в математике: в приближении слабого гравитационного поля написано уравнение Эйнштейна. Можно забыть, как мы его получили, и считать (1.12.7) оригинальным уравнением Эйнштейна в *лоренцевой калибровке*:

$$\tilde{h}_{\alpha, \beta}^\beta = \frac{1}{2} \tilde{h}_{, \alpha}. \quad (1.12.8)$$

Пусть теперь источника нет, то есть $T_{\alpha\beta} = 0$. Решением волнового уравнения являются свободные волны кривизны. Покажем, что у полей $\tilde{h}_{\alpha\beta}$ есть только две независимых компоненты.

Действительно, у нас еще осталась свобода выбора вектора ξ_α в (1.12.5), (1.12.6), удовлетворяющего волновому уравнению:

$$\square \xi_\alpha = 0. \quad (1.12.9)$$

Это уравнение совпадает с (1.12.7) (при $T_{\alpha\beta} = 0$), поэтому мы всегда можем исключить из $\tilde{h}_{\alpha\beta}$ четыре компоненты. Определим их четырьмя условиями:

$$\tilde{h} = 0, \quad u^\alpha \tilde{h}_{\alpha\beta} = 0, \quad (1.12.10)$$

где $u^\alpha = \text{const}$ — вектор скорости одного из инерциальных наблюдателей. Заметим, что второе уравнение (1.12.10) содержит лишь три независимых условия, поскольку $u^\alpha \tilde{h}_{\alpha, \beta}^\beta = 0$ в силу уже выбранной лоренцевой калибровки. Таким образом, мы имеем четыре условия (1.12.8) и четыре условия (1.12.10): из десяти компонент $\tilde{h}_{\alpha\beta}$ независимыми остаются только две. В дальнейшем мы будем придерживаться именно этой калибровки гравитационных волн, она называется *поперечно-бесследовой*.

Элементарное решение (1.12.7) в виде плоской волны имеет вид:

$$\mathfrak{S}_{\alpha\beta} \equiv \frac{1}{2} \tilde{h}_{\alpha\beta} = \mathcal{P}_{\alpha\beta} e^{-ik_\gamma x^\gamma} + \mathcal{P}_{\alpha\beta}^* e^{ik_\gamma x^\gamma}, \quad (1.12.11)$$

где $k_\alpha = (k, -\mathbf{k}) = \text{const}$ — волновой вектор ($k_\alpha k^\alpha = 0$), $\mathcal{P}_{\alpha\beta}(\mathbf{k}) = \text{const}$ — тензор поляризаций, звездочка означает комплексное сопряжение. В силу поперечно-бесследовой калибровки

$$k^\beta \mathcal{P}_{\alpha\beta} = u^\beta \mathcal{P}_{\alpha\beta} = \eta^{\alpha\beta} \mathcal{P}_{\alpha\beta} = 0 \quad (1.12.12)$$

отличны от нуля только пространственные компоненты:

$$\mathfrak{S}_{0\alpha} = 0, \quad \mathfrak{S}_{ij} \neq 0. \quad (1.12.13)$$

Если ось x^1 совместить с пространственным направлением волнового вектора $k^\alpha = (1, 1, 0, 0)$, то в собственной системе отсчета наблюдателя $u^\alpha = (1, 0, 0, 0)$ и отличны от нуля только следующие компоненты тензора поляризаций:

$$\mathcal{P}_{22} = -\mathcal{P}_{33} = \frac{1}{2} e_\oplus, \quad \mathcal{P}_{23} = \mathcal{P}_{32} = \frac{1}{2} e_\otimes. \quad (1.12.14)$$

Эффективный тензор энергии-импульса плоской волны имеет вид:

$$\langle \tau_{\mu\nu} \rangle = \frac{k_\mu k_\nu}{8\pi G} (|e_\oplus|^2 + |e_\otimes|^2), \quad (1.12.15)$$

где угловые скобки означают усреднение по масштабу больше длины волны. Сам тензор $\tau_{\mu\nu}$ можно получить путем разложения действия Гильберта до второго порядка по возмущениям поля (1.12.11) с последующей вариацией по метрике $\eta^{\mu\nu}$ полученного действия гравитационной волны:

$$L_T = \mathcal{L}_g(\mathfrak{S}_i^j) = \frac{1}{16\pi G} \eta^{\mu\nu} \mathfrak{S}_{i,\mu}^j \mathfrak{S}_{j,\nu}^i, \quad (1.12.16)$$

$$\tau_{\mu\nu} = \frac{1}{8\pi G} \mathfrak{S}_{i,\mu}^j \mathfrak{S}_{j,\nu}^i - L_T \eta_{\mu\nu}. \quad (1.12.17)$$

Сделаем ряд пояснений технического характера. Мы приписали смешанные индексы полю $\mathfrak{S}_{ij} \equiv \mathfrak{S}_i^j$, поскольку используется процедура суммирования по повторяющимся нижним и верхним индексам. Эквивалентно, можно считать, что манипуляции с внутренними полевыми индексами \mathfrak{S}_{ij} выполняются с помощью единичного тензора δ_{ij} и компоненты со смешанными индексами являются физическими компонентами поля гравитаци-

онной волны. При выводе (1.12.16) из действия Гильберта (1.9.1) удобно использовать тетрадное представление метрики (1.12.1):

$$e_{a\alpha} = \eta_{a\alpha} + \mathfrak{S}_{a\alpha}. \quad (1.12.18)$$

Тогда в лоренцевой калибровке $K_\alpha = 0$ и в действии (1.9.1) остается лишь член

$$-R = e_{a\alpha;\beta} e^{a\beta;\alpha}. \quad (1.12.19)$$

Подставляя выражение для ковариантной производной вектора

$$e_{a\alpha;\beta} = \mathfrak{S}_{\alpha\beta,a} - \mathfrak{S}_{a\beta,\alpha}$$

и переходя к пространственным компонентам (1.12.13), получаем лагранжиан плоской волны в форме (1.12.16). Компоненты типа $\mathfrak{S}_{\alpha\beta,a}$, $\mathfrak{S}^{a\beta,\alpha}$ в действие не войдут из-за условий поперечности (1.12.12).

Рассмотрим теперь предел слабого гравитационного поля в плоской модели Фридмана.

В поперечно-бесследовой калибровке имеем:

$$ds^2 = dt^2 - a^2 (\delta_{ij} - 2\mathfrak{S}_{ij}) dx^i dx^j, \quad (1.12.20)$$

$$|\mathfrak{S}_{ij}| \ll 1, \quad \mathfrak{S}_i^i = \mathfrak{S}_{i,j}^j = 0, \quad (1.12.21)$$

где манипуляции с пространственными индексами осуществляются с помощью единичного тензора δ_{ij} . Заметим, что пространственные координаты x^i метрики (1.12.20) являются координатами свободных пылинок. Измеряя переменную часть собственного расстояния $\delta\ell$ между соседними пылинками, находящимися в точках x^i и $x^i + \delta x^i$, мы измеряем компоненты гравитационно-волнового поля:

$$\delta\ell^j = a \mathfrak{S}_i^j \delta x^i. \quad (1.12.22)$$

Действуя аналогично предыдущему случаю, получаем уравнение для распространения волн в метрике Фридмана:

$$\ddot{\mathfrak{S}}_{ij} + 3H \dot{\mathfrak{S}}_{ij} - a^{-2} \Delta \mathfrak{S}_{ij} = 0, \quad (1.12.23)$$

и решение в виде плоской волны:

$$\mathfrak{S}_{ij} = \mathcal{P}_{ij} \nu_k e^{i\mathbf{k}\mathbf{x}} + \mathcal{P}_{ij}^* \nu_k^* e^{-i\mathbf{k}\mathbf{x}}, \quad (1.12.24)$$

где $\mathbf{k} = \text{const}$ — пространственный волновой вектор. Функция времени $\bar{\nu}_k = \bar{\nu}_k(\eta) \equiv a\nu_k$ удовлетворяет уравнению

$$\bar{\nu}_k'' + (k^2 - U) \bar{\nu}_k = 0 \quad (1.12.25)$$

с нормировочным условием

$$\bar{v}_k \bar{v}_k^{*'} - \bar{v}'_k \bar{v}_k^* = 2ik,$$

где штрих означает производную по времени η , а функция

$$U = U(\eta) \equiv \frac{a''}{a} = (1 - \bar{\gamma}) \bar{H}^2$$

описывает эффективный потенциал гравитационных волн в модели Фридмана ¹⁾. Аналогично получаем действие и тензор энергии-импульса плоской гравитационной волны:

$$S_k[\mathfrak{G}_i^j] = \int L_T a^3 dt dx, \quad L_T = \frac{1}{16\pi G} \mathfrak{G}_{i,\nu}^j \mathfrak{G}_j^{i,\nu}, \quad (1.12.26)$$

$$\tau_{\mu\nu} = \frac{1}{8\pi G} \mathfrak{G}_{i,\mu}^j \mathfrak{G}_j^{i,\nu} - L_T g_{\mu\nu}^{(F)}.$$

Полученные уравнения легко обобщаются на ансамбль гравитационных волн с разными волновыми векторами. В этом случае удобно перейти к представлению элементарных осцилляторов, каждый из которых описывает отдельную плоскую волну с соответствующим лагранжианом, получаемым с помощью пространственного интегрирования плотности (1.12.26):

$$L_k = a^3 \int L_T dx. \quad (1.12.27)$$

Тогда общее действие для ансамбля гравитационных волн представимо в виде суммы действий элементарных осцилляторов (1.12.27) (подробнее см. главу 4). Как видим, в поперечно-бесследовой калибровке возможно непротиворечивое введение тензора энергии-импульса гравитационных волн: уравнение $\tau_{\mu;\nu}^\nu = 0$ совпадает с (1.12.23).

¹⁾ Напомним, что $g_{\mu\nu}^{(F)}$ — это метрический тензор Фридмана (см. (1.10.6)),
 $\eta = \int \frac{dt}{a}$, $\bar{H} = \frac{a'}{a}$, $\bar{\gamma} = -\frac{\bar{H}'}{\bar{H}^2}$ (подробнее см. раздел 4.1).

ГРАВИТИРУЮЩИЕ СИСТЕМЫ

В этой главе мы продолжим изучение геометрических свойств пространства-времени и исследуем ряд задач, в которых гравитационное взаимодействие играет первостепенную роль.

2.1. Отклонение света массивным телом

Рассмотрим распространение света в пределе геометрической оптики вне тела массы M на расстояниях много больше его гравитационного радиуса:

$$r_{\min} \gg r_g \equiv 2GM, \quad (2.1.1)$$

где r_{\min} — прицельный параметр — минимальное расстояние, на котором луч света проходит мимо гравитирующего тела.

Если бы тела (звезды) не было, то свет распространялся бы по прямой. Гравитационное поле звезды заставляет свет отклоняться.

При далеком пролете можно пользоваться приближением слабого гравитационного поля:

$$g_{\alpha\beta} = \eta_{\alpha\beta} + 2\Phi\delta_{\alpha\beta}, \quad -\Phi = \frac{GM}{r} \ll 1, \quad (2.1.2)$$

и угол отклонения от первоначального направления движения будет мал ¹⁾. Как описать это отклонение?

Траектория света описывается нулевой геодезической (см. (1.5.7), (1.6.4)):

$$\begin{cases} \frac{dk_\alpha}{d\lambda} = \Phi_{,\alpha} (k^{02} + \mathbf{k}^2), \\ \frac{dk_0}{d\lambda} = 0, \quad \frac{dk_i}{d\lambda} = 2\Phi_{,i}, \end{cases} \quad (2.1.3)$$

где $k^\alpha = (k^0, \mathbf{k})$ — волновой вектор, касательный к траектории. Поскольку поле (2.1.2) не зависит от времени, k_0 является интегралом движения, далее положим $k_0 = 1$. Совместим начало

¹⁾ Для Солнца $r_g \simeq 3$ км. По сравнению с радиусом Солнца (700 000 км) это совсем небольшая величина, что делает использование предела слабого поля в этом случае обоснованным. Полный угол смещения изображения далекого источника, проходящего вблизи видимого края Солнца, равен $\psi_\odot = 1,75''$.

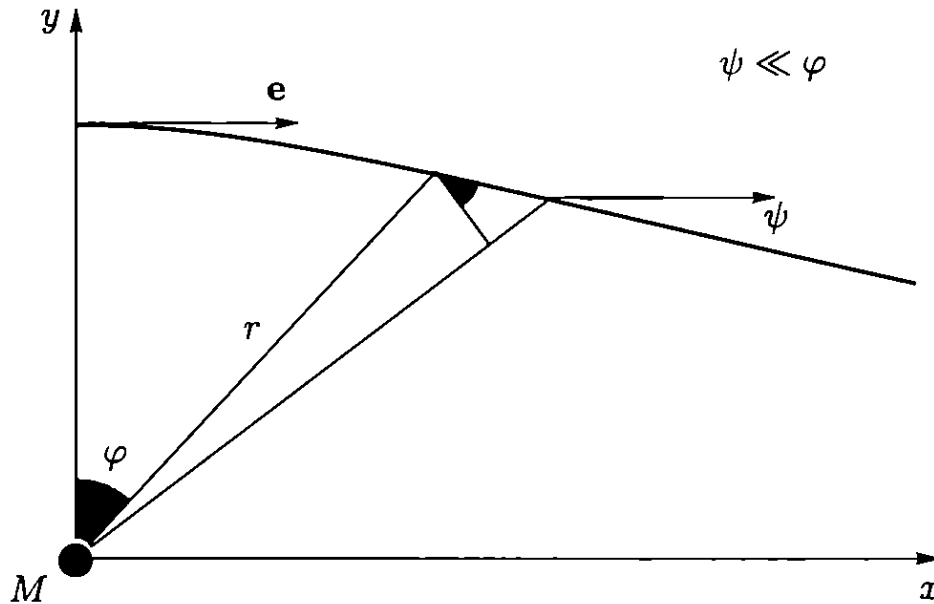


Рис. 2.1. Плоскость (x, y) траектории луча света в поле массивного тела M ; $\mathbf{k} = (1, -\psi)$ — пространственная часть волнового вектора света, вектор $\mathbf{e} = (1, 0)$ — это \mathbf{k} на минимальном расстоянии $\varphi = 0$ траектории от M

координат с центром массы M , а плоскость, в которой движется свет, с плоскостью $z = 0$. Ось x направим в сторону распространения света, а ось y проведем через ближайшую к M точку траектории (рис. 2.1). Компоненты волнового вектора получаем из условия нормировки $k_\alpha k^\alpha = 0$:

$$\begin{aligned} k^\alpha &= (1 - 2\Phi, 1, -\psi, 0), \\ k_\alpha &= (1, -1 + 2\Phi, \psi, 0), \end{aligned}$$

где $\psi \equiv -k^y/k^x$ — угол отклонения луча от оси x ($|\psi| \ll 1$).

Из (2.1.3) следует уравнение для угла отклонения:

$$\frac{d\psi}{d\lambda} = 2\Phi_{,y}. \quad (2.1.4)$$

Решать это уравнение не надо, поскольку полный угол отклонения луча света в гравитационном поле массивного тела получается интегрированием изменения ψ вдоль траектории распространения света:

$$\psi_{\text{tot}} = \int_{-\infty}^{\infty} \frac{d\psi}{d\lambda} d\lambda = 4 \int_0^{\infty} \Phi_{,y} d\lambda = 2r_g \int_0^{\pi/2} \frac{d\varphi}{r} = \frac{2r_g}{r_{\text{min}}},$$

а значит,

$$\psi_{\text{tot}} = \frac{4GM}{r_{\text{min}}}. \quad (2.1.5)$$

Эта формула учитывает как притяжение к телу ($g_{00,y}$), так и релятивистское искривление пространства в окрестности массивного тела ($g_{xx,y}$), что удваивает угол отклонения по сравнению с ньютоновской теорией.

Обратимся теперь к задаче об отклонении света вблизи компактной массы на расстоянии $r > r_g$.

2.2. Метрика Шварцшильда

Найдем сферически-симметричную метрику вблизи точечной гравитирующей массы.

В статическом гравитационном поле все компоненты метрического тензора зависят только от координат и не зависят от времени:

$$ds^2 = A(r) dt^2 - B(r) dr^2 - r^2 d\Omega, \quad (2.2.1)$$

где $d\Omega \equiv de^2 = d\theta^2 + \sin^2 \theta d\varphi^2$ — квадрат интервала на поверхности единичной сферы ($|e| = 1$). Координата r выбрана из соображений, чтобы площадь сферы на расстоянии r равнялась $4\pi r^2$, то есть r есть пространственная кривизна, а не радиальное расстояние. Однако мы будем использовать r в качестве радиальной координаты как метку расстояния, а функции A и B нам предстоит найти.

Итак, (2.2.1) — это общая статическая сферически-симметричная метрика. Поскольку источником является локализованная в $r = 0$ масса, то метрика задана в пустоте и тензор энергии-импульса материи равен нулю везде, кроме точки $r = 0$.

Ненулевые коэффициенты связности имеют вид:

$$\begin{aligned} \Gamma_{tr}^t &= \frac{A'}{2A}, & \Gamma_{rr}^r &= \frac{B'}{2B}, & \Gamma_{tt}^r &= \frac{A'}{2B}, \\ \Gamma_{\theta\theta}^r &= -\frac{r}{B}, & \Gamma_{\varphi\varphi}^r &= -\frac{r \sin^2 \theta}{B}, & \Gamma_{r\theta}^\theta &= \Gamma_{r\varphi}^\varphi = \frac{1}{r}, \\ \Gamma_{\varphi\varphi}^\theta &= -\sin \theta \cos \theta, & \Gamma_{\varphi\theta}^\varphi &= \operatorname{ctg} \theta, \end{aligned}$$

где штрих означает производную по r . Соответственно компоненты тензора Риччи равны

$$\begin{aligned} R_{tt} &= \frac{A''}{2B} - \frac{A'B'}{4B^2} - \frac{(A')^2}{4AB} + \frac{A'}{rB}, \\ R_{rr} &= -\frac{A''}{2A} + \frac{A'B'}{4AB} + \frac{(A')^2}{4A^2} + \frac{B'}{rB}, \\ R_{\theta\theta} &= 1 - \frac{1}{B} + \frac{r}{2B} \left(\frac{B'}{B} - \frac{A'}{A} \right) = \frac{R_{\varphi\varphi}}{\sin^2 \theta}. \end{aligned}$$

Поскольку вне центра в силу уравнений Эйнштейна

$$R_{tt} = R_{rr} = R_{\theta\theta} = 0,$$

то, разделив R_{tt} на A , а R_{rr} на B и сложив оба выражения, получаем:

$$\frac{R_{tt}}{A} + \frac{R_{rr}}{B} = \frac{1}{rB} \left(\frac{A'}{A} + \frac{B'}{B} \right) = 0, \quad (2.2.2)$$

откуда следует, что

$$A(r) \cdot B(r) = \text{const.} \quad (2.2.3)$$

Теперь воспользуемся ньютоновским пределом: на больших расстояниях от звезды наблюдается метрика Минковского плюс ньютоновский потенциал. Тогда константа в (2.2.3) равна единице и $B = A^{-1}$.

Используя уравнение $R_{\theta\theta} = 0$, найдем уравнение для A :

$$R_{\theta\theta} = 1 - A - rA' = 0. \quad (2.2.4)$$

Решая, в свою очередь, это уравнение, получаем:

$$A = 1 + \frac{\text{const}}{r}.$$

И опять, припоминая, как выглядит ньютоновское поле, заключаем, что эта константа должна равняться $2GM$:

$$A = 1 - \frac{2GM}{r}. \quad (2.2.5)$$

Предельный радиус $r_g = 2GM$, на котором $A = 0$, называется *радиусом Шварцшильда*.

Таким образом, искомая сферически-симметричная метрика точечного тела имеет вид:

$$ds^2 = \left(1 - \frac{2GM}{r}\right) dt^2 - \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 - r^2 d\Omega. \quad (2.2.6)$$

Метрика Шварцшильда обладает очень важным свойством: статическое решение имеется только при $r > r_g$. Почему внутри радиуса Шварцшильда никакой статики нет?

Физически это объясняется высокой интенсивностью гравитационного поля. Математически все дело в том, что внутри этого радиуса координата r имеет смысл времени. Напомним, что с геометрической точки зрения время — это координата с положительной сигнатурой (вернее, с сигнатурой, противоположной сигнатуре пространственных координат). Иногда говорят, что при $r < r_g$ пространство и время меняются местами, но, по-видимому, корректнее говорить, что метрика нестатична. Это

звучит менее впечатляюще, но зато понятно и ясно по смыслу. Статическая область $r > r_g$ называется R -областью, а нестатическая $r < r_g$ — T -областью.

В T -области сечения $r = \text{const}$ описывают однородное пространство второго типа по классификации Бианки (см. раздел 3.3). Если тело попало внутрь радиуса Шварцшильда, то за конечное время оно достигает сингулярной гиперповерхности $r = 0$; при этом вдоль направления t на тело действует бесконечная сила растяжения, а в поперечных направлениях — сила сжатия. Подобные по структуре силы (но конечные) действуют на тело в любой точке его траектории, в R -области растяжка происходит вдоль радиального направления. Такая деформация характерна для гравитационных приливных сил.

Метрика (2.2.6) представляет собой гравитационное поле массы, локализованной в области $r = 0$, и описывает как *черную*, так и *белую* дыры в зависимости от начальных условий и границ геометрии. Мы рассмотрим вопрос о взаимосвязи черных дыр и космологии в последующих разделах. Здесь же мы обратимся к черной дыре астрофизического происхождения и посмотрим, как свет распространяется в ее гравитационном поле. При близком пролете около черной дыры ее гравитационное поле сильно отклоняет мировую линию света и становится возможным даже «наматывание» траектории в окрестности сферы Шварцшильда.

Поскольку движение происходит в центрально-симметричном поле, мы можем рассматривать траекторию луча и массивное тело в одной плоскости $\theta = \pi/2$. Таким образом, траектория луча зависит только от угла φ , то есть нам нужно определить, как при пролете изменяется радиальная координата в зависимости от угла φ :

$$r = r(\varphi).$$

В принципе, надо решать уравнения геодезических. На практике в этом необходимости нет, поскольку в силу симметрии имеются интегралы движения. Запишем уравнение геодезической в виде (1.6.4):

$$\frac{dk_\alpha}{d\lambda} = \frac{1}{2} g_{\beta\gamma, \alpha} k^\beta k^\gamma, \quad k^\alpha = \frac{dx^\alpha}{d\lambda}.$$

Так как метрика не зависит от времени t и угла φ , то и дифференциалы соответствующих компонент k_α также равны нулю. А это и есть интегралы движения:

$$k_0 = \text{const}, \quad k_\varphi = \text{const}, \quad (2.2.7)$$

означающие сохранение энергии и момента импульса движения.

Распишем квадрат импульса

$$k_\alpha k^\alpha = k_0 k^0 + k_r k^r + k_\theta k^\theta + k_\varphi k^\varphi = 0 \quad (2.2.8)$$

и заметим, что $k_\theta = 0$, $k_r = g_{rr} k^r$, $k_0 = g_{00} k^0$, а также $g_{00} g_{rr} = -1$. Тогда выражение (2.2.8) можно переписать в виде:

$$\left(\frac{k^r}{k_0}\right)^2 = 1 - \frac{\kappa_0^2}{r^2} \left(1 - \frac{r_g}{r}\right), \quad (2.2.9)$$

где

$$\kappa_0 \equiv \frac{k_\varphi}{k_0} = \text{const}$$

— прицельный параметр — мера углового момента, выраженная в единицах энергии.

От зависимости $k^r = k^r(r)$ нам надо перейти к $r = r(\varphi)$. Сначала преобразуем функцию k^r . По определению, это $dr/d\lambda$, но сейчас удобнее записать ее как $(dr/dt) \cdot (dt/d\lambda)$, а $dt/d\lambda$ — это есть k^0 . Итак,

$$\frac{dr}{d\lambda} = \frac{dr}{dt} \frac{dt}{d\lambda} = \frac{dr}{dt} g^{00} k_0,$$

откуда, с учетом $\dot{\varphi} = k^\varphi/k^0$, следует:

$$\frac{k^r}{k_0} = \frac{\dot{r}}{1 - r_g/r} = -\frac{\kappa_0 r_{,\varphi}}{r^2}. \quad (2.2.10)$$

Далее, возвращаясь к (2.2.9), получаем искомое уравнение траектории:

$$-\frac{\kappa_0}{r^2} \frac{dr}{d\varphi} = \sqrt{1 - u^2}, \quad (2.2.11)$$

где

$$u = \frac{\kappa_0}{r} \sqrt{1 - \frac{r_g}{r}} = \frac{\kappa_0}{r} \left(1 - \frac{r_g}{2r} + O\left(\frac{r_g^2}{r^2}\right)\right).$$

В момент, когда r достигает своего минимального значения, квадратный корень в (2.2.11) обращается в нуль, то есть $u = 1$ — это точка наиболее близкого пролета. Однако эта функция не всегда может дорасти до 1. Это означает, что свет будет «проглочен» черной дырой, и точки разворота никогда не будет.

Максимум функции u соответствует $r = 3r_g/2$. Таким образом, если фотон попал внутрь орбиты с радиусом $3r_g/2$, то он неизбежно будет «проглочен» черной дырой. Если же минимальное расстояние между его траекторией и черной дырой больше этого значения, то фотон может сделать несколько витков и толь-

ко после этого улетит. А может получиться и так, что он вечно останется на орбите $3r_g/2$.

Теперь обратим уравнение (2.2.11) и определим угол поворота траектории:

$$\Delta\varphi = \int_0^1 \frac{d(\kappa_0/r)}{\sqrt{1-u^2}}. \quad (2.2.12)$$

Это эллиптический интеграл. Вычислим его в пределе слабого гравитационного поля, полагая $r_g/r \ll 1$. Тогда

$$\begin{aligned} \Delta\varphi &= \int_0^1 \frac{d(\kappa_0/r)}{\sqrt{1-u^2}} = \int_0^1 \frac{du}{\sqrt{1-u^2}} \left(1 + u \frac{r_g}{\kappa_0} + \dots \right) = \\ &= \arcsin u + \frac{r_g}{\kappa_0} \left(1 - \sqrt{1-u^2} \right) \rightarrow \frac{\pi}{2} + \frac{r_g}{\kappa_0} \end{aligned} \quad (2.2.13)$$

при $u \rightarrow 1$ ($r_{\min} = \kappa_0$). В этом выражении первый член описывает траекторию луча света в пространстве Минковского, второй — поправку, связанную с отклонением реального пространства от пространства Минковского. Полный угол поворота — это удвоенная величина (2.2.13), причем π есть поворот в неискривленном пространстве Минковского, а $2r_g/r_{\min}$ — угол отклонения луча света массивным полем. Таким образом, мы подтвердили результат (2.1.5).

В гравитационном поле Солнца величина угла отклонения луча света составляет (см. (1.7.10))

$$\psi_{\odot} = \frac{4GM_{\odot}}{r_{\odot}} \simeq 1,75''.$$

Характерные углы отклонения света другими астрофизическими объектами — галактиками, скоплениями и сверхскоплениями галактик — также весьма малы. Тем не менее, эффекты гравитационного линзирования присутствуют повсеместно в космологических наблюдениях, и их необходимо учитывать при определении космологических параметров.

В силу важности этого эффекта для наблюдательной и теоретической космологии мы рассмотрим основы теории гравитационного линзирования в разделах 2.7 и 2.8, а сейчас продолжим исследование простейших гравитирующих систем и свойств пространства-времени.

2.3. Связь космологии с белой дырой

Согласно релятивистской теории гравитации, черные дыры образуются в результате коллапса массивных звезд. На конечной стадии эволюции звезды внутреннее давление становится недостаточным для того, чтобы противостоять ее гравитационному сжатию под действием собственного гравитационного поля.

Ниже мы рассмотрим сферически-симметричную модель коллапса, в которой центральное тяготеющее тело описывается однородной пылевидной средой без давления. Пространство-время внутри такой «звезды» соответствует космологии Фридмана, поскольку движение пылинок на расстоянии R от центра зависит только от полной массы внутри шара радиуса R . Обращая эту задачу на случай расширения, мы увидим, как космология внутри шара выглядит для внешнего наблюдателя.

Для однородного пылевого шара метрика имеет вид:

$$\begin{cases} d\tau^2 - a^2 (dR^2 + R^2 d\Omega), & R \in (0, R_0), \\ \left(1 - \frac{r_g}{r}\right) dt^2 - \frac{dr^2}{1 - r_g/r} - r^2 d\Omega, & r \geq r_0. \end{cases} \quad (2.3.1)$$

Как видно, снаружи шара радиуса R_0 метрика шварцшильдовская, а внутри этого шара — фридмановская. Если мы хотим рассмотреть коллапс, то $a(t)$ — это убывающая функция времени.

Можно ли иметь в шаре ту же метрику, что и в бесконечном однородном фридмановском мире? Можно, если давление равно нулю везде вплоть до границы шара. Если давление не равно нулю, то на границе появятся градиенты давления и шар начнет деформироваться, а значит, он перестанет быть однородным и нарушится условие фридмановской метрики.

Итак, у нас есть две метрические сетки — одна внутри, а другая снаружи. Эти сетки можно «сшить», для чего удобно использовать условие, что пылинка, лежащая на границе, с точки зрения внутреннего наблюдателя принадлежит фридмановской метрике с координатой $R = R_0$, а с точки зрения наружного наблюдателя движется по радиальной геодезической $r_0(t)$ в метрике Шварцшильда. Это означает

$$r_0 = aR_0, \quad (2.3.2)$$

что обеспечивает непрерывность кривизны при переходе через границу шара.

Теперь необходимо произвести сшивку временной части метрики. С точки зрения фридмановской системы отсчета, которая для пылинки является сопутствующей, интервал на мировой линии пылинки на границе шара равен $d\tau^2$, а с точки зрения шварцшильдовской метрики надо учитывать обе координаты t и r :

$$d\tau^2 = \left(1 - \frac{r_g}{r}\right) dt^2 - \frac{dr^2}{1 - \frac{r_g}{r}}.$$

После несложных преобразований получаем:

$$\frac{dt}{d\tau} \left(1 - \frac{r_g}{r}\right) = \sqrt{1 - \frac{r_g}{r} + \left(\frac{dr}{d\tau}\right)^2}. \quad (2.3.3)$$

Решим внутреннюю задачу и выясним, как эволюционирует масштабный фактор. Запишем уравнение Фридмана:

$$H^2 = \frac{8\pi G}{3} \rho = \frac{r_g}{r_0^3}, \quad (2.3.4)$$

где $H \equiv \frac{\dot{a}}{a} = \frac{2}{3\tau}$ — функция Хаббла (точка над буквой означает производную по времени τ), ρ — плотность шара, $r_g = 2GM_0$ — гравитационный радиус шара массой $M_0 = 4\pi\rho r_0^3/3$.

В ходе коллапса совокупная масса всех частиц в шаре и внешняя гравитирующая масса шара, измеряемая шварцшильдовским наблюдателем, не изменяются со временем. Полная масса шара складывается из трех частей. Первая часть — это сумма масс покоя всех его частиц, вторая — гравитационная связь (потенциальная энергия), и третья — кинетическая энергия шара. В данном случае масса шара совпала с массой покоя всех составляющих его частиц: это значит, что сумма потенциальной и кинетической энергий шара равна нулю.

Получилось так потому, что фридмановское пространство плоское, и следовательно, реализовано параболическое движение шара как при его сжатии, $\tau \in (-\infty, 0)$, так и при расширении, $\tau \in (0, \infty)$. Соответственно масштабный фактор меняется как

$$a = \frac{r_g}{R_0} \left(\frac{3\tau}{2r_g}\right)^{2/3}.$$

Это позволяет нам определить r_0 в единицах шварцшильдовского радиуса внешней массы M :

$$r_0 = r_g \left(\frac{3\tau}{2r_g}\right)^{2/3}. \quad (2.3.5)$$

Поясним смысл этого соотношения. Пусть шар первоначально имеет большие размеры и покоится. Если его «отпустить», то он начнет коллапсировать под действием гравитации. Снаружи шара метрика шварцшильдовская, значит имеется гравитационный радиус, которого граница шара достигает в некоторый момент времени. При дальнейшем коллапсе величина $3|\tau|/2r_g$ становится меньше единицы и шар уходит под r_g . Начиная с этого момента, с точки зрения внешнего наблюдателя имеется только черная дыра с размером r_g ¹⁾.

Теперь можно взять производную $dr/d\tau$:

$$\frac{dr_0}{d\tau} = \frac{2r_0}{3\tau} = Hr_0 = \sqrt{\frac{r_g}{r_0}}. \quad (2.3.6)$$

Скорость $dr/d\tau$ удовлетворяет хаббловскому закону, как и должно быть с точки зрения фридмановского наблюдателя. А с точки зрения шварцшильдовского наблюдателя это $\sqrt{r_g/r_0}$. Возвращаясь немного назад к выражению (2.3.3), получаем

$$\frac{dt}{d\tau} \left(1 - \frac{r_g}{r}\right) = \sqrt{1 - \frac{r_g}{r} + \left(\frac{dr}{d\tau}\right)^2} = 1, \quad (2.3.7)$$

откуда следует связь между шварцшильдовским и фридмановским временами:

$$\left(1 - \frac{r_g}{r}\right) dt = d\tau. \quad (2.3.8)$$

С точки зрения внешнего наблюдателя имеется постоянный гравитационный радиус r_g . А с точки зрения внутреннего наблюдателя никакого r_g нет, а есть характерный хаббловский радиус H^{-1} — радиус пространственно-временной кривизны, зависящий от времени. Это радиус причинно-связанной области, все точки внутри которой успевают обменяться световыми сигналами за время изменения масштабного фактора вдвое. В момент, когда граница шара r_0 пересекает r_g , размер причинно-связанной области тоже становится равным r_g .

Рисунок 2.2 демонстрирует эволюцию масштабов в космологии (при расширении шара). Хотя с точки зрения внутреннего фридмановского наблюдателя он живет в какой-то весьма обширной вселенной, это не исключает возможности того, что его мир снаружи выглядит гравитирующим телом массы M . H^{-1} — это мера времени. В момент, когда H^{-1} достигает величины r_g ,

¹⁾ Хотя внешний наблюдатель и видит границу шара, приближающуюся к r_g , на нее невозможно высадиться или достичь ее.

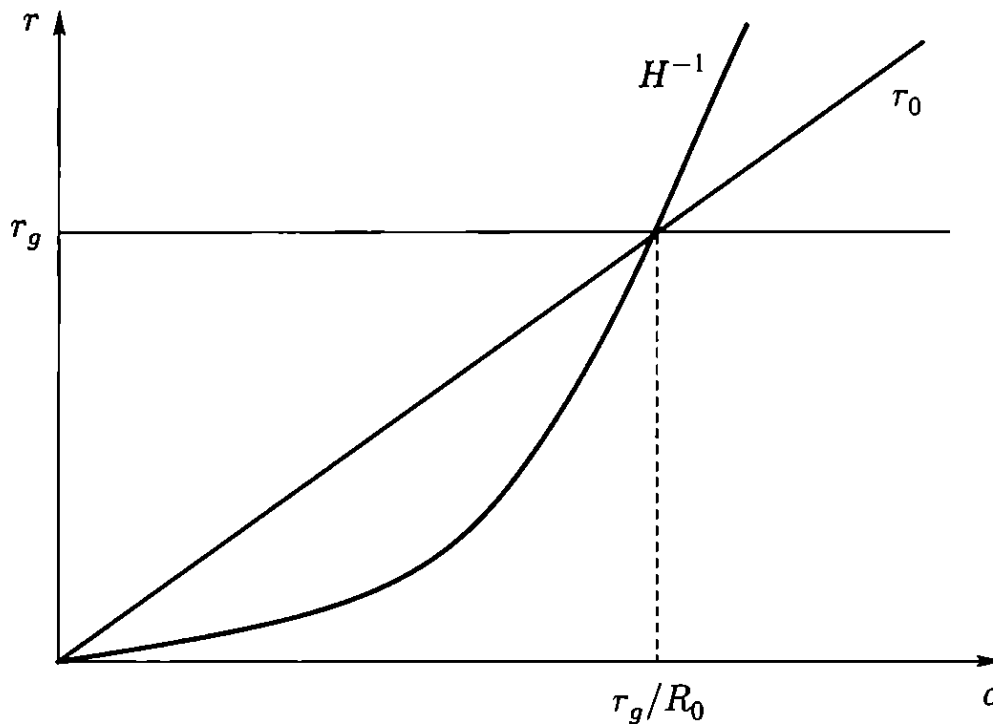


Рис. 2.2. Радиусы $r_0 = aR_0$, $H^{-1} = a/\dot{a}$ и $r_g = 2GM_0$ как функции масштабного фактора a ; $r < r_0$ — область космологии, $r > r_0$ — белая дыра

внешний наблюдатель регистрирует пересечение горизонта событий белой дыры. А внутренний наблюдатель сказал бы, что масштаб r_0 «вошел под горизонт» в его причинную область.

Локальный наблюдатель ощущает себя по-разному вне и внутри гравитирующего шара. В поле Шварцшильда приливные силы разрывают падающее тело, а внутри шара никакого приливного воздействия нет: тело находится в покое, поскольку его внутренняя часть оказывается под гравитационным воздействием меньшей массы, чем внешняя.

2.4. Падение в поле Шварцшильда

Мы исследовали параболическое движение гравитирующего шара. При расширении граница шара выходит из-под r_g и появляется в R -области внешнего наблюдателя. При сжатии шар уходит под r_g и для внешнего наблюдателя остается черная дыра. Такие движения и порождаемые ими геометрии — это следствия выбранных начальных условий. Параболическое движение соответствует плоскому трехмерному пространству внутри шара.

С точки зрения внешнего шварцшильдовского наблюдателя граница шара движется по радиальной геодезической ($u_\varphi = u_\theta = 0$). Прежде всего, это траектория массивной частицы, то есть квадрат 4-скорости равен единице. Кроме того, u_0 является интегралом движения. Уравнение радиальной геодезической в мет-

рике Шварцшильда выглядит следующим образом:

$$u_0 = \text{const}, \quad u_0^2 - (u^r)^2 = 1 - \frac{r_g}{r}. \quad (2.4.1)$$

То, что мы рассматривали эту геодезическую при условии параболического падения, означает, что $u_0 = 1$. Тогда $u^r = \sqrt{r_g/r}$ (см. (2.3.6)). В общем случае

$$u^r = u_0 \frac{dr}{dt} = \frac{u_0 dr}{\left(1 - \frac{r_g}{r}\right) dt},$$

в результате чего уравнение геодезической можно переписать в виде:

$$\frac{u_0^2 \dot{r}^2}{\left(1 - \frac{r_g}{r}\right)^2} - \frac{r_g}{r} = u_0^2 - 1. \quad (2.4.2)$$

Это уравнение описывает динамику радиальной траектории, и в этом виде оно напоминает закон сохранения энергии. Справа стоит полная энергия движения частицы (с точностью до двойки): первый член слева — это кинетическая энергия, второй — гравитационная энергия связи. Для параболического движения сумма членов в левой части (2.4.2) равна нулю. Однако можно рассмотреть не только параболическое движение.

Например, пусть шар конечного размера покоится в некоторый момент времени t . Тогда это гравитационно-связанный объект и $u_0 < 1$ (граница шара не может оказаться на бесконечности). Если $u_0 > 1$, то шар при расширении разлетится: скорость пылинок, из которых он состоит по постановке задачи, окажется больше второй космической скорости. Эти соображения понадобятся нам в дальнейшем при расчете полной энергии звезды.

Перепишем уравнение (2.4.2) в терминах локального импульса частицы:

$$p = m \sqrt{|u^r u_r|},$$

где m — масса частицы. Умножая (2.4.1) на m^2 , получаем

$$m^2 u_0^2 = \left(1 - \frac{r_g}{r}\right) (m^2 + p^2). \quad (2.4.3)$$

Мы видим, что $m u_0$ есть полная энергия частицы с точки зрения далекого наблюдателя: если $m u_0$ меньше массы покоя m , то такая частица никогда не сможет оказаться на бесконечности. Тогда условие удержания частицы с импульсом p , находящейся

в точке r , гравитационным полем массы M можно представить в более привычном виде:

$$\sqrt{m^2 + p^2} - \frac{m}{\sqrt{1 - \frac{2GM}{r}}} < 0. \quad (2.4.4)$$

2.5. Полузамкнутый мир

Обратимся к геометрии однородного шара с $u_0 < 1$. В этом случае пространство внутри шара искривлено и является частью замкнутого мира Фрийдмана.

Метрика полузамкнутого мира состоит из внутренней и внешней частей:

$$\begin{cases} d\tau^2 - a^2 \left(dR^2 + \frac{\sin^2(\sqrt{\kappa} R)}{\kappa} d\Omega \right), & R \in (0, R_0), \\ \left(1 - \frac{r_g}{r} \right) dt^2 - \frac{dr^2}{1 - r_g/r} - r^2 d\Omega, & r \geq r_0. \end{cases} \quad (2.5.1)$$

Здесь $\kappa = \text{const}$ — постоянная трехмерной кривизны. Если $\kappa = 0$, то мир пространственно-плоский¹⁾; если $\kappa > 0$, то мир замкнут; а если $\kappa < 0$, то мир открыт и вместо синуса в (2.5.1) надо писать гиперболический синус.

В геометрии, которую нам предстоит рассмотреть, внутреннее пространство неплоское, и связано это с тем, что начальные условия для шара выбраны не на бесконечности, а на конечном расстоянии. В этом случае энергия связи отрицательна ($\kappa > 0$, $R_0 \leq \pi/\sqrt{\kappa}$) и полная масса уже не равна просто сумме масс пылинок: шварцшильдовская масса шара меньше, чем сумма масс покоя всех составляющих его частиц. В частности, можно выбрать такие начальные условия, что для внешнего наблюдателя масса будет равна нулю или очень мала (например, равна массе электрона), а для внутреннего — это будет огромная вселенная. Может быть, мы живем в таком мире.

Этот кусок замкнутого мира (полузамкнутый мир) построен на поверхности 4-сферы. Представить это можно только по аналогии. Если бы мы были двумерными людьми, это означало бы, что в таком мире мы жили бы на поверхности трехмерного мяча.

В полузамкнутом мире существует горловина, локализованная при $r_0 = r_0(t)$. Теперь, как и в предыдущем примере, необ-

¹⁾ По имеющимся наблюдательным данным, главным образом анизотропии реликтового излучения, наш мир близок к пространственно-плоскому.

ходимо считать два решения. Сшивка кривизны на сфере $R = R_0$ (она же $r = r_0$) дает

$$r_0 = a \frac{\sin(\sqrt{\kappa} R_0)}{\sqrt{\kappa}}, \quad (2.5.2)$$

а сшивка по времени приводит к выражению

$$\left(1 - \frac{r_g}{r}\right) \frac{dt}{d\tau} = \sqrt{1 - \frac{r_g}{r} + \left(\frac{dr}{d\tau}\right)^2} = u_0 = \text{const.} \quad (2.5.3)$$

Решая уравнение Фридмана

$$H^2 = \frac{8\pi G}{3} \rho - \frac{\kappa}{a^2}, \quad (2.5.4)$$

получаем точное решение в форме циклоиды:

$$\begin{cases} a = a_0 (1 - \cos(\sqrt{\kappa} \eta)), \\ \tau = a_0 \left(\eta - \frac{\sin(\sqrt{\kappa} \eta)}{\sqrt{\kappa}} \right), \end{cases} \quad (2.5.5)$$

где $\eta = \int \frac{d\tau}{a} \in (0, 2\pi/\sqrt{\kappa})$ — конформное время, а $4\pi G \rho a^3 \equiv \equiv 3a_0 \kappa = \text{const.}$

Уравнение Фридмана (2.5.4) для границы шара можно переписать в терминах метрики Шварцшильда (см. (2.5.3)):

$$\left(\frac{dr_0}{d\tau}\right)^2 = \frac{r_g}{r_0} - \sin^2(\sqrt{\kappa} R_0), \quad u_0 = \cos(\sqrt{\kappa} R_0). \quad (2.5.6)$$

Таким образом, полная внешняя масса полузамкнутого мира (масса горловины) равна

$$M = \frac{4\pi}{3} \rho a^3 \left(\frac{\sin(\sqrt{\kappa} R_0)}{\sqrt{\kappa}} \right)^3. \quad (2.5.7)$$

При $R_0 \rightarrow \pi/\sqrt{\kappa}$ масса горловины стремится к нулю.

По аналогии с радиусом горловины (2.5.2) можно записать и радиус Шварцшильда: $r_g = a_g r_0 / a$, где $a_g = 2a_0 \sin^2(\sqrt{\kappa} R_0)$. Тогда из (2.5.5) следует, что геодезическая граница шара r_0 пересекает r_g в моменты времени $\eta_1 = 2 \min \{ R_0, \pi \kappa^{-1/2} - R_0 \}$ при расширении и $\eta_2 = 2 \max \{ R_0, \pi \kappa^{-1/2} - R_0 \}$ при сжатии. Чтобы разобраться во всех радиальных геодезических в геометрии Шварцшильда, нам следует обратиться к геодезически полным картам.

2.6. Испарение черных дыр

Для построения геодезически полных карт удобно использовать координаты, в которых траектории света — это прямые линии, расположенные под углом 45° к координатным осям. Прямым суммированием таких координат легко перейти к опережающим и запаздывающим нулевым координатам, вдоль которых распространяется свет. Чтобы проследить судьбу всех световых геодезических, используемая метрика не должна иметь особенностей.

Для перехода к нулевым координатам в метрике Шварцшильда введем вместо радиуса r черепашьёю координату x :

$$x = \int \frac{dr}{1 - \frac{r_g}{r}} = r + r_g \ln \left| \frac{r}{r_g} - 1 \right|.$$

В световых координатах

$$u = t + x, \quad v = t - x \quad (2.6.1)$$

метрика (2.2.6) принимает вид:

$$ds^2 = \left(1 - \frac{r_g}{r}\right) du dv - r^2 d\Omega. \quad (2.6.2)$$

Линии $u = \text{const}$ или $v = \text{const}$ и есть радиальные световые геодезические.

Итак, требуемые траектории света мы построили, но особенность в метрике при $r = r_g$ все еще осталась. Ее можно убрать экспоненциальным преобразованием:

$$\begin{cases} t' = \exp\left(\frac{u}{2r_g}\right) - \exp\left(-\frac{v}{2r_g}\right), \\ x' = \exp\left(\frac{u}{2r_g}\right) + \exp\left(-\frac{v}{2r_g}\right). \end{cases} \quad (2.6.3)$$

Безразмерные координаты (t', x') называются координатами Крускала (рис. 2.3). В этих координатах

$$\begin{aligned} ds^2 &= \frac{r_g^3}{r} \exp\left(-\frac{r}{r_g}\right) (dt'^2 - dx'^2) - r^2 d\Omega, \\ t'^2 - x'^2 &= 4 \left(1 - \frac{r}{r_g}\right) \exp\left(\frac{r}{r_g}\right). \end{aligned} \quad (2.6.4)$$

Как видно, гиперповерхность $r = r_g$ является нулевой (горизонт событий). Теперь все пространство-время можно исследовать спомощью световых геодезических.

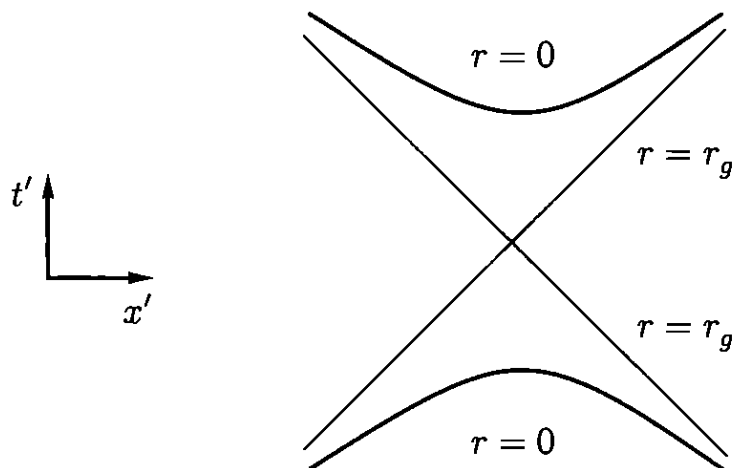


Рис. 2.3. Координаты Крускала

Как и предполагалось, источник вакуумной геометрии находится в T -областях на сингулярных пространственноподобных гиперповерхностях $r = 0$. Таких гиперповерхностей две: нижнюю можно назвать источником белой дыры, а на верхней заканчиваются геодезические, попавшие в черную дыру. Кроме того, появилась вторая R -область, причинно не связанная с первой. Как это понимать?

Мы получили так называемую *вечную* белую/черную дыру. Эта геометрия является следствием поставленных граничных условий: источник находится в области $r = 0$. Для ее реализации необходимо «пристыковать» левую и правую R -области в центральной точке $t' = x' = 0$, тогда особенности в вакуумной метрике не возникнут.

Координаты Крускала можно компактифицировать:

$$\begin{cases} t'' = \operatorname{arctg} \left(\frac{u'}{2} \right) + \operatorname{arctg} \left(\frac{v'}{2} \right), \\ x'' = \operatorname{arctg} \left(\frac{u'}{2} \right) - \operatorname{arctg} \left(\frac{v'}{2} \right), \end{cases} \quad (2.6.5)$$

где $u' = t' + x'$, $v' = t' - x'$. Компактное изображение всей истории пространства-времени в световых координатах называется диаграммой Пенроуза. На рис. 2.4 изображена диаграмма Пенроуза вечной белой/черной дыры в асимптотически плоском пространстве Минковского.

В реальности классические вечные дыры не могут существовать в вакууме Минковского: их нестационарное гравитационное поле, существующее в T -областях, является причиной параметрического рождения пар частиц из вакуума, что приводит к динамической перестройке решения и к изменению всей геометрии.

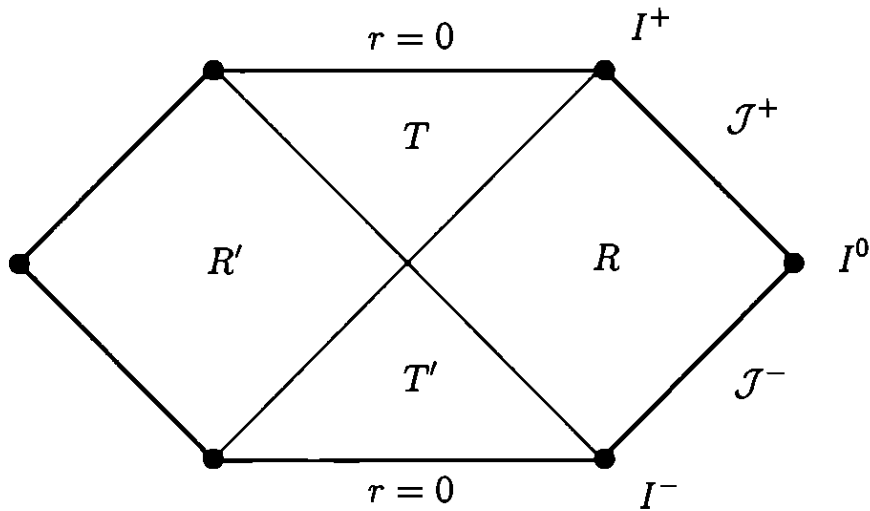


Рис. 2.4. Диаграмма Пенроуза вечной белой/черной дыры. I^- , I^+ и J^- , J^+ — соответственно временные и световые бесконечности прошлого и будущего, I^0 — пространственная бесконечность R -области

Тем не менее, черные дыры существуют, но с другой историей и с другими диаграммами.

Мы уже упоминали о том, что черные дыры могут образовываться в естественных условиях на конечной стадии эволюции массивных звезд, а также в результате динамических процессов, происходящих в центральных областях галактик (газовая и звездная релаксация, слияние, аккреция и пр.). Однако и здесь необходимо принимать во внимание эффект рождения частиц.

Диаграмма Пенроуза астрофизической черной дыры представлена на рис. 2.5. Зона, занятая материей звезды, затемнена. Светлая (вакуумная) зона также содержит нестационарную T -область, которая вызывает процесс спонтанного рождения пар частиц (приливные силы с частотой $\omega \sim r_g^{-1}$ (см. раздел 1.11, (4.3.7)) разрывают виртуальные пары частиц в радиальном направлении). В силу нелокальности эффекта рождения в R -области существует поток частиц, испаряющихся с «поверхности» черной дыры и уносящих массу на световую J^+ и временную I^+ бесконечности (процесс испарения Хокинга). Парные к ним частицы остаются внутри горизонта событий черной дыры и падают на сингулярную гиперповерхность $r = 0$, уменьшая ее полную массу. Они имеют отрицательную энергию по отношению к внешнему наблюдателю из-за гравитационного дефекта масс (см. (2.4.3)). Поэтому полная энергия пары частиц, одна из которых находится в области R , а другая в T , равна нулю, и закон сохранения энергии при рождении частиц из вакуума в такой геометрии не нарушается.

Спектральные характеристики хокинговского излучения можно получить из общих законов термодинамики. Часть

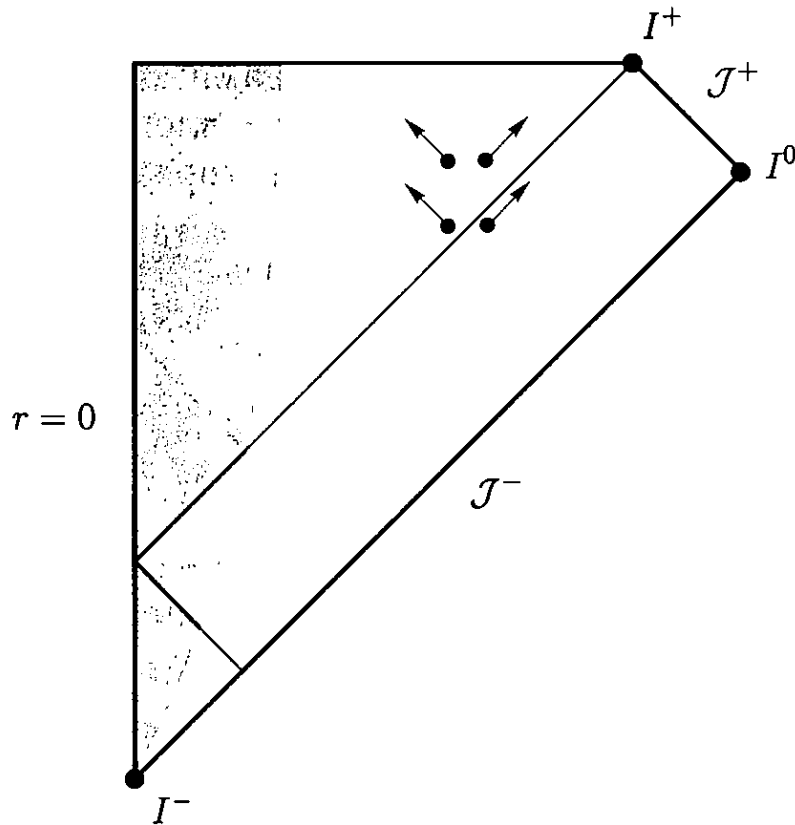


Рис. 2.5. Диаграмма Пенроуза астрофизической черной дыры. Линия $r = 0$ имеет времениподобную часть (центр звезды) и пространственноподобную часть (сингулярность)

степеней свободы, описывающих микроструктуру пространства-времени и материальных полей, спрятана внутри горизонта событий и недоступна внешнему наблюдателю. На макроскопическом уровне эти степени свободы можно характеризовать энтропией и температурой самой двухмерной поверхности горизонта (*голографический принцип* — информация о недоступной области находится на ее границе)¹⁾. По закону термодинамики уменьшение массы черной дыры происходит пропорционально потере ее энтропии S , которая, в свою очередь, пропорциональна площади горизонта черной дыры s :

$$s = s(M) = 4\pi r_g^2, \quad dM = T_H dS, \quad (2.6.6)$$

$$S = \frac{s}{4\ell_P^2} = 4\pi \frac{M^2}{M_P^2}. \quad (2.6.7)$$

¹⁾ Здесь мы сталкиваемся с универсальностью ОТО и термодинамики как *единой* макроскопической теории (классической физики), описывающей мир с точки зрения разных наблюдателей. В частности, отсюда следует, что оба важнейших понятия современной физики — «пространство-время» и «материя» — являются единым квазиклассическим пределом более общей микроскопической (квантовой) теории (см. также раздел 5.4).

Коэффициент пропорциональности $1/4$ между энтропией и площадью черной дыры вычисляется методами квантовой статистики: энтропия оказывается равной числу площадок планковского размера, которыми можно покрыть горизонт событий черной дыры.

Отсюда находим эффективную температуру черной дыры, измеряемую неподвижным наблюдателем в асимптотическом пространстве Минковского ¹⁾,

$$T_H = \frac{dM}{dS} = \frac{1}{4\pi r_g} = \frac{M_P^2}{8\pi M}$$

и полное время испарения массы M

$$\frac{dM}{dt} \simeq -sT_H^4, \quad t_0 \simeq 10^{71} \left(\frac{M}{M_\odot} \right)^3 \text{ с.} \quad (2.6.8)$$

Мы видим, что процессы квантово-гравитационного испарения для реальных астрофизических черных дыр несущественны. За все время существования Вселенной успевают полностью испариться лишь дыры с массами, не превышающими 10^{15} г. Столь малые массы могут представлять интерес лишь при исследовании гипотетических *первичных* черных дыр, которые могли образоваться на очень ранних стадиях эволюции Вселенной.

Рисунок 2.6 демонстрирует диаграмму Пенроуза испаряющейся черной дыры, образовавшейся в результате параболического коллапса однородного пылевого шара в асимптотически плоском пространстве Минковского. Если бы траектория границы шара начиналась на пространственноподобной гиперповерхности прошлого $r = 0$, то мы получили бы историю испаряющейся горловины полузамкнутого мира.

Какие частицы испаряются из черной дыры во внешнее пространство R ? В принципе, это все частицы с массами, не превышающими T_H : спектр частиц тепловой с учетом конечного размера дыры («серое» тело). В первую очередь испаряются безмассовые и маломассивные частицы: фотоны, гравитоны, нейтрино и пр. По мере уменьшения внешней массы черной дыры

¹⁾ Можно поступить иначе и вычислить T_H с помощью эффекта Унру, устанавливающего, что наблюдатель, движущийся с ускорением a в вакууме Минковского, детектирует его как тепловую баню частиц с температурой $T_U = a/2\pi$. Температура Хокинга T_H равна температуре Унру T_U свободно падающего вакуума Минковского на горизонте черной дыры, измеряемой неподвижным наблюдателем ($a = GM/r_g^2$, подробнее см. цитируемую литературу). Тогда, подставляя $T_H = GM/(2\pi r_g^2)$ в (2.6.6), получаем энтропию S в форме (2.6.7).

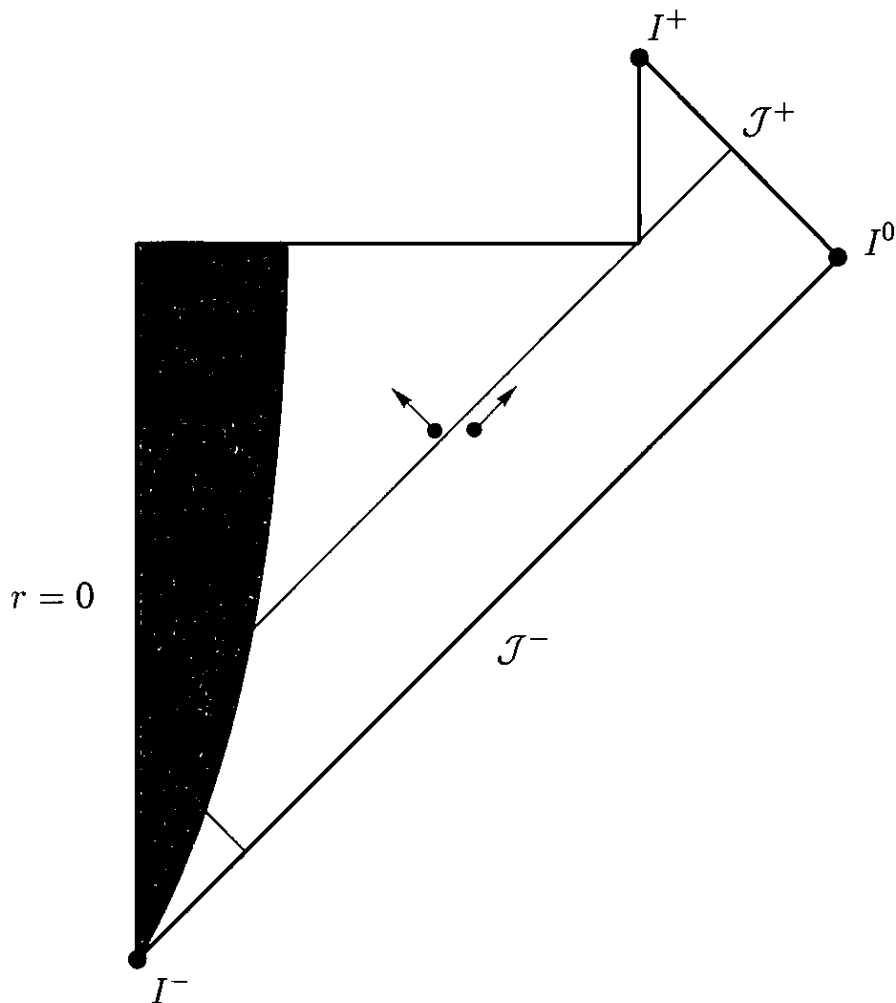


Рис. 2.6. Диаграмма Пенроуза испаряющейся черной дыры

могут испаряться и более массивные частицы. Однако при приближении массы к планковской мы уже не можем пользоваться квазиклассическим пределом для расчета энергии потока и должны учитывать эффекты квантовой гравитации. В частности, пока неизвестно, исчезает ли внешняя масса полностью (полузамкнутый мир замыкается и отщепляется от внешнего пространства) или на месте черной дыры остается топологический дефект (монополь, квантовая горловина полузамкнутого мира).

Сам факт существования черных дыр означает, что во внешнем пространстве не сохраняется барионный заряд: черная дыра образуется путем коллапса звезды (барионов), а испаряется, в основном, в виде безмассовых частиц. Кстати, с сохранением энергии-импульса и электрического заряда все в порядке.

Существование черных дыр демонстрирует также и нестроенность квантовой механики. Как известно, в квантовой теории используется понятие состояний. С учетом эффекта испарения из вакуума могут рождаться виртуальные черные дыры. Если задать чистое состояние в некоторый момент времени на *начальной* гиперповерхности Коши (например, вакуум), то часть

информации о системе может с течением времени попасть внутрь черной дыры и потеряться после ее испарения. Тогда на *конечной* гиперповерхности Коши состояние системы будет другим. Это означает, что возможны такие флуктуации вакуума, когда чистое состояние перестает быть чистым: происходит потеря когерентности. К счастью для отцов-основателей квантовой механики, в те времена еще не было известно об испарении черных дыр. По-видимому, виртуальные черные дыры не должны «портить» чистые состояния.

2.7. Гравитационное линзирование

Гравитационное линзирование — это способность гравитационного поля фокусировать свет. Понятно, что не всякое гравитирующее тело может быть эффективной гравитационной линзой.

При заданном положении наблюдателя относительно линзы способность к линзированию определяется размером и массой линзы. Реальные протяженные объекты во Вселенной отклоняют свет на очень малые углы, однако и расстояния до них велики. Поэтому прицельные параметры лучей света часто оказываются больше размера необходимой для линзирования массы. В наблюдательных данных эффект гравитационного линзирования присутствует повсеместно и широко используется для решения различных астрофизических задач.

Начнем с пространства Минковского. Совместим ось x с направлением наблюдатель–линза, а декартовы координаты в ортогональных направлениях будем помечать двумерным вектором $\mathbf{y} = (y_1, y_2)$. Угол отклонения света элементом массы M задается выражением (2.1.5):

$$\psi = \frac{4GM}{y} \ll 1,$$

где y — минимальное расстояние от массы до луча света (прицельный параметр).

С точки зрения наблюдателя, к которому приходит свет, удобно ввести понятие плоскости линзы, ориентированной в картинной плоскости наблюдателя «0» и совмещенной с центром линзы «1». Тогда угол $\boldsymbol{\theta} = (\theta_1, \theta_2)$, под которым наблюдатель видит вектор \mathbf{y} , находится из условия

$$\mathbf{y} = \boldsymbol{\theta} R_1, \quad (2.7.1)$$

где $R_1 \equiv |x_1 - x_0|$ — расстояние до линзы. Поскольку расстояние от наблюдателя до линзы R_1 много больше прицельного пара-

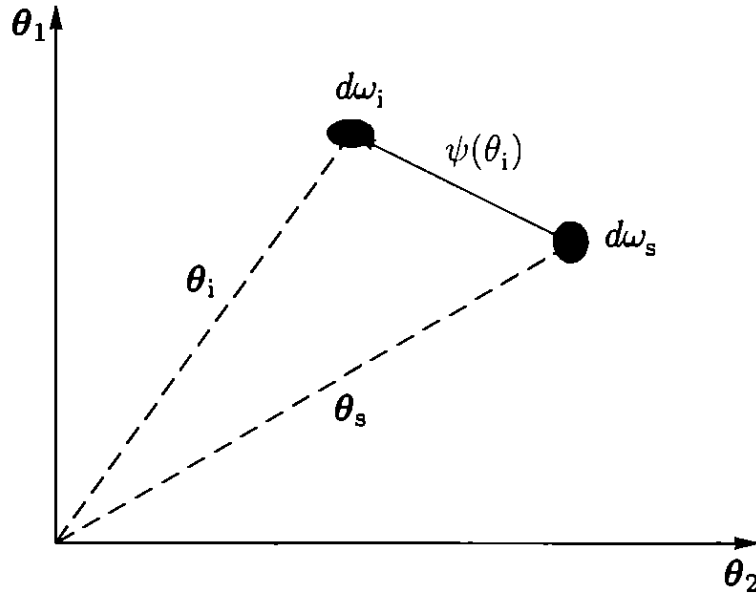


Рис. 2.7. Плоскость гравитационной линзы; $d\omega$ — элемент телесного угла метра y , то все рассматриваемые углы θ малы и полный угловой вектор отклонения света протяженной линзой равен

$$\psi(\theta) = 4G \int \frac{\sigma(y')(y - y')}{|y - y'|^2} dy', \quad (2.7.2)$$

где $\sigma(y) = \int \rho(x, y) dx$ — поверхностная плотность массы гравитационной линзы.

Выведем уравнение линзы для определения изображений источника. Для этого рассмотрим плоскость линзы в угловых координатах θ (рис. 2.7); индекс «s» относится к источнику излучения, а «i» — к его изображению. Если источник протяженный, то помимо смещения наблюдается и деформация формы источника, то есть, например, из круглого он становится овальным.

Пусть θ_s — угол, под которым был бы виден нелинзированный элемент источника, а $\theta = \theta_i$ — одно из его изображений. Вектор угла смещения элемента $\theta - \theta_s$ связан с углом отклонения луча света (2.7.2) простым тригонометрическим соотношением (рис. 2.8):

$$\theta - \theta_s = -\psi(\theta) \frac{R_{ls}}{R_s}, \quad (2.7.3)$$

где R_s — расстояние до источника, R_{ls} — расстояние между линзой и источником, знак «минус» учитывает, что угловой вектор отклонения противоположен вектору смещения.

Это и есть уравнение гравитационной линзы: корни уравнения (2.7.3) определяют все изображения $\theta_i(\theta_s)$ элемента источника. Интегрируя по источнику, мы находим карту изображений. Заметим, что вектор смещения непостоянен по источнику: у изображения имеются деформации и уярчения. Можно решить

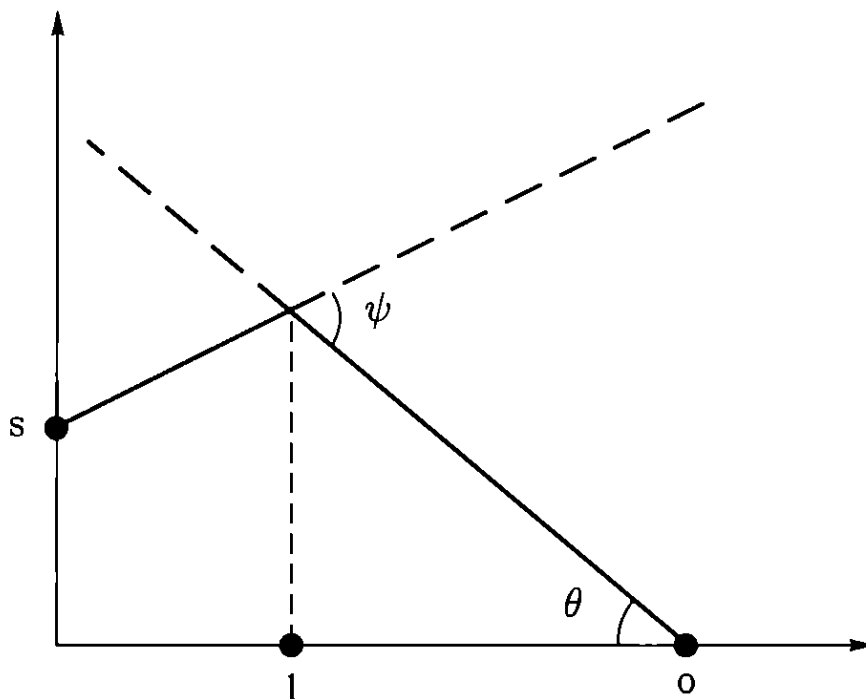


Рис. 2.8. Плоскость луча света; ψ — угол отклонения луча света, θ — угол, под которым наблюдатель «0» видит изображение источника «s»

и обратную задачу: по известным изображениям источника для заданной линзы построить функцию невозмущенного источника $\theta_s(\theta_i)$.

Важное свойство гравитационной линзы — это ее *ахроматичность*: она одинаково воздействует на свет любой длины волны. Это свойство становится ключевым, когда, располагая только кривой блеска некой удаленной звезды, нужно уметь различить явление гравитационного линзирования и возможную собственную переменность звезды, которая может быть обусловлена, например, ее двойственностью.

Гравитационная линза сохраняет поверхностную яркость источника. Это следует из того, что потоки излучения, принимаемые наблюдателем, зависят от расстояния, проходимого светом, таким же образом ($F = F_0 R^{-2}$), как и соответствующие телесные углы, под которыми видны источник или его изображения ($\omega = s_0 R^{-2}$):

$$dF = \frac{F_0}{s_0} d\omega.$$

Коэффициент усиления яркости равен отношению телесных углов изображения и источника:

$$\begin{cases} \mu_i = \frac{d\omega_i}{d\omega_s} = \left| \det \frac{\partial \theta_s}{\partial \theta_i} \right|^{-1}, \\ \mu = \sum \mu_i, \end{cases} \quad (2.7.4)$$

где μ — полное усиление яркости источника. Почему в результате линзирования яркость источника увеличивается? Потому что свет собирается с большей поверхности источника: к нам приходят лучи, которые без линзирования прошли бы мимо нас.

Рассмотрим подробнее эффект линзирования на примере аксиально-симметричной линзы. В такой линзе гравитирующая масса зависит от модуля прицельного параметра y . Тогда угол отклонения и коэффициент усиления яркости могут быть записаны в общем виде:

$$\psi(y) = \frac{4GM(y)}{y}, \quad \mu_i = \frac{\theta_i d\theta_i}{\theta_s d\theta_s}. \quad (2.7.5)$$

Что такое способность массы произвести линзирование? Предположим, что заданы положения источника излучения, предполагаемой линзы и наблюдателя, и все они находятся на одной линии (аксиальная симметрия). Будет ли изображение источника линзировано данной массой, имеющей определенный размер? Ответ на этот вопрос зависит от того, будет ли луч света, проходящий на расстоянии y от линзы, в достаточной мере отклонен линзой, чтобы попасть к наблюдателю. Для существования линзированного изображения (при заданном прицельном параметре) должно выполняться неравенство (рис. 2.9)

$$\psi \geq \psi_0 \equiv \theta \frac{R_s}{R_{ls}}. \quad (2.7.6)$$

В обратном случае никакого эффекта линзирования не будет и луч пройдет мимо наблюдателя.

Условие линзирования $\psi \geq \psi_0$ в терминах поверхностной плотности линзы приобретает вид:

$$\sigma(< y) \geq \sigma_c, \quad (2.7.7)$$

то есть поверхностная плотность внутри прицельного параметра должна быть больше, чем некоторое критическое число σ_c , определяемое фактическим взаиморасположением источника, линзы и наблюдателя:

$$\sigma(< y) = \frac{M(y)}{\pi y^2}, \quad \sigma_c = \frac{R_s}{4\pi G R_l R_{ls}}. \quad (2.7.8)$$

Например, если линзой является прозрачный однородный диск с поверхностной плотностью $\sigma = \sigma_c$, то все лучи, прошедшие через линзу, достигнут наблюдателя (он окажется в фокусе).

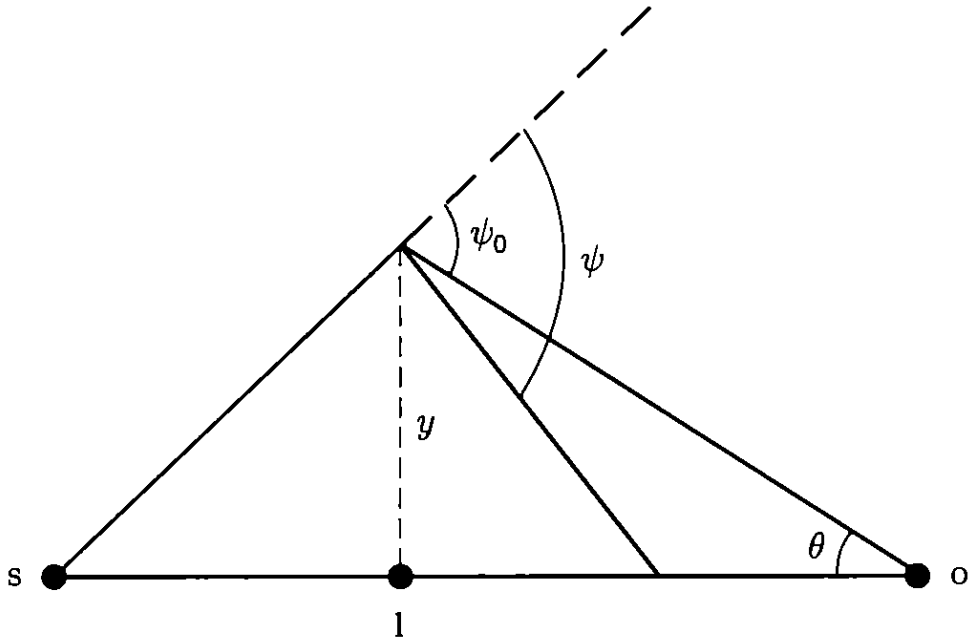


Рис. 2.9. Соотношение между углом отклонения луча света ψ и геометрическим углом ψ_0 для реализации эффекта линзирования

Условие $\sigma(< y) = \sigma_c$ фиксирует угловой радиус окружности изображения θ_{CE} , называемый углом Хвольсона–Эйнштейна:

$$\theta_{CE}^2 = \frac{4GM(\theta_{CE}R_l) R_{ls}}{R_l R_s}. \quad (2.7.9)$$

Размер θ_{CE} является характеристикой линзирующей массы и определяет ее способность к линзированию. Если линза находится посередине между источником и наблюдателем, мы получаем следующую оценку для угла и радиуса кольца Хвольсона–Эйнштейна:

$$\theta_{CE} = \sqrt{\frac{r_g}{R_l}}, \quad y_{CE} = \sqrt{r_g R_l}, \quad (2.7.10)$$

где r_g — гравитационный радиус линзы массой M .

Если источник находится на одной линии с аксиальной линзой и наблюдателем и если выполнено условие линзирования, то изображение превращается в окружность с угловым радиусом θ_{CE} . Это дает другую трактовку угла Хвольсона–Эйнштейна как угла между источником и линзой, при котором наблюдается характерный эффект гравитационного линзирования. Для этого источник должен оказаться внутри окружности Хвольсона–Эйнштейна.

Предположим, что у нас есть некий удаленный компактный источник, а линзой является звезда, которая движется в картинной плоскости. Источник смещается в картинной плоскости относительно линзы, попадая на какой-то интервал времени внутрь кольца Хвольсона–Эйнштейна этой звезды. Наблюдатель в это

время видит характерное проявление эффекта линзирования: появление двойного изображения источника с сопоставимыми по яркости компонентами (см. (2.8.4)).

Оценим вероятность линзирования фоновых источников излучения массивными объектами некоторого класса, находящимися на более близком расстоянии, чем источники. Каждая линза окружена своим кольцом Хвольсона–Эйнштейна. Полное сечение линзирования — это сумма всех площадей окружностей Хвольсона–Эйнштейна. Каждая площадь пропорциональна массе линзы. Таким образом, вероятность (частота) линзирования пропорциональна полной массе всех объектов данного класса. Если окружности перекрываются на небесной сфере, то любой источник испытывает гравитационное линзирование.

Все результаты для пространства Минковского легко обобщить на случай модели Фридмана, поскольку световые геодезические конформно-инвариантны, а фридмановская геометрия, как известно, конформна плоскому пространству-времени:

$$dt^2 - a^2 dx^2 = a^2 (d\eta^2 - dx^2), \quad (2.7.11)$$

где $\mathbf{x} = (R, \mathbf{y})$, \mathbf{y} — двухкомпонентный вектор в картинной плоскости линзы (см. (2.7.1)), а роль времени Минковского играет конформное время η . Сопутствующее расстояние между наблюдателем и объектом связано с красным смещением z точки наблюдения:

$$R = R(z) = \eta_0 - \eta = \int_0^z \frac{dz}{H(z)}, \quad (2.7.12)$$

а определение угловых размеров (2.7.1) и уравнение линзы (2.7.3) остаются справедливыми в общем случае с той лишь оговоркой, что

$$\psi(\theta) = \frac{4GM}{r},$$

где $r = a_1 y = \frac{\theta R_1}{1 + z_1}$ — минимальное физическое расстояние от элемента массы до луча зрения. Отсюда получаем угол и радиус Хвольсона–Эйнштейна для аксиальной линзы массой $M(r)$ в модели Фридмана:

$$\theta_{CE}^2 = \frac{4GM(r_{CE}) R_{ls}(1 + z_1)}{R_1 R_s}, \quad r_{CE} = \frac{\theta_{CE} R_1}{1 + z_1}. \quad (2.7.13)$$

Эта формула приобретает «стандартный» вид, если пользоваться не сопутствующим геодезическим расстоянием R , а *угловым расстоянием* D , называемым также *расстоянием по*

угловому диаметру. Это расстояние до объекта с физическим размером r , имеющим красное смещение z и наблюдаемого под углом θ :

$$D = D(z) \equiv \frac{r}{\theta} = aR \equiv \frac{R}{1+z}, \quad (2.7.14)$$

причем соотношение между D , r и θ определено формально, как в геометрии Минковского. Тогда формулу (2.7.13) можно переписать следующим образом:

$$\theta_{CE}^2 = \frac{4GM(\theta_{CE}D_1) D_{ls}}{D_1 D_s}, \quad r_{CE} = \theta_{CE} D_1, \quad (2.7.15)$$

где $D_{ls} = \frac{R_{ls}}{1+z_s}$ — расстояние по угловому диаметру, под которым наблюдатель, находящийся в окрестности линзы, видит объекты в плоскости источника.

Характеристики линзирования на различных объектах Вселенной приведены в табл. 2.1. В первой колонке содержится название объекта, во второй — его масса, выраженная в массах

Таблица 2.1. Характеристики линзирования на различных объектах во Вселенной

Линза	M [M_{\odot}]	R_1 [пк]	r [пк]	$\sigma(< r)$ [σ_c]	θ_{CE} ["]	r_{CE} [пк]
звезда в Галактике	1	10^4	$2 \cdot 10^{-8}$	$2 \cdot 10^6$	$6 \cdot 10^{-4}$	$3 \cdot 10^{-5}$
звезда в галактике	1	10^9	$2 \cdot 10^{-8}$	$2 \cdot 10^{11}$	$2 \cdot 10^{-6}$	10^{-2}
галактика	10^{12}	10^9	$5 \cdot 10^3$	4	2	10^4
скопление галактик	10^{14}	10^9	10^5	1	20	10^5

Солнца, в третьей — характерное расстояние до линзы в парсеках, в четвертой — физический размер линзы в парсеках, который также можно использовать в качестве оценки прицельного параметра. В пятой колонке указано отношение поверхностной плотности линзы к критическому значению (см. (2.7.8)). То, что это отношение для звезд больше единицы с колоссальным запасом, означает, что звезды являются эффективными линзами. В шестой колонке указан угол Хвольсона–Эйнштейна, выраженный в угловых секундах. Для звезд он очень мал, то есть звезда является эффективной линзой, но попасть на луч зрения звезды очень трудно. В седьмой колонке приведены радиусы Хвольсона–Эйнштейна, выраженные в парсеках. Для звезд этот радиус больше размера звезды, то есть звезда и по этому параметру

является эффективной линзой. Но если мы подсчитаем долю полной массы звезд во Вселенной, что даст нам представление об эффективности линзирования именно на звездах, то она мала.

Что касается линзирования на галактиках и скоплениях галактик, то здесь ситуация иная. Основное отличие состоит в том, что поверхностная плотность галактик и скоплений довольно мала — гораздо меньше, чем у звезд, хотя все-таки достаточна для того, чтобы произвести эффект линзирования. Для космологического источника и линзы с красными смещениями $z_s \simeq 2$ и $z_l \simeq 0,5$ характерная величина критической поверхностной плотности $\sigma_c \simeq 1 \text{ г} \cdot \text{см}^{-2}$. Другое важное отличие космологических линз — гораздо большие углы Хвольсона–Эйнштейна (по сравнению с звездами) из-за гораздо большей массы, в результате чего открываются прекрасные возможности для наблюдения линзирования на массивных линзах как в радио-, так и в оптическом диапазоне. Ограничения по прицельному параметру в этом случае также велики. При сопоставлении размера галактики с прицельным параметром становится ясно, что основной эффект производит ее центральная часть с размером несколько килопарсек. Это же замечание относится и к линзированию на скоплении галактик, где основной эффект дает ядро размером не более 100 кпк.

Так устроена гравитация и тела во Вселенной. Удивительно, что гравитационное линзирование реализуется на пределе для галактик и скоплений и оказывается типичным явлением в космологии.

Эффект линзирования широко используется для определения космологических параметров.

Различают сильное и слабое гравитационное линзирование. Рисунок 2.10 иллюстрирует сильное линзирование: далекие источники света (галактики), попавшие на луч зрения в направлении линзы (скопления галактик), расщеплены на множество изображений, имеющих характерную форму изогнутых дужек. Рисунок 2.11 демонстрирует другой пример сильного линзирования: компактный источник конечного размера проецируется на центр почти сферической линзы, в результате чего получилось изображение в виде кольца Хвольсона–Эйнштейна.

Сильное линзирование (источник проецируется внутрь окружности Хвольсона–Эйнштейна) можно использовать для определения постоянной Хаббла. Для этого необходимо измерить угловое расстояние между изображениями и знать распределение материи в линзе ($\Delta\theta \simeq 2\theta_{CE} \rightarrow M/D_1 \rightarrow MH_0$, см. раздел 2.8). Получающиеся значения H_0 сопоставимы по точности с изме-

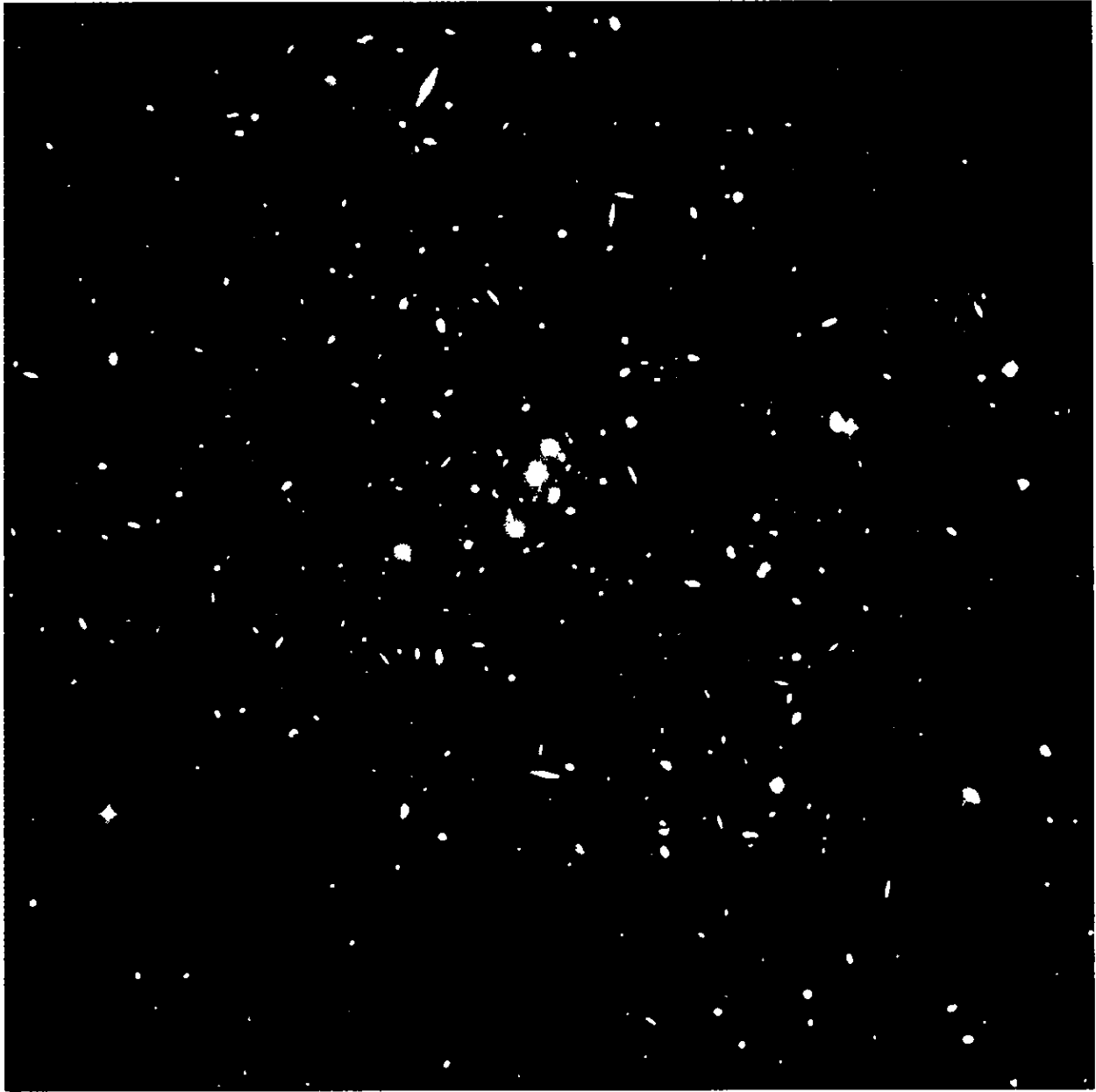


Рис. 2.10. Пример сильного гравитационного линзирования. Голубые дужки — изображения далеких галактик, искаженные при линзировании на близком скоплении галактик C10024+1654 (hubble-site.org/gallery/album/)

рениями локальной хаббловской постоянной с помощью цефеид и сверхновых.

При слабом линзировании источник находится вне колец Хвольсона–Эйнштейна и его смещение и деформация изображения невелики. Подавляющее большинство далеких галактик слабо линзированы близлежащей структурой.

Эффект слабого гравитационного линзирования позволяет по искажению формы фоновых источников восстановить пространственное распределение материи как в отдельных линзирующих системах (например, в скоплениях галактик), так и на больших масштабах (сверхскопления, войды и др., см. главу 12). Рисунок 2.12 демонстрирует распределение поверхностной плотности

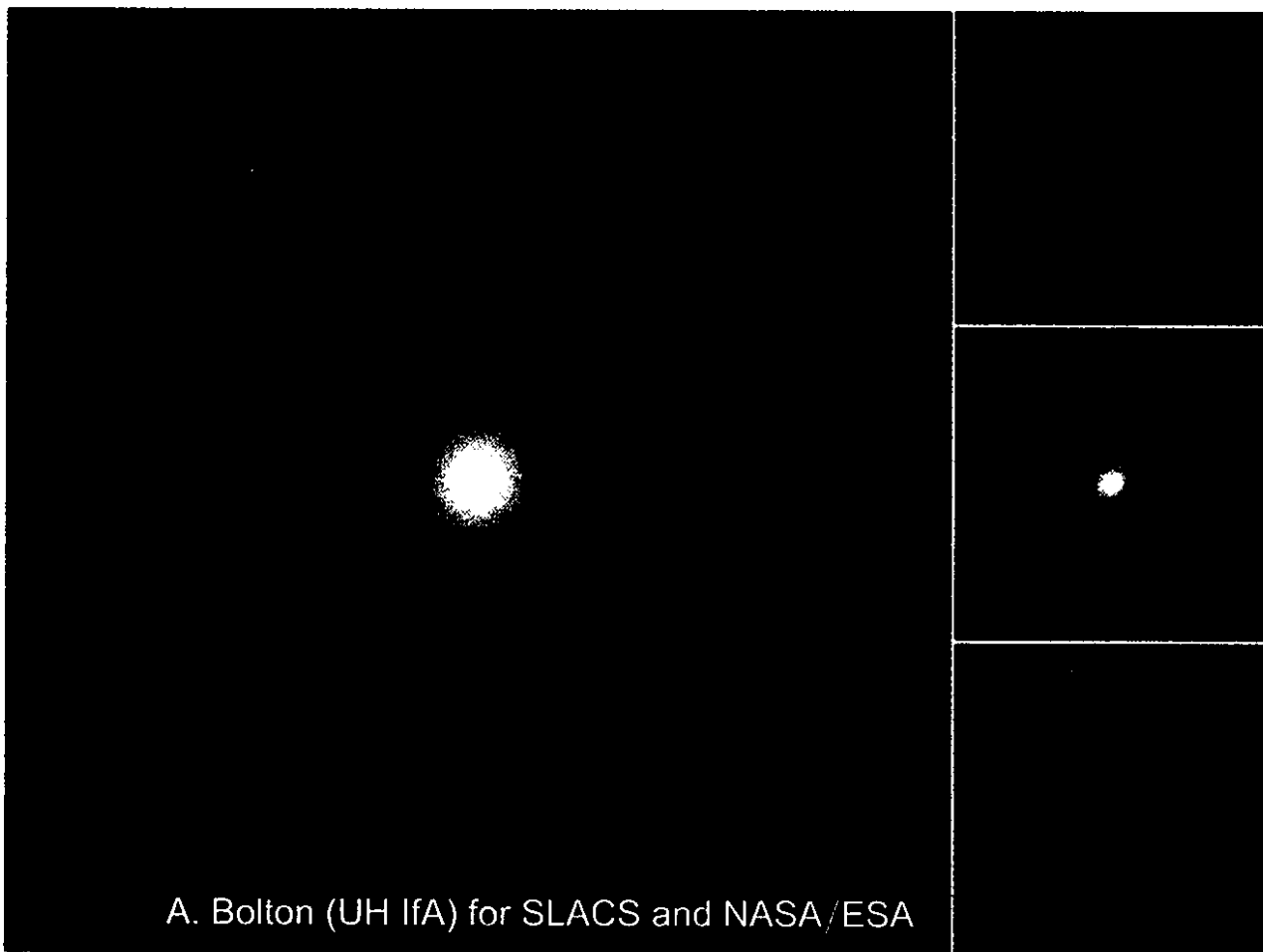


Рис. 2.11. Кольцо Хвольсона–Эйнштейна SDSS J1430. В колонке справа сверху вниз — восстановленный источник, линза, наблюдаемое изображение источника [Bolton, 2008]

темной материи в одном из скоплений галактик, полученное по данным анизотропии деформации многочисленных галактик фона, проектирующихся в область локализации линзирующей массы. Поверхностная плотность массы связана двумерным преобразованием Лапласа с гравитационным потенциалом в плоскости линзы $\phi = \phi(\mathbf{y})$, который восстанавливается из наблюдаемого сдвигового тензора приливных сил $\phi_{,ab}$ по гравитационному искажению галактик путем усреднения десятков локальных изображений для подавления собственной случайно ориентированной эллиптичности галактик поля.

Другой пример слабого линзирования представлен на рис. 2.13. Здесь с помощью линзирования получено распределение невидимой темной материи в двух скоплениях галактик (известных как скопление «пуля») после их прямого столкновения несколько млрд лет назад. Мы видим, что темная материя одного скопления свободно прошла сквозь темную материю другого, и гравитирующие центральные области практически не искажены, однако газ оказался выбит из гравитирующих центров в ходе прямого столкновения и теперь возвращается в свои гравита-

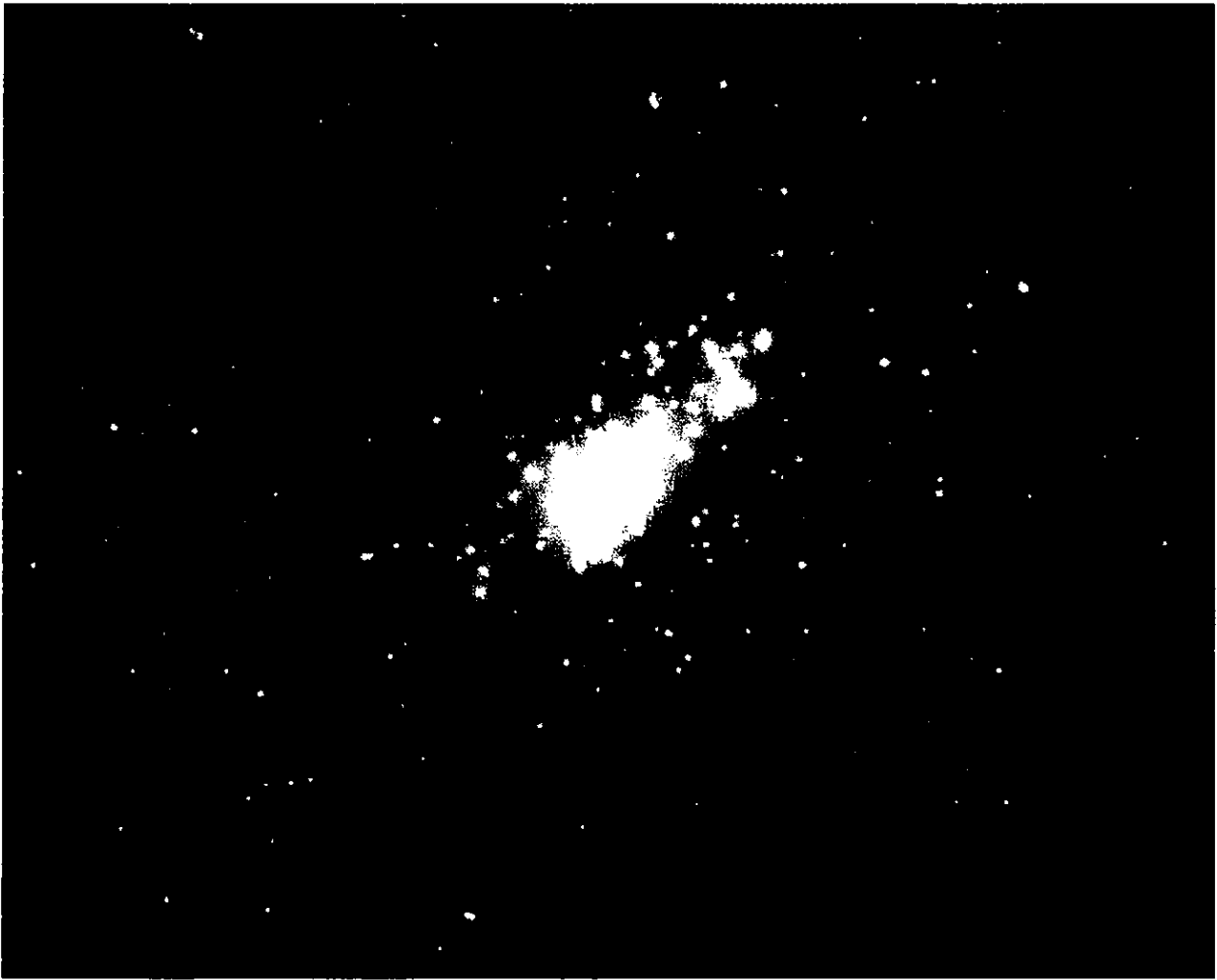


Рис. 2.12. Скопление галактик C10024+1654. Голубым цветом показано распределение темной материи, восстановленное по слабому гравитационному линзированию (antwgr.gsfc.nasa.gov/apod/ap030814.html)

ционные ямы в сверхзвуковом режиме падения. Этот пример свидетельствует о том, что темная материя состоит из нерелятивистских бесстолкновительных частиц небарионной природы, а не является формой модифицированной гравитации или зеркальной материей, близкой по свойствам нашему «левому» миру.

Наблюдения слабой деформации изображений далеких галактик из-за приливного воздействия гравитационного поля близлежащей крупномасштабной структуры Вселенной дают бесценную информацию о распределении неоднородной части гравитационного потенциала, что позволяет «взвесить» ту долю массы Вселенной, которая входит в ее структуру. Исследуя источники с разными красными смещениями, мы можем измерять эволюцию структуры во времени и определять точные параметры космологической модели (например, эволюцию темной энергии, см. раздел 11.3). Эти и другие методы исследования, основанные на эффекте гравитационного линзирования, входят в фундамент точной космологии XXI века.

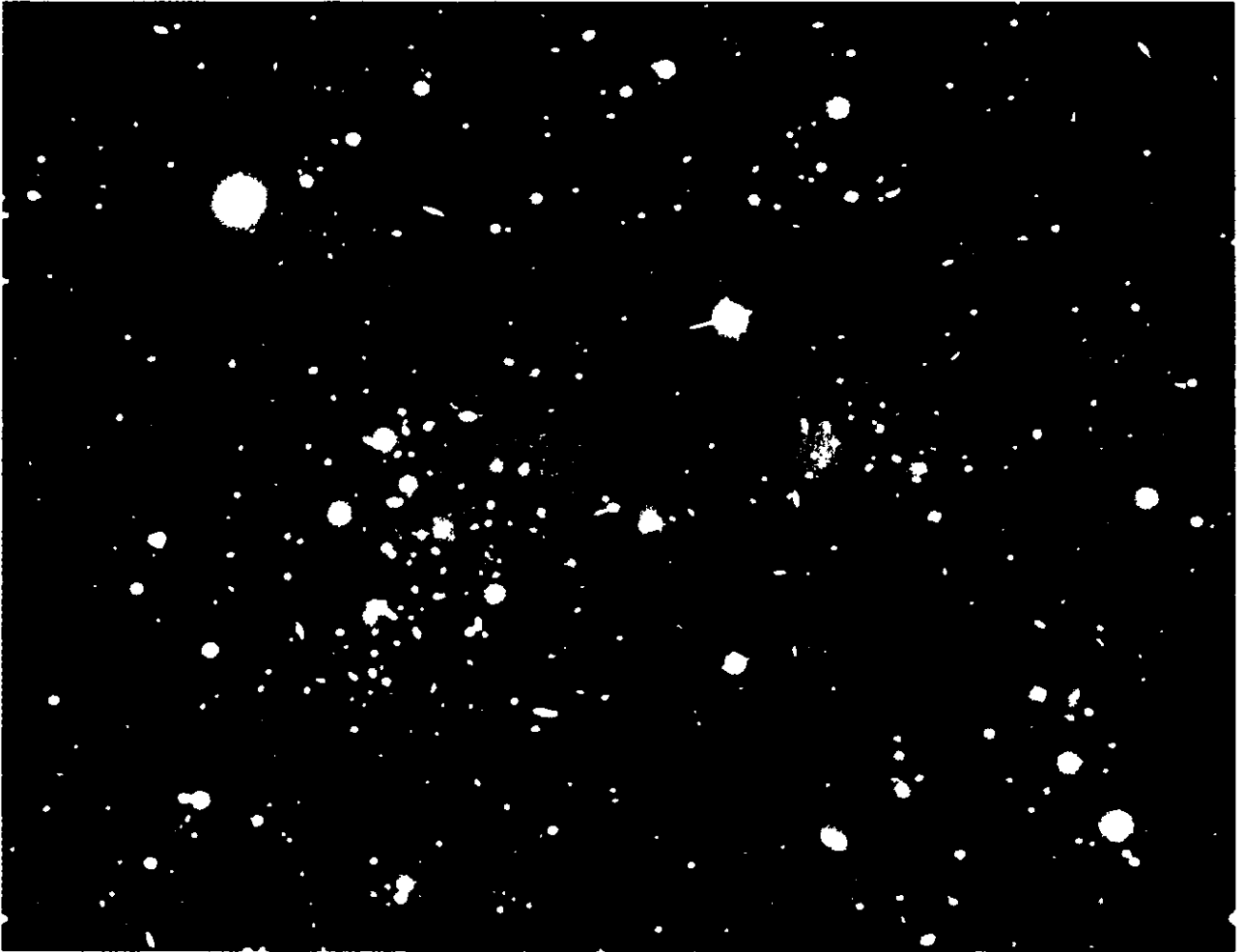


Рис. 2.13. Пример слабого гравитационного линзирования. Скопление галактик 1E 0657–56. Голубым цветом выделено восстановленное распределение темной материи, розовым — распределение горячего газа, полученное по рентгеновскому изображению скопления (antwrp.gsfc.nasa.gov/apod/ap080823.html)

2.8. Гравитационные линзы

Рассмотрим несколько примеров аксиально-симметричных линз. Для краткости ограничимся моделями линз в плоском пространстве-времени. Для обобщений на случай фридмановской геометрии следует использовать формулы (2.7.12)–(2.7.15).

Начнем с точечной массы. Уравнение линзы в этом случае принимает вид:

$$\theta - \theta_s = \frac{\theta_{CE}^2}{\theta}, \quad \theta_{CE}^2 = \frac{2r_g R_{ls}}{R_l R_s}, \quad (2.8.1)$$

и имеет два решения, а значит, два изображения — A и B :

$$\theta_{A,B} = \frac{1}{2} (\theta_s \pm \Delta\theta), \quad \Delta\theta = \sqrt{\theta_s^2 + 4\theta_{CE}^2}. \quad (2.8.2)$$

Соответствующие коэффициенты уярчения изображений задаются выражениями:

$$\mu_{A,B} = \frac{1}{4} \left(\frac{\Delta\theta}{\theta_s} + \frac{\theta_s}{\Delta\theta} \pm 2 \right). \quad (2.8.3)$$

Если их сложить, то суммарное уярчение составит

$$\mu = \frac{1}{2} \left(\frac{\Delta\theta}{\theta_s} + \frac{\theta_s}{\Delta\theta} \right) \rightarrow \frac{2\theta_{CE}}{d\theta_s}. \quad (2.8.4)$$

Источник, оказавшийся внутри кольца Хвольсона–Эйнштейна точечной линзы, имеет выигрыш в яркости $\mu \geq 1,34$. При $\theta_s \rightarrow 0$ яркость изображений увеличивается, но расходимости нет, поскольку θ_s не может быть меньше половины размера самого источника $d\theta_s$.

Гравитационное линзирование, связанное с изменением видимой яркости источников во времени при их прохождении внутри кольца Хвольсона–Эйнштейна (при этом компоненты изображений, как правило, не разрешаются и их размеры составляют микросекунды дуги), называется *микрولينзированием*¹⁾. Примером является микрولينзирование звезд Галактики или ближайших к нам галактик, таких как Туманность Андромеды, Большое и Малое Магеллановы Облака и др., компактными темными объектами гало звездных и субзвездных масс (см. табл. 2.1).

Для набора статистики производится одновременное патрулирование большого числа звезд. Если кривая яркости какой-то из звезд похожа на то, что предполагается в случае гравитационного микрولينзирования, то такую звезду исследуют на собственную переменность. В установлении события микрولينзирования очень помогают ахроматичность линзированного потока, а также то, что собственная физическая переменность звезды — явление повторяющееся, тогда как повторяющееся гравитационное линзирование — это редкое явление.

Такие эксперименты уже проводились и привели к важным результатам: 80% темной материи гало Галактики не может состоять из компактных объектов гало, имеющих массы в диапазоне от массы Луны до характерных масс звезд (рис. 2.14). В противном случае частота событий микрولينзирования оказалась бы больше, чем наблюдаемая в эксперименте.

¹⁾ Более подробная классификация учитывает разные интервалы масс линз и углов расщеплений. От этого зависят характерные времена уярчений изображений.

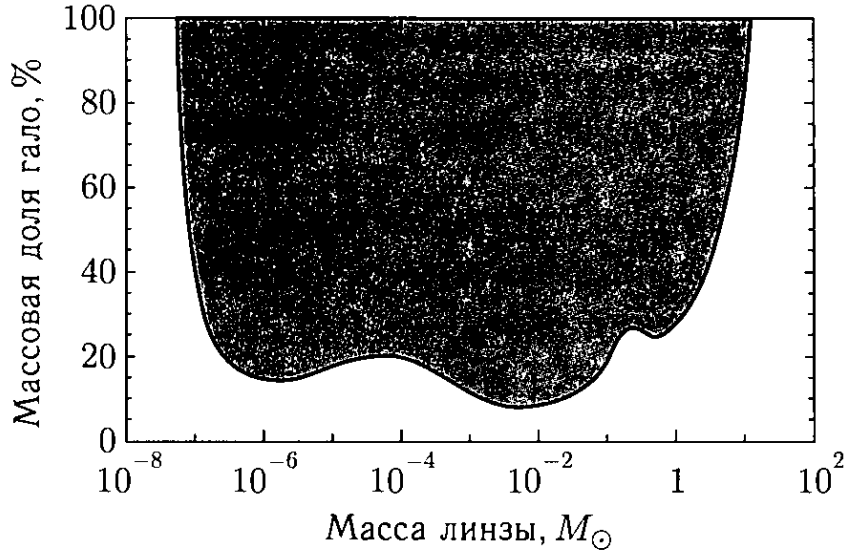


Рис. 2.14. Массовая доля гало Галактики в форме компактных объектов с массой M . Темное поле — запрещенная область по частоте событий микролинзирования (по данным [Afonso et al., 2003]). См. также <http://eros2.in2p3.fr>

Одной из наиболее часто используемых моделей линзы является модель изотермической сферы, для которой

$$M(y) \sim y. \quad (2.8.5)$$

Плотность материи и скорость вращения частиц в такой линзе ведут себя как

$$\rho(r) \sim r^{-2}, \quad v^2 = \frac{GM}{r} = \text{const.}$$

Если эти соотношения подставить в выражения для углов отклонения и Хвольсона–Эйнштейна, то получим

$$\psi(y) = \psi_0 = 4v^2, \quad \theta_{CE} = \psi_0 \frac{R_{ls}}{R_s},$$

то есть угол отклонения не зависит от прицельного параметра. В этих условиях при $\theta_s \leq \theta_{CE}$ происходит раздвоение изображения источника:

$$\theta_{A,B} = \theta_s \pm \theta_{CE}, \quad (2.8.6)$$

и коэффициенты уярчения равны

$$\mu_{A,B} = \frac{\theta_{CE}}{\theta_s} \pm 1, \quad \mu = 2 \frac{\theta_{CE}}{\theta_s} \rightarrow \frac{4\theta_{CE}}{d\theta_s}. \quad (2.8.7)$$

Таким образом, изотермическая сфера в два раза эффективнее усиливает яркость источника, чем точечная линза.

Последний пример модели гравитационной линзы, который мы рассмотрим, — это однородный диск. В этом случае

$$\sigma = \text{const}, \quad M(y) = \pi\sigma y^2, \quad \psi(y) = 4\pi G\sigma y, \quad (2.8.8)$$

то есть чем больше прицельный параметр, тем больше угол отклонения. Однородный диск является собирающей линзой с фокальной длиной $(4\pi G\sigma_0)^{-1}$. Уравнение линзы имеет только одно решение:

$$\theta_A = \frac{\theta_s}{1 - \kappa}, \quad \kappa = \frac{\sigma}{\sigma_c}, \quad (2.8.9)$$

и коэффициент уярчения равен

$$\mu_A = \frac{1}{(1 - \kappa)^2}. \quad (2.8.10)$$

В рассмотренных нами примерах линзы с расходящейся поверхностной плотностью в центре создавали два изображения, а линза с конечной плотностью (однородный диск) — одно. Справедлива следующая общая теорема: сингулярная линза образует четное число изображений (каждое изображение имеет двойника), а несингулярная — нечетное.

Если гравитационное линзирование приводит к появлению нескольких изображений, возникает еще один эффект — временная задержка между изображениями. Характерные величины временной задержки для космологических объектов — месяцы и годы. По величине задержки можно определять постоянную Хаббла. Анализируя данные, полученные в оптическом диапазоне, исследователи получают значения постоянной Хаббла в интервале от 60 до 70 км·с⁻¹·Мпк⁻¹; неопределенность в эти результаты вносит неточность знаний моделей линз.

Рассмотрим источник света «s» и два его изображения A и B, создаваемые точечной линзой. Рассчитаем интервал времени между приходами импульса излучения от изображений A и B. Рисунок 2.8 иллюстрирует пространственную траекторию светового импульса, приходящего от одного из изображений. Используя интеграл движения нулевой геодезической в статическом гравитационном поле

$$p_0 = 1 = g_{00} \frac{dt}{d\lambda},$$

получаем выражение для временной задержки в виде интеграла по траекториям:

$$\Delta t \Big|_B^A = \int_s^0 \frac{d\lambda}{g_{00}} \Big|_B^A = R_s \Big|_B^A + r_g \int_s^0 \frac{d\lambda}{r} \Big|_B^A. \quad (2.8.11)$$

Первый член суммы есть разница геометрических путей фотонов из A и B в плоском пространстве Минковского ($g_{00} = 1$), а вто-

рой член учитывает релятивистскую задержку времени между приходами сигналов из-за искривления пространства-времени (эффект Шапиро). Оба члена суммы первого порядка малости по углам отклонений.

Считая, что свет распространяется в геометрии Минковского, геометрическую задержку можно вычислить следующим образом:

$$R_s \Big|_B^A = \frac{R_{ls}}{\cos(\psi - \theta)} \Big|_B^A + \frac{R_l}{\cos \theta} \Big|_B^A = r_g \left(\frac{\theta_B}{\theta_A} - \frac{\theta_A}{\theta_B} \right),$$

где

$$\psi \theta = \frac{2r_g}{R_l}, \quad \theta_A \theta_B = \theta_{CE}^2 = \frac{2r_g R_{ls}}{R_s R_l}.$$

Задержка Шапиро связана с неоднородностью хода времени вдоль луча зрения (вблизи линзы ход времени замедляется) и с увеличением геометрического пути света из-за искривления пространства вблизи тяготеющего тела. Эту величину можно вычислить следующим образом ($\lambda \equiv y \operatorname{sh} \alpha$):

$$\int \frac{d\lambda}{\sqrt{\lambda^2 + y^2}} = \alpha_{ls} + \alpha_{lo} = \ln \frac{4R_{ls}R_l}{y^2} \Big|_B^A = 2 \ln \left| \frac{\theta_B}{\theta_A} \right|. \quad (2.8.12)$$

В результате получаем полный интервал времени между событиями приходов импульсов от двух изображений источника для точечной линзы:

$$\Delta t \Big|_B^A = r_g \left(\frac{\theta_B}{\theta_A} - \frac{\theta_A}{\theta_B} + 2 \ln \left| \frac{\theta_B}{\theta_A} \right| \right). \quad (2.8.13)$$

При $\theta_s \ll 2\theta_{CE}$ (см. (2.8.2)) $\theta_A^2 - \theta_B^2 \simeq 2\theta_s \theta_{CE}$ и

$$\Delta t \Big|_A^B \simeq 4r_g \frac{\theta_s}{\theta_{CE}}. \quad (2.8.14)$$

ГРАВИТАЦИОННОЕ УДЕРЖАНИЕ МАТЕРИИ

Важную роль в наблюдательной космологии играют гравитационно-связанные системы, состоящие из нерелятивистских частиц обычного вещества (барионов) и темной материи (бесстолкновительных частиц небарионной природы). Это соответственно звезды, звездные скопления и центральные области галактик, а также гало темной материи галактик, групп и скоплений галактик. В силу того, что частицы обычного вещества активно взаимодействуют с излучением и сами с собой, они легко разогреваются при сжатии материи и только малая их доля (менее 10%) участвует в образовании звезд. Основная масса барионов находится в состоянии горячего разреженного газа, взаимодействующего с ультрафиолетовым излучением звезд и активных ядер галактик и подогреваемого ударными волнами и взрывами сверхновых. Напротив, темная материя (ТМ) практически вся сосредоточена в гравитационно-связанных гало широкого диапазона масс: от карликовых систем, не способных удержать своим гравитационным полем барионную компоненту, до богатых скоплений галактик с массами порядка $10^{15} M_{\odot}$. Из-за того, что космологическая плотность ТМ в пять раз больше плотности барионного вещества, ТМ сыграла определяющую роль в образовании нелинейных структур во Вселенной.

3.1. Горячие звезды

Рассмотрим системы из барионов, которые удерживаются собственным гравитационным полем. Многие из окружающих нас небесных тел — звезды и планеты — являются самогравитирующими системами такого рода.

Характерную массу звезды можно выразить через массу протона m_p и константу гравитации $G = M_P^{-2}$:

$$M \sim \frac{M_P^3}{m_p^2} \simeq 1,8 M_{\odot}. \quad (3.1.1)$$

Какой смысл скрыт в этой формуле? Почему масса звезд столь велика?

Чтобы ответить на эти вопросы, попробуем собрать протоны вместе. Возьмем два протона и сблизим их. Потом отпустим. Сила Кулона оттолкнет их один от другого — гравитационное притяжение протонов друг к другу слишком мало, чтобы протоны остались вместе. Но если мы будем добавлять еще протоны, их число будет становится все больше и больше, пока, наконец, их масса не станет достаточной для того, чтобы собственное гравитационное поле могло удержать их вместе.

Протоны в этой куче будут иметь среднеквадратичную скорость, определяемую из условия равенства энергии кулоновского взаимодействия и кинетической энергии частицы в куче:

$$\langle v_p^2 \rangle \simeq \frac{e^2}{m_p r},$$

где e — кулоновский заряд протона, r — среднее расстояние между протонами в куче. Если эта скорость окажется меньше второй космической скорости $\sqrt{GM/R}$, определяемой полной массой протонов и размером кучи, то они не смогут разлететься. Конечно, мы рассматриваем электронейтральную плазму, однако ее масса определяется, главным образом, массой протонов, поскольку масса электронов в 2000 раз меньше.

Сколько нужно собрать протонов вместе, чтобы такая куча не разлетелась?

Условие удержания водородной плазмы собственным гравитационным полем имеет вид:

$$T \simeq m_p \langle v_p^2 \rangle \simeq \frac{e^2}{r} \leq \frac{GMm_p}{R} \equiv \frac{\alpha_N}{r}, \quad (3.1.2)$$

где $M = m_p N_p$ и $R = r N_p^{1/3}$ — полная масса и размер гравитационно-связанной системы, а N_p — общее число протонов в куче. С помощью безразмерного параметра задачи

$$\alpha_N = Gm_p^2 N_p^{2/3} = N_p^{2/3} \left(\frac{m_p}{M_p} \right)^2$$

условие самоудержания протонов можно переписать в виде:

$$\alpha_N \geq e^2 \simeq 10^{-2},$$

откуда следует, что

$$N_p \geq \left(\frac{M_p}{10m_p} \right)^3 \simeq 10^{54}. \quad (3.1.3)$$

Столь большое число протонов, необходимое для их гравитационного удержания, обусловлено тем, что гравитационное

взаимодействие очень слабое — константа G мала. Отношение гравитационной и кулоновской энергий взаимодействия двух протонов равно:

$$\frac{G m_p^2}{e^2} \simeq 10^{-36},$$

и именно это отношение определяет число протонов N_p .

Из (3.1.3) получаем минимальную массу гравитационно-связанной водородной плазмы:

$$M = m_p N_p = \alpha_N^{3/2} \frac{M_p^3}{m_p^2} \simeq 10^{-3} M_\odot, \quad (3.1.4)$$

что близко к массе Юпитера. Если в системе присутствуют нейтроны, то коэффициент α_N должен определяться *полным* числом барионов (протонов и нейтронов):

$$\alpha_N \equiv G m_p^2 N_b^{2/3} = N_b^{2/3} \left(\frac{m_p}{M_p} \right)^2. \quad (3.1.5)$$

Напомним, что эта оценка относится к плазме. Как мы знаем, Земля тоже самогравитирующее тело, но она состоит не из водородно-гелиевой плазмы, а содержит тяжелые химические элементы.

Для Солнца $\alpha_N \simeq 1$. Полное число барионов и средняя температура Солнца соответственно равны:

$$N_b \simeq 10^{57},$$

$$T_\odot = m_b \langle v_b^2 \rangle \simeq \frac{1}{r} \simeq \frac{N_b^{1/3}}{R_\odot} \simeq 10^8 \text{ см}^{-1} \simeq 1 \text{ кэВ}.$$

Разница между Солнцем и Юпитером заключается в разных источниках газового теплового давления. У Юпитера это кулоновское взаимодействие между частицами и энергия, выделяемая при стратификации вещества и радиоактивном распаде изотопов, а у Солнца источником тепла является энергия, выделяемая в термоядерных реакциях, идущих в недрах Солнца. На Юпитере такие реакции не идут из-за того, что его масса недостаточна для обеспечения необходимых для этого условий (ограниченные температура и плотность).

Наблюдательная функция масс звезд в Галактике аппроксимируется функцией Солпитера:

$$dn = f(m) d \ln m,$$

$$f(m) = A \cdot m^{-1,35} \exp \left(-\frac{0,3}{m} \right), \quad (3.1.6)$$

где $m \equiv M/M_{\odot}$, A — нормировочная константа. Как видно, характерная масса индивидуальных звезд Галактики порядка массы Солнца. Существование таких устойчивых и долгоживущих звезд, как Солнце, связано с условиями их образования, в первую очередь — с химическим составом дозвездного вещества, обогащенного тяжелыми элементами. Первые звезды во Вселенной образовывались из водородно-гелиевой плазмы и были совсем другими: более массивными, короткоживущими и неустойчивыми по отношению к коллапсу.

Сам факт того, что отношение масс протона и нейтронной звезды (или белого карлика или звезды типа Солнца) равно кубу отношения массы протона к планковской массе, свидетельствует о том, что имеется тесная связь между сильным и гравитационным взаимодействиями, возможно, напоминающая связь между электромагнитным и слабым взаимодействиями.

Это может означать, что сильное и гравитационное взаимодействия, ответственные за существование столь стабильных систем, как протоны и звезды, что-то «знали» друг о друге в прошлом и могли возникнуть из единого типа взаимодействия — сильно-гравитационного — таким же образом, как электромагнитное и слабое взаимодействия возникли из единого электро-слабого. Только речь идет о совершенно разных уровнях энергий *фазовых переходов*, изменивших состояние материи: если электрослабый масштаб ~ 1 ТэВ, то сильно-гравитационный переход отодвигается в область $\sim 10^{10}$ ГэВ и, возможно, еще дальше к планковской энергии.

Такой подход отличается от «традиционного» Великого объединения (предположительно при энергиях $\sim 10^{16}$ ГэВ), не затрагивающего гравитацию. На пути к столь высоким энергиям лежит «пустыня» неизведанного и доступного лишь с помощью неускорительных экспериментов (наблюдений космических лучей, активных галактических ядер и др.). Сильная гравитационная связь может характеризоваться громадным безразмерным числом $\mathcal{N}_p \simeq 10^{19}$ (см. (1.1.1)), которое обеспечивает иерархию масштабов и масс в нашем мире.

Возвратимся к теме гравитационного удержания материи.

Мы исследовали условия удержания тепловых барионов собственным гравитационным полем и убедились в том, что существование звезд и планет обусловлено балансом тепловой и гравитационной энергий: в условии (3.1.4) не входит постоянная Планка. Теперь рассмотрим гравитационно-связанные системы, состоящие из холодных барионов, и определим условия

существования релятивистских звезд — белых карликов, нейтронных и кварковых звезд, — у которых температура недостаточна для поддержания их в равновесии.

3.2. Холодные звезды

Внутреннее давление холодных барионных систем определяется законами квантовой статистики, которые запрещают двум фермионам находиться в одном состоянии. Помимо этого, нам придется учитывать релятивизм, поскольку размер гравитирующих систем близок к радиусу Шварцшильда.

Условие удержания частиц — электронов, протонов и нейтронов — гравитационным полем массы M с размером R имеет вид (см. (2.4.4)):

$$\mathcal{E}_b = \sum_{e,p,n} \frac{n_i}{n_b} \sqrt{m_i^2 + x_i^2} - \frac{m_b}{\sqrt{1 - 2GM/R}} < 0, \quad (3.2.1)$$

где \mathcal{E}_b — полная энергия частиц, приходящаяся на один барион, n_i и x_i — средняя плотность и импульс частиц соответствующих компонент вещества,

$$M = \alpha_N^{3/2} \frac{M_P^3}{m_b^2}, \quad m_b = \sum_{e,p,n} \frac{n_i}{n_b} m_i, \quad n_b = n_p + n_n \simeq \frac{N_b}{R^3}.$$

Каждая из частиц сорта $i = (e, p, n)$ локализована в масштабе $n_i^{-1/3}$. В соответствии с принципом неопределенности Гейзенберга их импульсы соответственно равны:

$$\begin{aligned} x_e = x_p &\simeq n_e^{1/3}, & x_n &\simeq n_n^{1/3}, \\ x &\equiv n_b^{1/3} \simeq \frac{N_b^{1/3}}{R} \simeq (x_e^3 + x_n^3)^{1/3} \end{aligned} \quad (3.2.2)$$

(здесь мы учли, что $n_e = n_p$ в силу электронейтральности). Будем искать решения уравнений (3.2.1) и (3.2.2) как функцию полной энергии частиц на один барион $\mathcal{E}_b(x)$, зависящую от среднего импульса частиц x в системе.

Пусть x меньше разницы между массами нейтрона и протона. Тогда в системе отсутствуют свободные нейтроны, поскольку им выгодно распасться на протоны и электроны (и антинейтрино, которые покидают звезду будучи релятивистскими частицами и обладая малым сечением взаимодействия). Таким образом, при

$$x = x_e < \Delta \equiv m_n - m_p \simeq 1,3 \text{ МэВ}, \quad x_n \simeq 0,$$

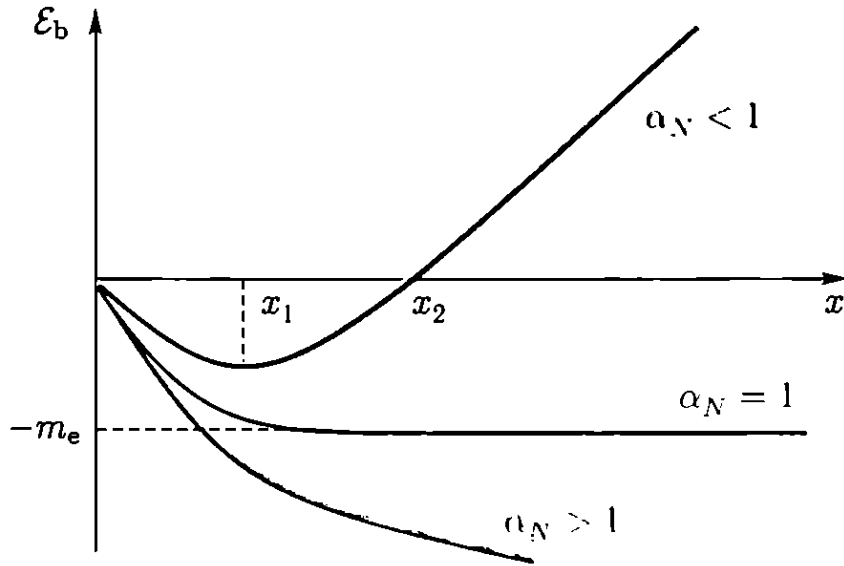


Рис. 3.1. Функция $\mathcal{E}_b(x)$ при $x < \Delta$ для разных значений параметра α_N . Точка минимума синей линии соответствует состоянию белого карлика

условие связанной конфигурации принимает вид:

$$\mathcal{E}_b(x) = -\alpha_N x + \sqrt{m_e^2 + x^2} - m_e < 0. \quad (3.2.3)$$

В этом выражении учитывается релятивизм электронов в силу малости их массы, однако импульсом протонов по сравнению с их массой можно пренебречь. Гравитационный потенциал написан в ньютоновском пределе: белый карлик — это еще слабо релятивистская звезда.

Функция $\mathcal{E}_b(x)$ для разных значений α_N представлена на рис. 3.1. Устойчивое состояние системы реализуется только при $\alpha_N < 1$ и соответствует точке $x = x_1$ локального минимума функции \mathcal{E}_b :

$$x_1 = \frac{\alpha_N m_e}{\sqrt{1 - \alpha_N^2}}, \quad x_2 = \frac{2\alpha_N m_e}{1 - \alpha_N^2} < \Delta. \quad (3.2.4)$$

В точке $x = x_2$: $\mathcal{E}_b(x_2) = 0$. Из условия существования устойчивого минимума в области отрицательных энергий $x_2 < \Delta$ получаем более точное ограничение: $\alpha_N < 0,7$.

Итак, мы нашли значение x устойчивой по отношению к малым возмущениям конфигурации частиц, а значит можем вычислить массу, размер и другие параметры звезды. Как уже было сказано, такие звезды называются белыми карликами. В реальном белом карлике присутствуют тяжелые элементы, поэтому связь α_N с массой оказывается более жесткой, чем в соотношении (3.1.4):

$$\alpha_N < 0,7 \quad \rightarrow \quad M_{WD} = M_1 \leq 1,4 M_{\odot}. \quad (3.2.5)$$

Белый карлик с массой, равной солнечной, имеет радиус порядка 10^3 км.

Таким образом, сделанная нами качественная оценка очень хорошо согласуется с наблюдениями. Верхний предел массы белого карлика (3.2.5) называется *чандрасекаровским пределом*. Белые карлики с массами выше чандрасекаровского предела существовать не могут, поскольку давление вырожденных электронов недостаточно для того, чтобы противостоять гравитационному сжатию звезды.

Перейдем к несколько большим значения x :

$$\Delta < x < \ell_n^{-1} \simeq 100 \text{ МэВ.} \quad (3.2.6)$$

При этих энергиях более выгодным оказывается образование свободных устойчивых нейтронов: они не могут распасться обратно на электроны и протоны из-за занятых уровней релятивистских электронов. Верхнее ограничение импульса в (3.2.6) соответствует условию существования нейтрона как частицы. Размер нейтрона $\ell_n \simeq 10^{-13}$ см, а его плотность $\rho_n = 2 \cdot 10^{14}$ г·см⁻³. Соответственно, приходящийся на одну частицу объем не может быть меньше объема этой частицы. При $x_n \sim 100$ МэВ нейтроны плотно упакованы, а при дальнейшем повышении энергии они «раздавливаются» и превращаются в кварк-глюонную плазму.

Равновесное состояние нейтронов определяется энергетическим условием:

$$x > \Delta: \quad x_e + \sqrt{m_p^2 + x_e^2} = \sqrt{m_n^2 + x_n^2}, \quad (3.2.7)$$

решая которое, получаем импульсы и, следовательно, концентрации электронов и нейтронов в системе:

$$x_e \simeq \Delta + \frac{x_n^2}{2m_n}, \quad x_n \simeq x \cdot \Theta(x - \Delta), \quad (3.2.8)$$

где $\Theta(x)$ — функция Хевисайда ($\Theta(x \leq 0) = 0$, $\Theta(x > 0) = 1$). Как мы видим, появление нейтронов имеет пороговый характер.

На рис. 3.2 показано, как уменьшается доля электронов x_e/x в системе с ростом энергии. Минимум достигается при $x = \sqrt{2m_n\Delta} \simeq 50$ МэВ. Вертикальная черная линия на рисунке обозначает границу применимости наших уравнений (см. (3.2.6)).

При исчезновении электронов из системы состояние белого карлика невозможно. Система теряет устойчивость и сжимается,

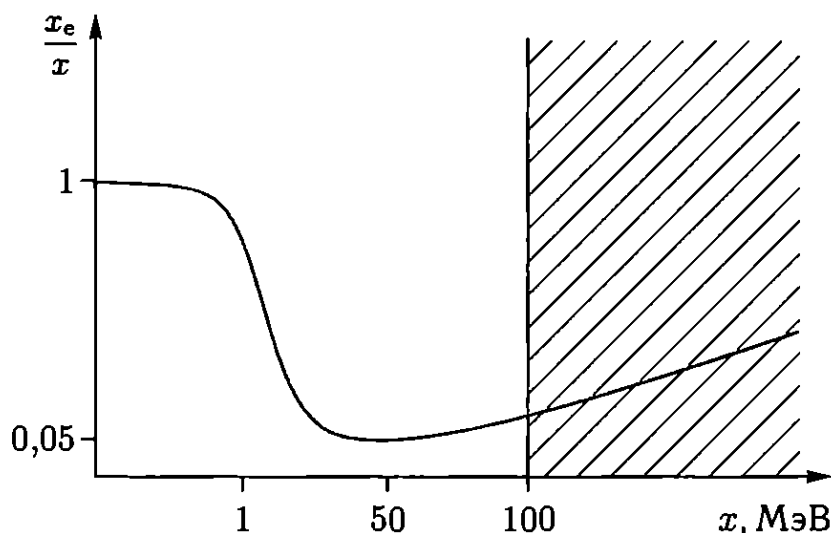


Рис. 3.2. Доля электронов в холодной звезде x_e/x как функция среднего импульса частиц x . В заштрихованной области нейтрон не существует

происходит нейтронизация вещества, и уравнение гравитационного удержания (3.2.1) принимает вид:

$$\mathcal{E}_b(x) = \sqrt{m_n^2 + x^2} - m_n \left(1 - \frac{2\alpha_N x}{m_n}\right)^{-1/2} < 0. \quad (3.2.9)$$

Система барионов «сваливается» в новый минимум ($\alpha_N < 0,5$), который близок к границе существования нейтрона (рис. 3.3). Такие объекты называются нейтронными звездами: их равновесие поддерживается отталкиванием вырожденных нейтронов. Размер нейтронной звезды с солнечной массой составляет около 10 км, что всего в 3 раза больше гравитационного радиуса этой компактной звезды.

По расчетам предельная масса нейтронной звезды близка к $1,8 M_\odot$. Теоретическое определение этой величины не столь надежно, как в случае белого карлика, из-за сложного уравнения состояния нейтронной жидкости. Масса обнаруженных нейтронных звезд не превышает $1,8 M_\odot$.

При дальнейшем увеличении импульса x нейтроны разрушаются, система теряет устойчивость и снова начинает сжиматься. Когда размер системы достигает своего гравитационного радиуса, происходит необратимый коллапс: система барионов «сваливается» в гравитационную яму, у которой нет дна. Успевают ли образоваться еще один минимум функции $\mathcal{E}_b(x)$ в состоянии кварк-глюонной плазмы или нет — очень интересный вопрос, который активно изучается в настоящее время. Некоторые исследователи считают, что успевают. Подтверждением этого могло бы быть наблюдение более массивной и/или компактной звезды, чем нейтронная. Здесь важно уметь определять массу звезды,

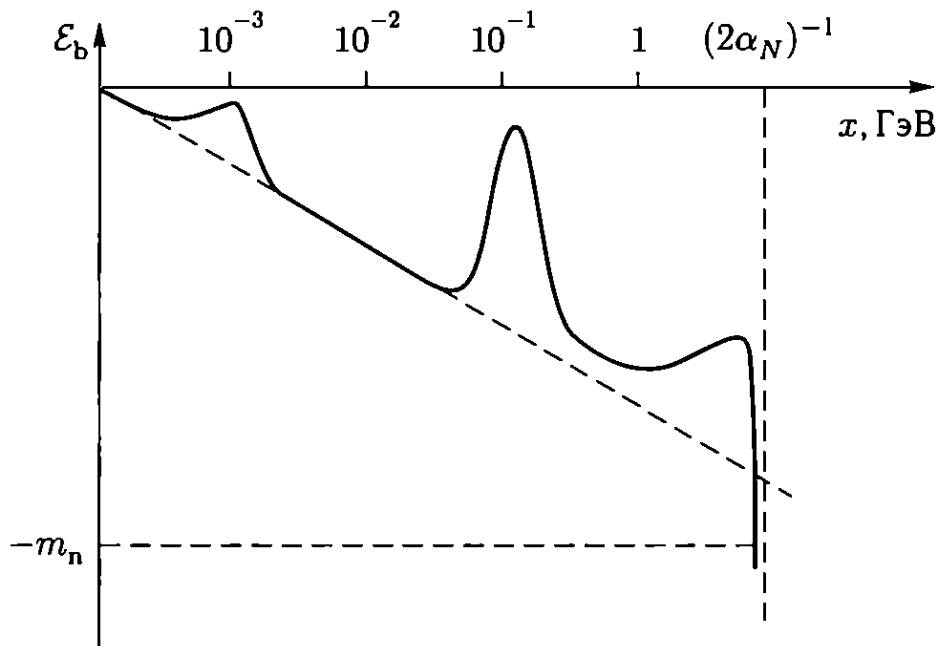


Рис. 3.3. Энергия частиц на один барион $\mathcal{E}_b(x)$ в холодной звезде

в нейтронной структуре которой есть основания сомневаться. Если ее масса окажется больше предельной массы нейтронной звезды, то это будет аргументом в пользу того, что мы имеем дело с кварковой (кварк-глюонной) звездой.

Конечно, наши вычисления носят оценочный характер: у реальной нейтронной звезды есть кора из железа, переходная нейтронная зона, а центральная ее часть состоит из кварк-глюонного конгломерата. Численные расчеты демонстрируют зависимость конечного результата от распределения плотности, уравнения состояния, вращения и других параметров звезды. Сделанные здесь оценки учитывают главную причину равновесия компактных нейтронных звезд — давление вырожденных нейтронов — и подтверждают правильность наших знаний о физике этих объектов.

3.3. Гало темной материи

Говоря о темной материи, мы имеем в виду нерелятивистские массивные частицы, взаимодействующие друг с другом и с другими частицами только гравитационным образом. Название *темная* материя (ТМ) связано с тем, что она невидима, имеет небарионную природу и не взаимодействует со светом¹⁾. Однако ее можно изучать с помощью динамических методов,

¹⁾ Материя может быть «слабо видима», если ее частицы взаимодействуют слабым образом, например аннигилируют, испуская фотоны, в местах высоких концентраций в ядрах галактик.

поскольку ТМ скучена и создает пространственные градиенты гравитационного потенциала, влияющие на движение видимых тел (галактик, звезд, газа), состояние барионов (горячий газ) и отклонение лучей света (гравитационное линзирование).

Средняя плотность ТМ в пять раз превышает космологическую плотность барионов, поэтому именно нерелятивистские темные частицы управляют процессом гравитационного сгущения и эволюцией неоднородной части гравитационного потенциала Вселенной. Контраст плотности слабо неоднородного пространственного распределения ТМ возрастает со временем. Поскольку изначально эта материя холодная, градиенты давления в ней отсутствуют и не могут являться препятствием к развитию гравитационной неустойчивости.

В областях повышенной плотности ТМ темп начального фридмановского расширения материи замедляется, в какой-то момент разбегание останавливается и переходит в коллапс. В ходе динамического сжатия и последующих осцилляций потоков вещества развиваются процессы бесстолкновительной гравитационной *бурной* релаксации: частицы оказываются в переменном (зависящем от времени) гравитационном потенциале, в результате чего часть частиц (около 10%) выбрасывается из системы и уносит избыток положительной энергии. Оставшаяся часть энергии перераспределяется внутри системы и образуется гравитационно-связанный объект ТМ, называемый *гало*. В этом разделе мы рассмотрим простейшие условия равновесия самогравитирующих систем ТМ.

Функция распределения нерелятивистских частиц в сферически-симметричном стационарном гало зависит от радиального расстояния от центра гало r и модулей радиальной и поперечной компонент импульсов частиц. В сферических координатах тензор энергии-импульса материи имеет диагональный вид:

$$T_{\mu}^{\nu} = \rho \operatorname{diag} (1, -\sigma_r^2, -\sigma_t^2, -\sigma_t^2), \quad (3.3.1)$$

где $\rho = \rho(r)$ — профиль плотности гало, $\sigma_r(r)$ и $\sigma_t(r)$ — соответственно радиальная и поперечная (трансверсальная) дисперсии скоростей частиц. В ньютоновском пределе из тождеств Бианки (1.6.3) получаем уравнение гидростатического равновесия сферически-симметричной конфигурации бесстолкновительных частиц:

$$-\frac{1}{\rho r^2} \frac{d}{dr} (\rho r^2 \sigma_r^2) + \frac{2}{r} \sigma_t^2 = \frac{d\Phi}{dr} = \frac{GM(r)}{r^2}, \quad (3.3.2)$$

где гравитационный потенциал Φ , определенный с точностью до аддитивной постоянной, и масса системы M зависят от r и связаны с плотностью следующим образом:

$$\Delta\Phi \equiv \frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d\Phi}{dr} \right) = 4\pi G\rho, \quad M = 4\pi \int_0^r \rho r'^2 dr'. \quad (3.3.3)$$

Частицы ТМ движутся внутри гало по разным геодезическим орбитам — от чисто радиальных до круговых. Каждая орбита характеризуется энергией и вектором углового момента, которые являются интегралами движения в статическом сферически-симметричном поле. Функция распределения частиц зависит от истории образования гало. Численный эксперимент и наблюдения показывают, что для большинства космологических гало распределения частиц по скоростям близки к изотропным, хотя иногда отклонения от изотропии могут быть порядка единицы.

Под гало мы понимаем отрелаксированную систему частиц, гравитационно связанную во всех трех пространственных направлениях. Наблюдаемые гало ограничены сверху по размерам и массам соответственно до нескольких Мпк и $10^{15} M_{\odot}$. Гало бóльших масс еще не успели образоваться во Вселенной, в этом случае эволюция идет по другому: типичными нелинейными образованиями крупномасштабной структуры Вселенной являются системы, сколлапсировавшие в двух или одном направлениях и называемые соответственно *филаментами* и *стенками*. В среднем же по Вселенной в масштабах более 10 Мпк контраст плотности остается меньше единицы ($|\delta\rho/\rho| < 1$) и можно говорить лишь об областях с повышенной и пониженной плотностями материи по сравнению с космологической.

Рассмотрим более подробно изотропную функцию распределения с $\sigma_r = \sigma_t = \sigma$. Равновесие гало определяется балансом градиента эффективного давления материи

$$p = nT = \rho\sigma^2 \quad (3.3.4)$$

и гравитационной силы притяжения полной массы $M = M(r)$:

$$\frac{dp}{\rho dr} = -\frac{GM(r)}{r^2}, \quad (3.3.5)$$

где функции радиуса $n = \rho/m$ и T соответствуют плотности и эффективной температуре частиц ТМ с массой $m = \text{const}$. В силу принципа эквивалентности массу m нельзя определить

из уравнений гравитации, поскольку все частицы движутся по геодезическим независимо от массы. Измеряемыми переменными являются плотность ρ и дисперсия скорости частиц σ .

Уравнение гидростатического равновесия позволяет найти профиль плотности в гало по известному распределению дисперсии скорости $\sigma(r)$, и наоборот, по заданному радиальному распределению плотности восстановить дисперсию скорости. При адиабатическом изменении гравитационного потенциала $\sigma \propto n^{1/3}$, где коэффициент пропорциональности должен зависеть от энтропии. По аналогии с идеальным газом мы можем ввести *функцию энтропии* $E = E(r)$ вириализованного гало следующим образом ¹⁾:

$$E = \sigma^2 \left(\frac{m_p}{\rho} \right)^{2/3} \propto \frac{T}{n^{2/3}} = \frac{p}{n^{5/3}}, \quad (3.3.6)$$

где m_p — масса протона, на которую нормирована масса m .

Функция $E(r)$ является мерой совокупной энтропии частиц гало, накопленной за всю историю его образования. Главные ее составляющие — это начальная (фоновая) энтропия, заданная космологическими мелкомасштабными течениями и неоднородностями материи в протогало, и приобретенная энтропия, сгенерированная в ходе бесстолкновительной и иерархической релаксации материи на нелинейной стадии формирования гало.

При моделировании внутренней структуры наблюдаемых гало часто используется приближение *изотермической сферы*, когда дисперсию скорости можно считать постоянной величиной, не зависящей от радиуса:

$$\sigma(r) = \sigma_0 = \text{const}, \quad \kappa_0 \equiv \frac{4\pi G}{\sigma_0^2} = \frac{4\pi G m}{T} = \text{const}. \quad (3.3.7)$$

Это приближение хорошо согласуется с наблюдательными данными в ограниченном диапазоне масштабов (см., напр., рис. 3.4).

¹⁾ Подчеркнем, что E является не самой энтропией, а *функцией* энтропии идеального газа. Здесь мы не рассматриваем общие вопросы применимости понятия энтропии к бесстолкновительным частицам, определяя функцию $E(r)$ лишь для стационарных *равновесных* систем с *изотропной* функцией распределения частиц по скоростям. Напомним, что в определении (3.3.6) $\sigma^2 = \overline{v^2}$, где v — *одномерная* пекулярная скорость частиц ТМ в точке r , черта означает усреднение в пространстве скоростей.

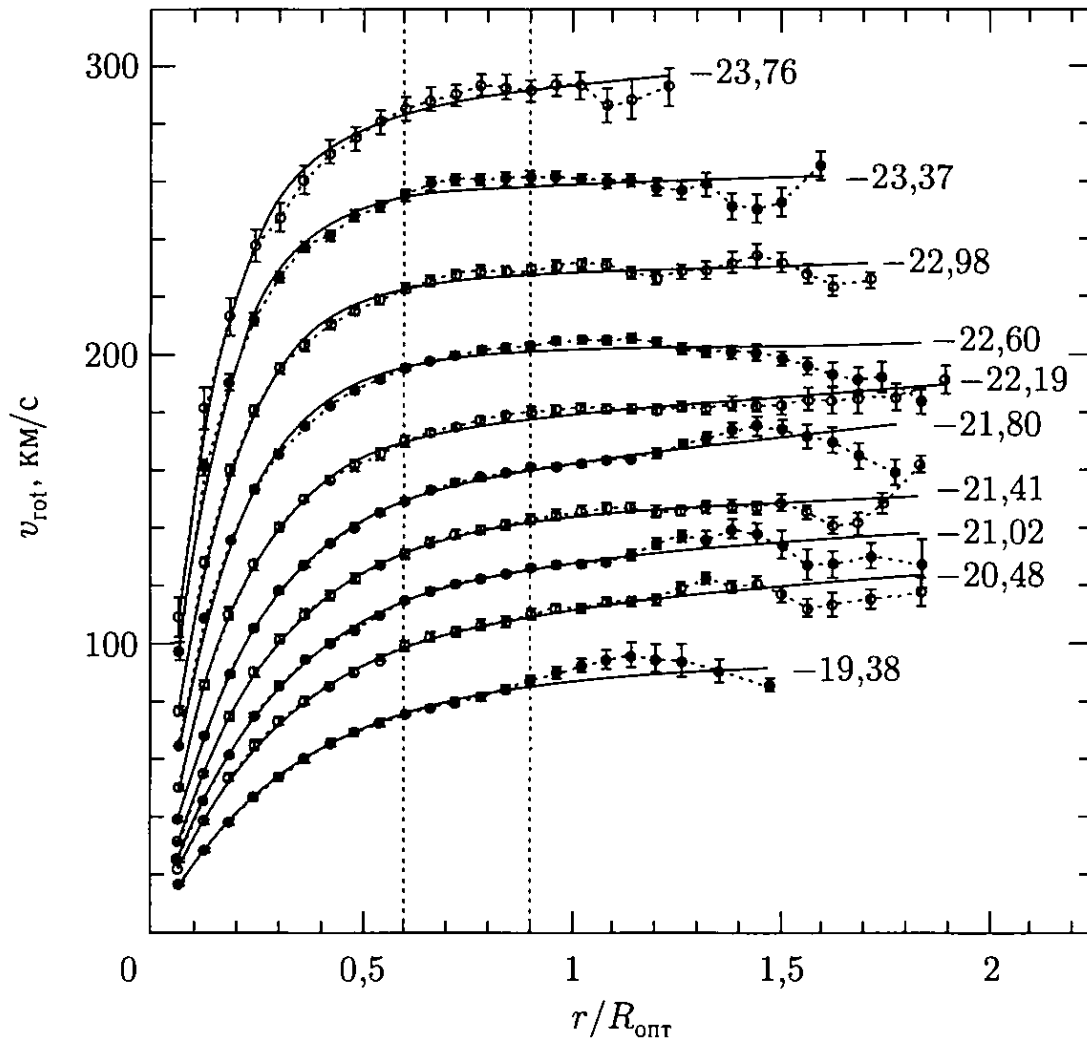


Рис. 3.4. Кривые вращения для 2155 галактик по данным [Catinella et al., 2006]. Кривые разбиты на несколько групп по светимости галактик и характеризуются средней абсолютной звездной величиной в I-фильтре (справа). $R_{\text{опт}}$ — «оптический» радиус галактик, определенный как радиус, внутри которого сосредоточено 83% интегрального потока излучения от галактики

Умножая обе части (3.3.5) на r^2 и дифференцируя по r , получаем уравнение для профиля плотности изотермической сферы:

$$\frac{d}{\rho r^2 dr} \left(\frac{r^2 d\rho}{\rho dr} \right) = -\kappa_0. \quad (3.3.8)$$

Это нелинейное уравнение содержит аттрактор — частное решение, которое является затягивающей сепаратрисой для динамических траекторий (3.3.8). Общее решение имеет вид:

$$\rho(r) = \begin{cases} \frac{\rho_0}{1 + \kappa_0 \rho_0 r^2 / 6}, & r < r_0 \equiv \sqrt{\frac{3}{\kappa_0 \rho_0}}, \\ \frac{2}{\kappa_0 r^2}, & r > r_0 \text{ (аттрактор)}, \end{cases} \quad (3.3.9)$$

где параметр ρ_0 задает центральную плотность ядра гало, а r_0 — размер ядра. Независимо от величины ρ_0 , профиль плотности и масса гало при $r > r_0$ зависят только от дисперсии скорости частиц:

$$M = \frac{8\pi r}{\kappa_0} \simeq 7 \cdot 10^{12} \left(\frac{\sigma}{300 \text{ км} \cdot \text{с}^{-1}} \right)^2 \left(\frac{r}{200 \text{ кпк}} \right) M_{\odot}. \quad (3.3.10)$$

Для характерных значений σ и r мы получили *типичную* массу нелинейных гало ТМ в наблюдаемой Вселенной — массу групп галактик. Напомним, что полная масса Местной группы, в которую входят Млечный Путь и Туманность Андромеды, составляет $2 \cdot 10^{12} M_{\odot}$.

Мы вывели уравнения (3.3.2) и (3.3.5) для ТМ в предположении ее доминирования, однако эти же уравнения описывают гидростатическое равновесие *любой* компоненты вещества (ТМ, звезд, газа) в гравитационном поле *полной* массы $M(r)$. Для этого в их левые части вместо p и ρ надо подставить давление и плотность соответствующей компоненты среды. Так, мы можем переписать массу гало в терминах эффективной температуры гидростатически равновесного газа:

$$M = \frac{2T_{\text{gas}} r}{G\mu_{\text{gas}}} \simeq 7 \cdot 10^{12} \left(\frac{T_{\text{gas}}}{1 \text{ кэВ}} \right) \left(\frac{r}{200 \text{ кпк}} \right) M_{\odot}, \quad (3.3.11)$$

где $\mu_{\text{gas}} \simeq m_p = 1 \text{ ГэВ}$ — молекулярная масса газа. Такой рентгеновский газ, находящийся в скоплениях галактик в течение миллиардов лет, наблюдается с помощью космических телескопов, что позволяет восстанавливать распределения гравитационного потенциала и плотности ТМ в нелинейных гало больших масс.

Вернемся вновь к решению (3.3.9). Плотность ядра $\rho \simeq \rho_0$ никак не связана с гравитацией материи и зависит только от начальных космологических условий, определивших историю образования гало. Размер ядра есть произведение скорости частиц на динамическое время центрального уплотнения:

$$r_0 \simeq \frac{\sigma_0}{2\sqrt{G\rho_0}}.$$

Это радиус, на котором начинает сказываться влияние самогравитации материи. При $r > r_0$ гравитация ТМ кардинально перестраивает внутреннюю структуру гало таким образом, что

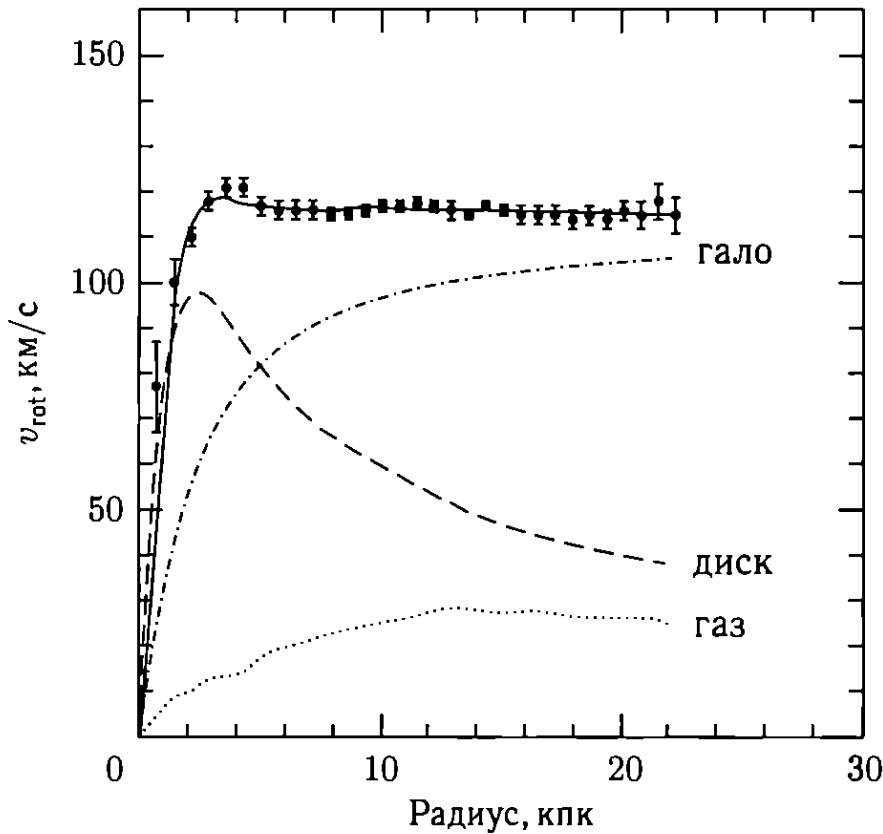


Рис. 3.5. Кривая вращения галактики NGC 6503 по данным [Freese, 2008]. Вклад гало темной материи обозначен штрихпунктирной линией

скорости вращения частиц перестают зависеть от радиуса:

$$v_{\text{rot}}(> r_0) = \sqrt{\frac{GM}{r}} = \sqrt{2} \sigma_0 = \text{const.} \quad (3.3.12)$$

Такие распределения реализуется во многих гравитационно-связанных космологических объектах. В спиральных галактиках наблюдаются *плоские* кривые вращения (рис. 3.5): циркулярные скорости вращения звезд и газа сначала растут с увеличением радиуса, а потом становятся постоянными или изменяются очень медленно. При этом вклад массы ТМ увеличивается с расстоянием, поддерживая плоский участок кривой вращения на протяжении нескольких динамических диапазонов радиуса.

В карликовых галактиках с низкой поверхностной яркостью ТМ доминирует, начиная уже с малых радиусов (см. раздел 3.6, рис. 3.15). Здесь при увеличении радиуса наблюдается *линейный* рост циркулярной скорости, говорящий о насыщении плотности ТМ центрального ядра (см. (3.6.3)). Мы видим, что и начальный подъем галактических кривых вращения также соответствует изотермической сфере распределения ТМ (3.3.9):

$$v_{\text{rot}}(< r_0) = \sqrt{\frac{GM}{r}} = \frac{\sigma_0 r}{r_0}. \quad (3.3.13)$$

Тот факт, что внутренние циркулярные скорости оказываются меньше скоростей частиц гало ($v_{\text{rot}} < \sigma_0$), свидетельствует о том, что в центральной области частицы ТМ движутся в пределах всего ядра $r \sim r_0$, а не находятся лишь в его части, ограниченной данным радиусом r .

Таким образом, приближение изотермической сферы хорошо описывает наблюдаемые распределения ТМ в центральных областях галактических гало (подробнее см. разделы 3.4–3.6). В пределах точности современных наблюдательных данных распределение плотности ТМ внутри гало может быть аппроксимировано следующей простой формулой (ср. (3.3.9)):

$$\rho \simeq \frac{\rho_0}{1+x^2}, \quad x = \frac{r}{r_0}. \quad (3.3.14)$$

3.4. Внутренняя структура гало

При $r_0 \rightarrow 0$ плотность темной материи в центре гало расходится ($\rho_0 \rightarrow \infty$), что получило название *центрального каспа*. При этом масса гало не расходится, а остается конечной. При численном моделировании процесса образования гало (задача N тел) в центрах гало образуются сингулярные каспы, однако в реальной Вселенной такие образования не наблюдаются.

Рисунок 3.6 демонстрирует возникающую в численных моделях расходимость плотности в центре гало темной материи, состоящей из нерелятивистских бесстолкновительных частиц. Сравнение с наблюдаемыми галактическими профилями плотности ТМ представлено на рис. 3.7. Вопрос о том, существуют ли каспы ТМ в природе, важен не только для понимания процессов образования галактик. Он может быть связан с физикой частиц ТМ. На гипотезу галактических каспов опирается один из методов поиска продуктов аннигиляции частиц ТМ.

Как же обстоит дело в реальности? Создает ли природа сингулярные каспы, на которые указывает численный эксперимент? Ведь в центральных областях карликовых галактик с низкой поверхностной яркостью наблюдаются *ядра* — гладкие распределения ТМ с конечной плотностью массы.

Проблема каспов еще ждет своего окончательного решения. Возможно, здесь мы сталкиваемся с ограниченностью стандартной холодной модели, не способной разрешить противоречие между наблюдениями и численным экспериментом. Дело в том, что каспы возникают при релаксации изначально *холодного* вещества, когда крупномасштабные потоки частиц, самопересекаясь, образуют каустики, из которых в дальнейшем

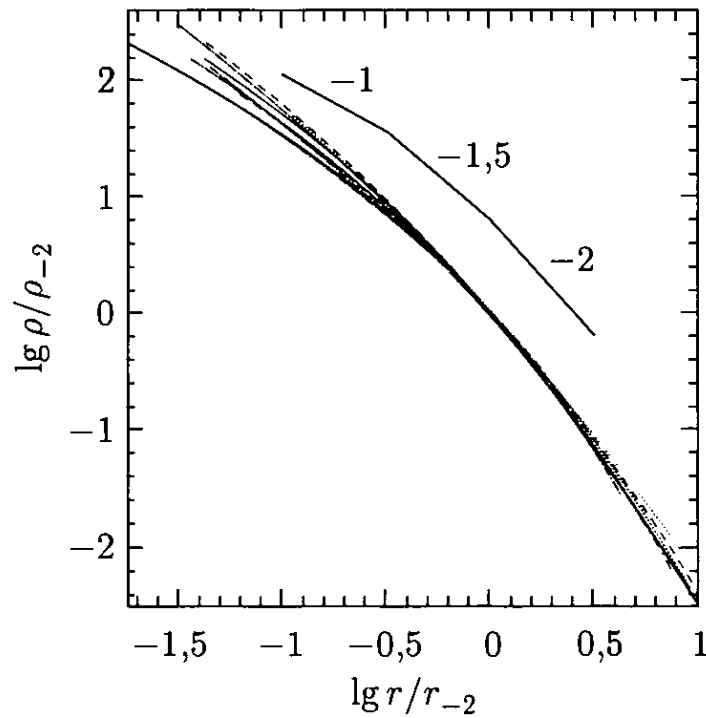


Рис. 3.6. Средние профили плотности гало в численных моделях по данным [Navarro, 2003]. Красная линия — карликовые галактики, зеленая — галактики средней массы, синяя — скопления галактик; цифрами указаны степенные индексы наклона $d \ln \rho / d \ln r$, нормировка соответствует точке наклона -2

и формируются каспы. Реальная же ТМ могла быть изначально *теплой* с небольшими случайными скоростями частиц. При таком начальном состоянии вещества каустики размываются и вместо каспов в центрах гало образуются ядра с более гладкими распределениями плотности материи.

Как получить теплую темную материю?

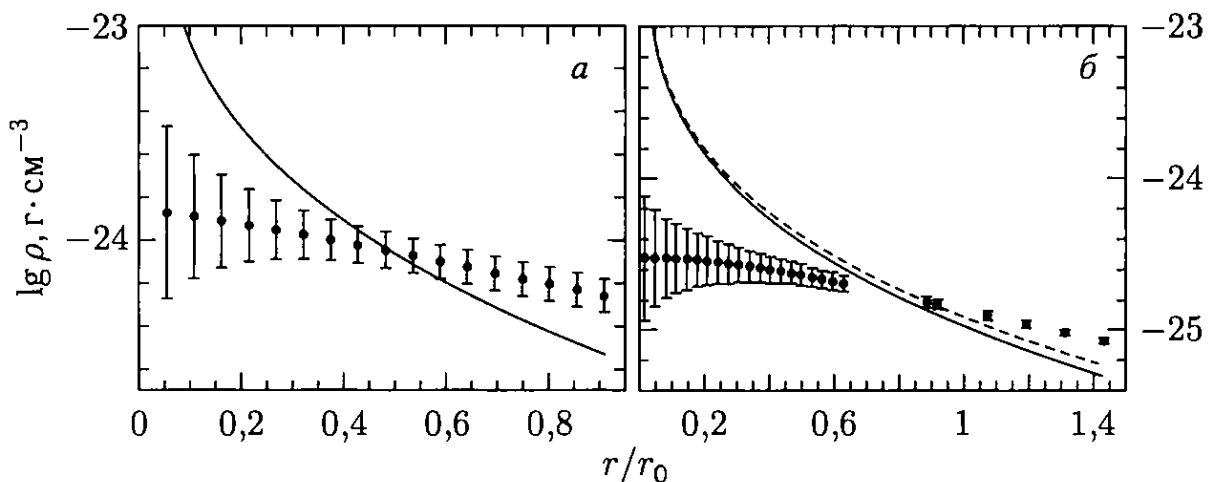


Рис. 3.7. Демонстрация проблемы каспов галактических гало темной материи по данным [Gentile et al., 2007]. Точки — наблюдаемые профили плотности в галактиках DDO 47 (а) и ESO 287-G13 (б), линии — средние профили плотности, полученные с помощью численного моделирования

Кардинальным решением могла бы стать модификация космологической стандартной модели (КСМ), постулирующая многокомпонентность ТМ. Для этого достаточно ввести дополнительную компоненту — теплые частицы небольшой массы (~ 10 кэВ), у которых сохранились остаточные тепловые скорости от горячей фазы их эволюции в ранней Вселенной. Однако для столь существенного усложнения фундаментальной модели, напоминающего скорее «тонкую подстройку», должны быть веские основания. Пока таких оснований нет. Вопрос до конца еще не исследован. Здесь возможны и более простые решения.

Отметим ряд принципиальных моментов, учет которых способен решить проблему каспов в рамках КСМ.

Прежде всего, следует усовершенствовать сам численный эксперимент. Имеющиеся в настоящее время возможности не позволяют пока корректно разрешить центральные области формирующихся гало. В космологических численных моделях разрешение по массе не превышает $10^8 M_{\odot}$, что недостаточно для надежного обнаружения ядер в распределении ТМ. Трюки с последовательным разрешением локальных объемов путем добавления в них «новых точек» на последующих этапах эволюции (так называемый адаптивный метод) также не выручают, так как при такой процедуре теряется информация о начальном мелкомасштабном поле космологических неоднородностей. Эволюцию малых масштабов невозможно последовательно просчитать, начиная, скажем, с $z \sim 40$, поскольку процедура уменьшения масштабов, которые мы хотим разрешить внутри возникающих в процессе счета гало, применяется только в определенных местах, занимающих ничтожную долю полного объема выборки, что приводит к эффективному обрезанию спектра со стороны коротких волн.

Вопрос о влиянии мелкомасштабных возмущений скорости и плотности материи на внутреннюю структуру гало поддается аналитическому исследованию с помощью энтропийной функции (3.3.6). Использование энтропии вместо плотности мотивировано тем, что энтропия только возрастает в процессе формирования и релаксации гало и является интегральной характеристикой всей истории его образования.

Для того чтобы существовал касп ТМ, нужно иметь много холодных частиц в центре гало. Возникает, как минимум, два вопроса.

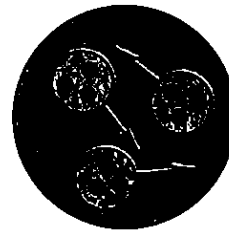
- Где найти холодные частицы в поле возмущений?
- Как перенести холодные частицы внутрь гало?

Идеальный газ



частицы
расстояние между частицами
скорость частицы
динамическое время

Коллапсирующее облако



возмущения
размер возмущения
пекулярная скорость
время коллапса

Рис. 3.8. Сравнение движений частиц в идеальном газе и в протогало темной материи. Энтропия частиц в области размером r определена на характерном динамическом масштабе времени, совпадающем с временем коллапса: $r/v \simeq H^{-1}$

Поскольку доля энергии, сброшенная при формировании гало, не превышает 10–15%, а степень сжатия вещества при релаксации конечна, то оба эти вопроса должны быть адресованы в первую очередь к численным программам N тел: где «холодильник» частиц, почему не учитывается «подогрев» ТМ мелкомасштабными движениями материи?

По определению, частицы ТМ изначально холодные, то есть обладают нулевой энтропией. Однако имеются мелкомасштабные пекулярные потоки и сгустки материи. При сближении друг с другом внутри коллапсирующих протогало они разрушаются приливным образом и рассеиваются, образуя складки и пересечения, «подогревая» тем самым частицы ТМ, привнося в их скорость хаотическую составляющую еще до наступления бурной релаксации, когда придет очередь диссипировать крупномасштабным потокам материи.

Таким образом, поле мелкомасштабных возмущений скорости трансформируется в ходе первого коллапса в хаотическое движение частиц ТМ. Этот процесс можно описать в терминах *начальной энтропии* по аналогии с идеальным газом.

Для введения энтропийной характеристики частиц в протогало требуется понятие *ансамбля*: им является подвыборка поля линейных возмущений, объединяющая все пространственно ограниченные области, которые сколлапсируют к текущему красному смещению z . Схематичное сравнение с идеальным газом дано на рис. 3.8.

Протогало характеризуется массой M_R , коллапсирующей в первый раз к красному смещению z , которое называют *мо-*

ментом образования гало. Массе M_R соответствует линейный сопутствующий размер R :

$$M_R = \frac{4\pi}{3} \rho_m R^3. \quad (3.4.1)$$

В сфере радиусом $R = 10$ Мпк содержится масса типичного скопления галактик $M_{10} \simeq 1,6 \cdot 10^{14} M_\odot$.

Для выделения протогало из общего поля плотности ему необходимо иметь начальный избыточный (*пороговый*) контраст массы в сфере радиусом R , обеспечивающий отрицательную полную энергию связи этой области и ее коллапс к заданному моменту времени:

$$\nu_R \equiv \frac{\delta_R}{\sigma_R} \geq \nu \equiv \frac{\delta_c(z)}{\sigma_R}. \quad (3.4.2)$$

Функция $\delta_c(z)$ рассчитывается аналитически и контролируется численными методами. Для КСМ $\delta_c = \bar{g}(z) \cdot \delta_c(z) = \text{const} \simeq 1,67$, где $\bar{g}(z)$ — фактор роста линейных возмущений плотности ($\bar{g}(0) = 1$). Интересующий нас начальный ансамбль протогало разных масс

$$M_h \geq M_R \quad (3.4.3)$$

включает в себя только те точки пространства \mathbf{x} , для которых средняя плотность материи в шаре радиуса R вокруг данной точки превышает пороговое значение ν (3.4.2), где $\delta(\mathbf{x}) \equiv \rho(\mathbf{x})/\rho_m - 1$ есть линейный контраст плотности по отношению к ее среднему значению во Вселенной $\rho_m = \langle \rho \rangle$,

$$\delta_R = \delta_R(\mathbf{x}) = \int \delta(\mathbf{x}') W_R(|\mathbf{x} - \mathbf{x}'|) d\mathbf{x}' \quad (3.4.4)$$

— сглаженный контраст плотности с фильтром радиусом R :

$$W_R(r) = \frac{3}{4\pi R^3} \begin{cases} 1, & r \leq R, \\ 0, & r > R \end{cases}$$

(*функция окна*). Очевидно, в функции (3.4.4) отсутствует информация о возмущениях в масштабах меньше R .

Дисперсия линейного контраста сглаженной плотности σ_R , аппроксимированная по линейному закону эволюции к сегодняшнему моменту времени $z = 0$, связана со спектром мощности скалярных неоднородностей $P(k)$ следующим интегральным

соотношением ¹⁾:

$$\sigma_R^2 \equiv \langle \delta_R^2 \rangle = \int_0^\infty P(k) W^2(kR) k^2 dk, \quad (3.4.5)$$

где $W(y) = 3y^{-3}(\sin y - y \cos y)$ — фурье-образ функции окна ($y = kR$). Из наблюдений известно, что $\sigma_{11} \simeq 0,8$ для радиуса $R \simeq 11$ Мпк, охватывающего массу скопления галактик $M_{11} \simeq 2 \cdot 10^{14} M_\odot$. В астрономии это эквивалентно нормировке «на сигму-восемь», где индекс «8» соответствует физическому радиусу $R = 8/0,7 \simeq 11$ Мпк (подробнее см. раздел 12.8).

Мы будем использовать полученные ансамбли для описания средних «термодинамических» характеристик ТМ в протогало. Полная энтропия вириализованного гало объединяет (интегрирует) энтропию всех неравновесных процессов, протекавших за время его формирования. Она включает как *фоновую* энтропию частиц, связанную с начальными космологическими возмущени-

¹⁾ Фурье-гармоники линейного поля возмущений δ -коррелированы (см. раздел 3.5):

$$\delta(\mathbf{x}) = (2\pi)^{-3/2} \int \delta_{\mathbf{k}} e^{i\mathbf{k}\mathbf{x}} d\mathbf{k}, \quad \langle \delta_{\mathbf{k}} \delta_{\mathbf{k}'}^* \rangle = 2\pi^2 P(k) \delta^{(3)}(\mathbf{k} - \mathbf{k}'),$$

где ожидание $\langle \dots \rangle$ берется по всем реализациям случайного гауссова поля, $\delta^{(3)}(\mathbf{k})$ — δ -функция Дирака. Опираясь на это положение, можно построить любые корреляторы поля космологических возмущений, например:

$$\langle \delta(\mathbf{x} + \mathbf{r}) \delta(\mathbf{x}) \rangle = \int_0^\infty P(k) \frac{\sin kr}{kr} k^2 dk,$$

$$\langle \mathbf{S}(\mathbf{x} + \mathbf{r}) \delta(\mathbf{x}) \rangle = -\frac{\mathbf{r}}{3} \int_0^\infty P(k) W(kr) k^2 dk,$$

$$\langle S_j(\mathbf{x} + \mathbf{r}) S_k(\mathbf{x}) \rangle = e_j e_k \int_0^\infty P(k) \left[\frac{\sin kr}{kr} - \frac{2}{3} W(kr) \right] dk + \frac{1}{3} p_{jk} \int_0^\infty P(k) W(kr) dk,$$

где $\mathbf{e} = e_j = r_j/r$ — единичный вектор в направлении радиуса \mathbf{r} , $p_{jk} = \sigma_{jk} - e_j e_k$ — проекционный тензор,

$$\delta = -\text{div } \mathbf{S}, \quad \mathbf{S} = S_j = i(2\pi)^{-3/2} \int \frac{k_j}{k^2} \delta_{\mathbf{k}} e^{i\mathbf{k}\mathbf{x}} d\mathbf{k}$$

— вектор смещения элементов материи (см. (3.5.1)–(3.5.8)). Таким же образом можно представить сглаженные поля, например:

$$\delta_R = (2\pi)^{-3/2} \int \delta_{\mathbf{k}} W(kR) e^{i\mathbf{k}\mathbf{x}} d\mathbf{k},$$

$$\langle \mathbf{S}(\mathbf{x} + \mathbf{r}) \delta_R(\mathbf{x}) \rangle = -\frac{\mathbf{r}}{3} \int_0^\infty P(k) W(kR) W(kr) k^2 dk$$

и т.д.

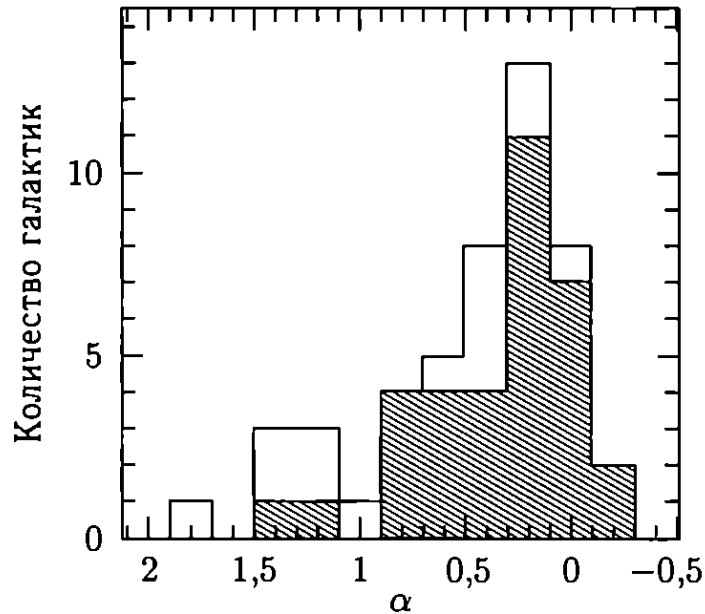


Рис. 3.9. Распределение степенного индекса плотности α в центральных областях галактик с низкой поверхностной яркостью [de Blok et al., 2001]

ями, так и *приобретенную* энтропию, сгенерированную в ходе бурной релаксации и иерархического сгущивания материи. Космологическую часть можно рассчитать аналитически, а для оценки нелинейной составляющей мы воспользуемся результатами численного моделирования.

Прежде всего следует разобраться в том, каким образом распределена энтропия внутри равновесного гало. Для упрощения вычислений ограничимся степенной аппроксимацией центрального профиля плотности гало:

$$\rho(r) \propto r^{-\alpha}, \quad (3.4.6)$$

$$M(r) = 4\pi \int_0^r \rho(r) r^2 dr \propto r^{3-\alpha},$$

где индекс наклона $\alpha \in (0, 2.5)$. В численных моделях возникают профили с $\alpha \geq 1$, тогда как наблюдаются чаще всего значения $\alpha < 1$ (рис. 3.9). В первом случае мы имеем дело с каспом ($1 \leq \alpha < 2.5$), а во втором — с ядром ($0 \leq \alpha < 1$).

Физическое различие между каспом и ядром следует из поведения эффективного давления, которое можно определить из уравнения равновесия (3.3.5):

$$p(r) = c_1 + c_2 r^{2(1-\alpha)}, \quad (3.4.7)$$

где c_1 и c_2 — константы интегрирования. Критической величиной является значение $\alpha = 1$: в ядре давление конечно, тогда как в центре каспа оно расходится.

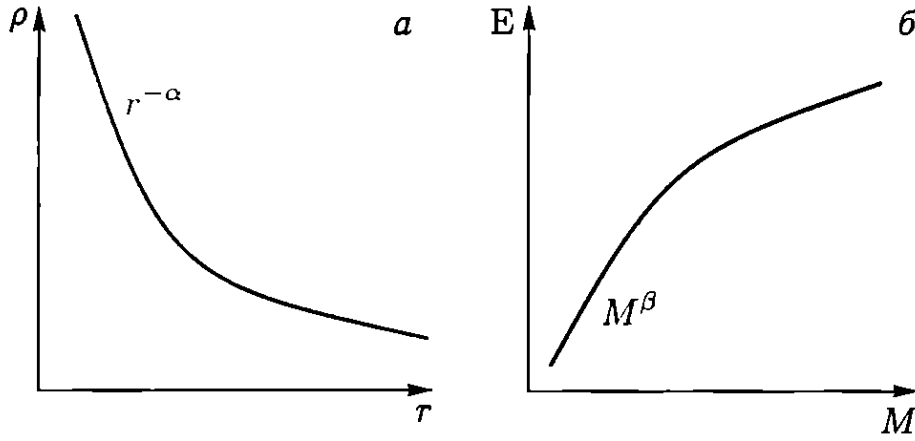


Рис. 3.10. Профили плотности (а) и энтропии (б) в равновесном гало темной материи

Из (3.3.6) получаем энтропийную функцию масс:

$$E(M) \propto c_1 M^{\beta_1} + c_2 M^{\beta_2} \propto M^{\beta}, \quad (3.4.8)$$

$$\beta_1 = \frac{5\alpha}{3(3-\alpha)}, \quad \beta_2 = \frac{6-\alpha}{3(3-\alpha)},$$

где параметр $\beta = d \ln E / d \ln M$ лежит в интервале значений между β_1 и β_2 , и его локальное значение зависит от текущей массы гало. Оба индекса $\beta_{1,2}$ возрастают с ростом α . При $\alpha = \alpha_c \equiv 1$ интервал возможных значений β стягивается в точку:

$$\beta_1 = \beta_2 = \beta_c \equiv \frac{5}{6}.$$

Вместо радиуса r мы перешли к сопутствующей массе M в шаре радиуса r , поскольку она сохраняется и для линейных возмущений, и в ходе релаксации.

Как мы уже отмечали, энтропийная функция масс аккумулирует воздействие необратимых процессов в ходе эволюции ТМ и задает профиль равновесного гало. В рассматриваемом интервале значений α мы имеем $\beta_{1,2} > 0$, так что $E(M) \rightarrow 0$ при $M \rightarrow 0$ (рис. 3.10). Однако в центре каспа энтропия частиц пренебрежимо мала:

$$\beta > 5/6: \quad M^{-5/6} E(M) \rightarrow 0 \quad \text{при} \quad M \rightarrow 0,$$

тогда как в центре ядра она существенно выше (что и делает невозможным возникновение каспа):

$$\beta < 5/6: \quad M^{-5/6} E(M) \rightarrow \infty \quad \text{при} \quad M \rightarrow 0.$$

Для критического значения $\beta_c = 5/6$ из (3.3.6) и (3.4.8) получаем:

$$M \propto \sigma^4. \quad (3.4.9)$$

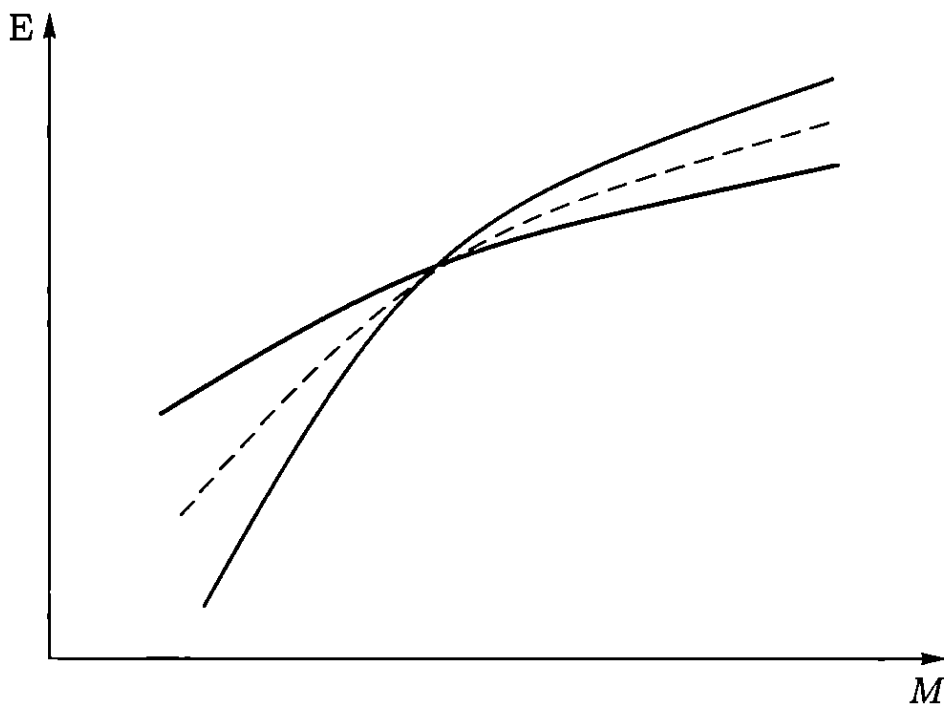


Рис. 3.11. Энтропийная функция масс равновесного гало. Красная линия — касп, синяя — ядро, штриховая линия соответствует значению $\beta_c = 5/6$

Эта формула воспроизводит известные эмпирические соотношения Фабер–Джексона и Талли–Фишера, связывающие наблюдаемую дисперсию скоростей звезд или газа с массой (светимостью) соответственно в эллиптических галактиках или сферических подсистемах спиральных галактик.

Рисунок 3.11 иллюстрирует различие между каспом и ядром равновесного гало с точки зрения распределения энтропии. Как видно, вопрос о формировании центрального каспа или ядра зависит от полной энтропии материи, приобретенной до, в процессе и после образования гало.

Теперь мы можем количественно ответить на вопрос о том, откуда берется температура у изначально холодных частиц в КСМ.

Историю образования гало можно разбить на следующие этапы.

- Начальные движения частиц ТМ потенциальны и сфазированы, относительные скорости близких частиц равны нулю, имеется гладкое поле скоростей ТМ.
- Коррелированы средние по ансамблю потоки (*средняя скорость*). Случайные хаотические отклонения от средней скорости (*вариация скорости*) увеличиваются с ростом масштаба, достигая насыщения на относительно большой корреляционной длине. Для возникновения хаотических движений необходимо сблизить частицы среды.

- В ходе первого коллапса протогало происходит сближение частиц и теряется начальная фаза *варьированной* части скорости: частицы ТМ «подогреваются» (возникает фоновая энтропийная функция масс E_b).
- Хаотизация *средней* скорости происходит несколько позже в ходе бурной релаксации и слияния сгустков (генерируется энтропия E_g).
- Результирующий энтропийный профиль равновесных гало получается объединением (суммированием) начальной E_b и приобретенной E_g энтропий.

В следующем разделе мы получим начальный *средний* энтропийный профиль $\langle E_b \rangle$, используя в (3.3.6) дисперсию космологических линейных возмущений скорости. Такой способ расчета предполагает почти адиабатическое сжатие материи в протогало и является оценкой снизу на функцию $E_b(M)$ при малых M .

3.5. Решение проблемы галактических каспов

Поле скалярных космологических неоднородностей в пост-рекомбинационную эпоху характеризуется тремя калибровочно-инвариантными переменными: пространственным вектором смещения материальной точки из невозмущенного положения

$$\mathbf{S} = \mathbf{S}(\mathbf{x}), \quad (3.5.1)$$

полной скоростью материи

$$\mathbf{V} = \mathbf{V}(z, \mathbf{x}) \quad (3.5.2)$$

и сопутствующим возмущением плотности

$$\delta_m \equiv \frac{\rho(z, \mathbf{x})}{\rho(z)} - 1 = \bar{g}(z) \cdot \delta(\mathbf{x}). \quad (3.5.3)$$

Зависимость плотности от красного смещения имеет вид:

$$\rho(z) = \rho_m (1 + z)^3,$$

где средняя плотность темной материи задается выражением:

$$\rho_m = \frac{3\Omega_m H_0^2}{8\pi G} \simeq 3 \cdot 10^{-30} \frac{\text{г}}{\text{см}^3}.$$

Среднее смещение во Вселенной равно нулю, а его дисперсия определяется интегралом от спектра мощности возмущений

плотности:

$$\sigma_{\mathbf{S}} \equiv \sqrt{\langle \mathbf{S}^2 \rangle} = \left(\int_0^{\infty} P(k) dk \right)^{1/2} \simeq 14 \text{ Мпк.} \quad (3.5.4)$$

Сфера радиусом $R = \sigma_{\mathbf{S}}$ охватывает массу богатого скопления галактик (ср. (3.4.1)):

$$M_{\mathbf{S}} = \frac{4\pi}{3} \rho_m \sigma_{\mathbf{S}}^3 \simeq 4 \cdot 10^{14} \text{ Мпк.}$$

Соотношения между переменными определяются линейной теорией возмущений (см. раздел 12.2):

$$\tau(z, \mathbf{x}) = (1+z)^{-1} (\mathbf{x} + \bar{g} \mathbf{S}), \quad (3.5.5)$$

$$\mathbf{V} \equiv \dot{\tau} = H (\tau - \bar{g}'_z \mathbf{S}), \quad \delta(\mathbf{x}) = -\text{div } \mathbf{S}, \quad (3.5.6)$$

$$\mathbf{v} = \mathbf{v}_{\text{pec}} \equiv \mathbf{V} - H \tau = -\bar{g}'_z H \mathbf{S}, \quad (3.5.7)$$

где $\tau(z, \mathbf{x})$ и \mathbf{x} — соответственно эйлеровы и лагранжевы координаты среды, штрих у функции \bar{g} означает здесь производную по красному смещению z . Фактор роста $\bar{g} = \bar{g}(z)$ нормирован на единицу при нулевом красном смещении (см. (12.1.15)), а при $z > 1$ равен

$$\bar{g} \simeq \frac{1,3}{1+z}.$$

Функция Хаббла и соотношение между смещением и пекулярной скоростью материи при $z > 1$ имеют вид:

$$\begin{aligned} H &\simeq 0,5 \cdot (1+z)^{3/2} H_0, \\ \mathbf{v} &\simeq 0,5 \bar{g} \sqrt{1+z} H_0 \mathbf{S} \simeq \frac{45}{\sqrt{1+z}} \frac{\mathbf{S}}{\text{Мпк}} \frac{\text{км}}{\text{с}}, \\ \sigma_{\mathbf{v}} &\equiv \sqrt{\langle \mathbf{v}^2 \rangle} \simeq 200 \sqrt{\frac{10}{1+z}} \frac{\text{км}}{\text{с}}. \end{aligned} \quad (3.5.8)$$

Перейдем к характеристикам мелкомасштабных движений материи.

Введем вектор относительного смещения точек среды, находящихся на расстоянии \mathbf{r} друг от друга (см. (3.5.1)):

$$\mathbf{s} = \mathbf{s}(\mathbf{r}, \mathbf{x}) = \frac{\mathbf{S}(\mathbf{x} + \mathbf{r}) - \mathbf{S}(\mathbf{x})}{\sqrt{2} \sigma_{\mathbf{S}}}. \quad (3.5.9)$$

Он описывает поле нормированной пекулярной скорости материи с текущим радиус-вектором \mathbf{r} , отсчитываемым от точки \mathbf{x} .

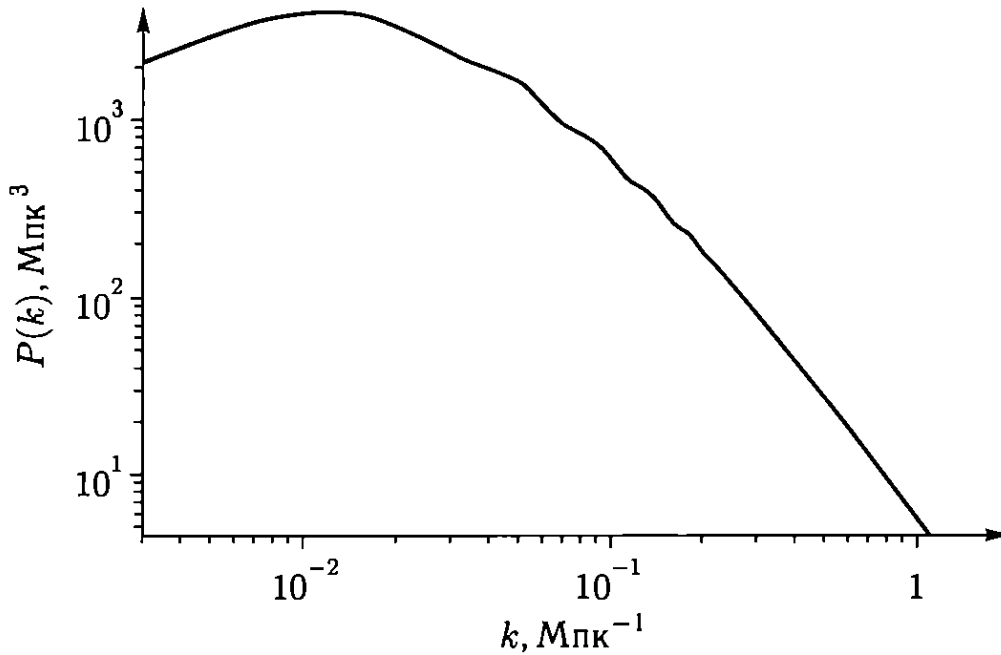


Рис. 3.12. Спектр мощности $P(k)$ в стандартной модели; $\sigma_{10} = 1$

Нас интересует средняя относительная скорость и ее вариация как функции r при усреднении по разным ансамблям (или точкам x) в разных областях Вселенной.

Модуль радиус-вектора $r = |\mathbf{r}|$ будем фиксировать в интервале линейных масштабов, соответствующих внутренним областям гравитационно-связанного гало размера R :

$$r < R. \quad (3.5.10)$$

При $R < \sigma_S$ эти масштабы относятся к коротковолновой части спектра $P(k)$ (справа от максимума, см. рис. 3.12):

$$k\sigma_S \gg 1: \quad P(k) \propto k^{-3} \ln^2(k\sigma_S).$$

Отсюда получаем следующую оценку для σ_r (см. (3.4.5)):

$$\sigma_r^2 = \int_0^{\infty} P(k) W^2(kr) k^2 dk \propto \ln^3(2\sigma_S/r). \quad (3.5.11)$$

Учитывая нормировку спектра мощности, согласно которой $\sigma_{10} \simeq 1$, приближенно можно считать

$$1 \lesssim \sigma_R \sim \ln^{3/2}(2\sigma_S/R) \quad \text{при} \quad R \lesssim 10 \text{ Мпк.}$$

При $R > 10$ Мпк необходимо пользоваться более точной оценкой для $\sigma_R < 1$, принимая во внимание изменение наклона спектра $P(k)$ в области богатых скоплениях галактик.

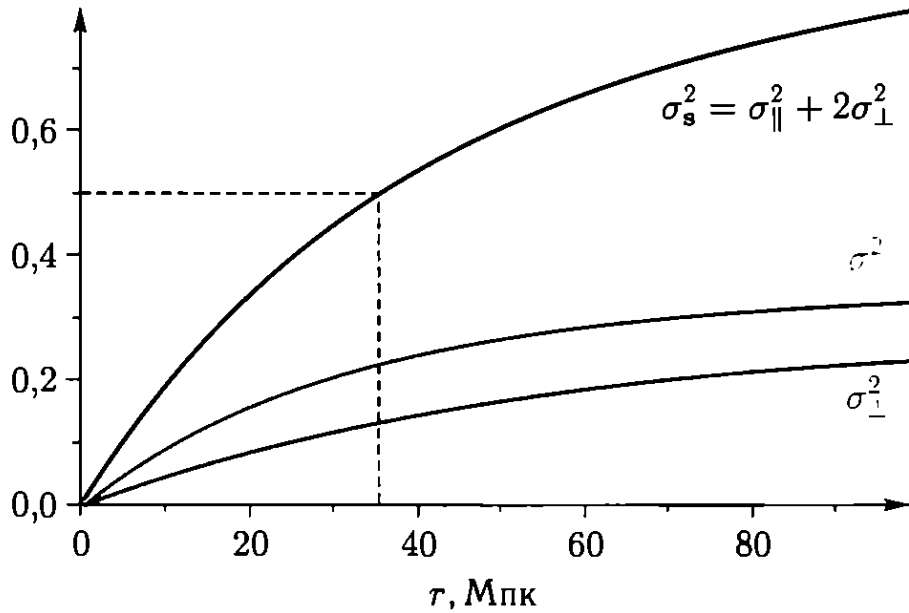


Рис. 3.13. Вариация относительной скорости $\sigma_s^2(r)$ (черная линия) и ее продольная $\sigma_{\parallel}^2(r)$ (зеленая) и поперечная $\sigma_{\perp}^2(r)$ (красная) компоненты

Усредняя по всей Вселенной, получаем выражения для средней относительной скорости и ее вариации:

$$\langle \mathbf{s} \rangle = 0, \quad (3.5.12)$$

$$\sigma_s^2 \equiv \langle \mathbf{s}^2 \rangle = \sigma_S^{-2} \int_0^{\infty} P(k) \left(1 - \frac{\sin kr}{kr} \right) dk. \quad (3.5.13)$$

Как мы видим, варьированная часть скорости $\sigma_s = \sigma_s(r)$ монотонно растет с увеличением масштаба r (рис. 3.13). При $r < \sigma_S$ ее рост имеет квазистепенной характер с плавным уменьшением индекса наклона, начиная от единицы:

$$\sigma_s \simeq \frac{r\sigma_r}{\sqrt{6}\sigma_S}.$$

При $r \gtrsim \sigma_S$ рост функции σ_s замедляется, а при $r \gg \sigma_S$ он достигает насыщения: $\sigma_s \rightarrow 1$. При $r = \sigma_S$: $\sigma_s \simeq 0,5$.

Корреляционный радиус функции (3.5.13), при котором $\sigma_s^2(r_s) = 1/2$, равен $r_s = 36$ Мпк. Заметим, что это *средний* корреляционный радиус относительной скорости материи. Покомпонентная вариация вектора (3.5.9) анизотропна и зависит от направления между точками среды $\mathbf{e} = \mathbf{r}/r$:

$$c_{ij} \equiv \langle s_i s_j \rangle = \sigma_{\parallel}^2 e_i e_j + \sigma_{\perp}^2 p_{ij}, \quad (3.5.14)$$

$$\sigma_s^2 = c_i^i = \sigma_{\parallel}^2 + 2\sigma_{\perp}^2, \quad (3.5.15)$$

где $p_{ij} = \delta_{ij} - e_i e_j$ — проекционный тензор на плоскость, ортогональную вектору e ,

$$\sigma_{\parallel}^2 = \sigma_S^{-2} \int_0^{\infty} P(k) \left[\frac{1 + 2W(kr)}{3} - \frac{\sin kr}{kr} \right] dk,$$

$$\sigma_{\perp}^2 = \frac{1}{3} \sigma_S^{-2} \int_0^{\infty} P(k) [1 - W(kr)] dk.$$

У продольной вариации корреляционный радиус равен 20 Мпк ($\sigma_{\parallel}^2 = 1/6$), тогда как у поперечной он составляет 49 Мпк ($\sigma_{\perp}^2 = 1/6$).

Усредняя по ансамблю протогало¹⁾ со средним контрастом плотности ν_R (см. (3.4.2)), получаем условные ожидания мелко-масштабных скоростей материи в протогало:

$$\langle \mathbf{s} | \nu_R \rangle = \mathbf{c} \nu_R, \quad (3.5.16)$$

$$\sigma_{\mathbf{s}|\nu}^2 \equiv \langle \mathbf{s}^2 | \nu_R \rangle - \langle \mathbf{s} | \nu_R \rangle^2 = \sigma_S^2 - \mathbf{c}^2, \quad (3.5.17)$$

¹⁾ Для вычисления условной вероятности $p(\mathbf{s} | \nu_R)$ найти \mathbf{s} при заданной величине ν_R рассмотрим гауссову функцию распределения безусловной вероятности для совместной переменной $x_A = (\mathbf{s}, \nu_R)$, где индекс A пробегает значения i, R :

$$p(\mathbf{s}, \nu_R) = \frac{\exp(-\kappa^2/2)}{(2\pi)^2 \sqrt{C}}, \quad \kappa^2 = c^{AB} x_A x_B,$$

где c^{AB} — матрица, обратная к матрице

$$c_{AB} = \langle x_A x_B \rangle = \begin{pmatrix} c_{ij} & \mathbf{c} \\ \mathbf{c} & 1 \end{pmatrix},$$

$$C \equiv \det(c_{AB}) = \bar{\sigma}_{\parallel}^2 \sigma_{\perp}^4, \quad \bar{\sigma}_{\parallel}^2 = \sigma_{\parallel}^2 - \mathbf{c}^2.$$

После несложных преобразований получаем:

$$\kappa^2 = \bar{\kappa}^2 + \nu_R^2, \quad \bar{\kappa}^2 = \bar{c}^{ij} \bar{s}_i \bar{s}_j, \quad \bar{\mathbf{s}} = \mathbf{s} - \mathbf{c} \nu_R,$$

где \bar{c}^{ij} — матрица, обратная к матрице

$$\bar{c}_{ij} \equiv \langle \bar{s}_i \bar{s}_j | \nu_R \rangle = \bar{\sigma}_{\parallel}^2 e_i e_j + \sigma_{\perp}^2 p_{ij},$$

откуда получаем условную функцию распределения вектора \mathbf{s} и дисперсию скоростей внутри протогало с контрастом плотности ν_R :

$$p(\mathbf{s} | \nu_R) = \frac{p(\mathbf{s}, \nu_R)}{p(\nu_R)} = \frac{\exp(-\bar{\kappa}^2/2)}{(2\pi)^{3/2} \bar{\sigma}_{\parallel} \sigma_{\perp}^2}, \quad \sigma_{\mathbf{s}|\nu}^2 \equiv \langle \bar{\mathbf{s}}^2 | \nu_R \rangle = \bar{\sigma}_{\parallel}^2 + 2\sigma_{\perp}^2.$$

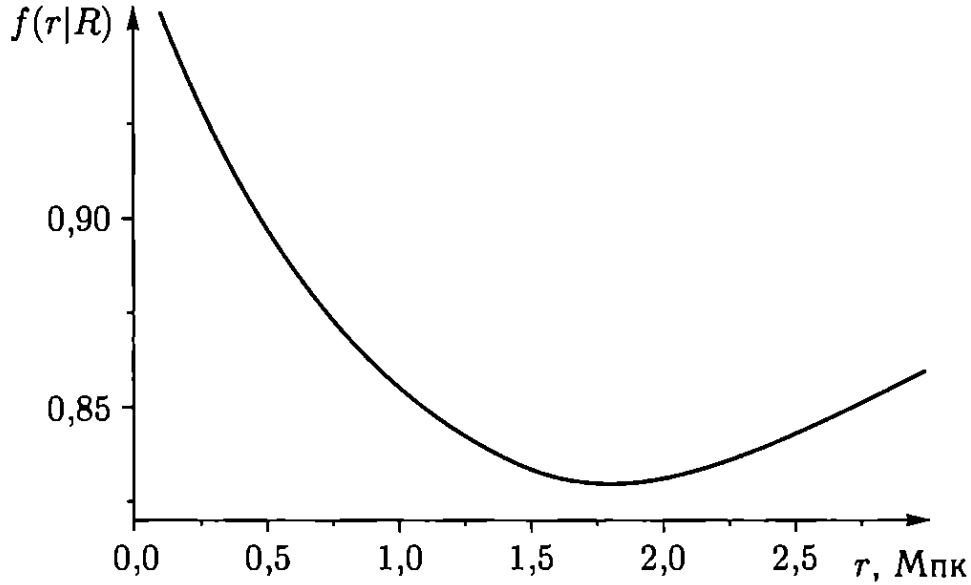


Рис. 3.14. Фактор подавления $f(r|R)$ пекулярных скоростей внутри гало с массой $M = 1,3 \cdot 10^{12} M_{\odot}$ (линейный размер $R = 2$ Мпк)

где коэффициент кросс-корреляции имеет вид (см. (3.5.10)):

$$c \equiv \langle s\nu_R \rangle = -\frac{r}{3\sqrt{2}\sigma_S\sigma_R} \int_0^{\infty} P(k) W(kR) W(kr) k^2 dk \simeq -\frac{r\sigma_R}{3\sqrt{2}\sigma_S}.$$

Функция (3.5.16) представляет собой средний универсальный профиль скорости коллапсирующего протогало массы M_R . Условная вариация (3.5.17) описывает дисперсии случайных отклонений скорости от средней внутри протогало:

$$\sigma_{s|\nu}^2(r) \simeq \sigma_S^2 - \frac{1}{18} \left(\frac{r\sigma_R}{\sigma_S} \right)^2. \quad (3.5.18)$$

Фактор подавления пекулярных движений ТМ в протогало по сравнению с относительными движениями материи в целом во Вселенной

$$f(r|R) \equiv \frac{\sigma_{s|\nu}^2}{\sigma_S^2} \simeq 1 - \frac{1}{18} \left(\frac{r\sigma_R}{\sigma_S} \right)^2$$

представлен на рис. 3.14 для галактической массы $1,3 \cdot 10^{12} M_{\odot}$ ($R = 2$ Мпк). Для галактик и групп галактик кросс-корреляция внутри протогало не существенна:

$$\sigma_{s|\nu}^2(r) \simeq \frac{r^2}{6\sigma_S^2} \left(\sigma_r^2 - \frac{1}{3}\sigma_R^2 \right) \simeq \sigma_S^2, \quad (3.5.19)$$

и дисперсия мелкомасштабных скоростей ТМ имеет *универсальный* вид, независимый от массы протогало (см. (3.5.8))

$$\sigma_v \equiv \sigma_v \sigma_{s|\nu} \simeq \sigma_v \sigma_S,$$

однако величина дисперсии зависит от момента образования гало z .

Предполагая, что энтропия $E(M)$ не изменяется при сжатии протогало (адиабатическое сжатие), получаем средний профиль фоновой энтропийной функции масс (3.3.6) в параметрическом представлении:

$$\langle E_b \rangle = \frac{\sigma_v^2}{3} \left(\frac{m_p}{\rho(z)} \right)^{2/3} \propto \sigma_s^2(r), \quad M = M_T \propto r^3. \quad (3.5.20)$$

Отсюда вычисляем степенной индекс наклона β_b , зависящий от показателя n центральной массы гало $M \equiv 10^n M_\odot < M_R < M_S \simeq 4 \cdot 10^{14} M_\odot$:

$$\beta_b = \frac{d \ln \langle E_b \rangle}{d \ln M} = \frac{2}{3} - \frac{1}{\ln(2\sigma_S/r)} \simeq \frac{2}{3} - \frac{1,3}{16 - n}. \quad (3.5.21)$$

Таблица 3.1. Степенной индекс фоновой энтропийной функции β_b для центральных масс гало $M \equiv 10^n M_\odot < M_R < 4 \cdot 10^{14} M_\odot$

n	12	10	8	6	4
β_b	0,34	0,45	0,50	0,53	0,56

В таблице 3.1 приведены численные значения β_b для разных n . Мы видим, что в гало галактик и групп галактик фоновая энтропия препятствует образованию центрального каспа:

$$\beta_b < 0,6 < \frac{5}{6}. \quad (3.5.22)$$

При увеличении R за область σ_S и пропорциональном увеличении внутреннего радиуса $r < R$ (то есть для больших гало $M_R \geq 4 \cdot 10^{14} M_\odot$)¹⁾ кросс-корреляцией $\langle s\nu_R \rangle$ пренебрегать уже нельзя (см. (3.5.18)). По этой причине влияние фоновой энтропии ТМ внутри богатых скоплений галактик ослаблено.

Таким образом, в центральной области гало галактик и их групп формируются ядра ТМ, которые, следуя астрономической традиции, уместно назвать сферической подсистемой или «балджем» ТМ (по аналогии с барионным балджем, см. сноску на с. 122). В частности, эллиптические галактики представляют со-

¹⁾ Заметим, что радиус r влияния ТМ нельзя считать малым, принимая во внимание «холодные» барионные потоки и массивные сD-галактики в центрах скоплений. По сходной причине нет смысла рассматривать центральные массы ТМ меньше массы черных дыр в центрах галактик. Напомним в этой связи, что характерные масштабы распределений ТМ много больше размеров центральных ярких барионных образований (балджей, от английского «buldge»).

бой единый большой балдж. Если предположить, что в области максимальной светимости $\alpha \sim 1$ (см. (3.4.9)), то для балджей должно выполняться соотношение Фабера–Джексона, что и наблюдается в эллиптических галактиках. В дисковых и спиральных галактиках балдж занимает относительно небольшую область гало, хотя светимость его в линии H_α может быть значительной, что, по-видимому, объясняет соотношение Тулли–Фишера (см. (3.4.9)).

В больших спиральных галактиках балдж ТМ деформирован из-за притока туда барионов¹⁾, однако его размер все же достаточно велик. Так, в галактике Млечный Путь балдж занимает несколько килопарсеков. При увеличении r он трансформируется в распределение ТМ $\rho \propto r^{-2}$ с почти плоской кривой вращения, которая, в свою очередь, простирается до десятка кпк и возможно далее за край диска, где о дальнейшем распределении ТМ можно судить по движениям карликовых галактик-спутников.

Что касается гравитационно-связанных групп галактик, то, согласно расчетам, центральный балдж ТМ успел образоваться только в *компактных* группах. Существует обширный класс *рассеянных* групп галактик, прошедших лишь частичную релаксацию. К их числу относится *Местная группа*, масса которой ($\sim 2 \cdot 10^{12} M_\odot$) складывается из масс двух больших спиральных галактик, имеющих собственные названия, — Млечный Путь и Туманность Андромеды, а также включает небольшую добавку из трех десятков карликовых галактик, гравитационно связанных с Млечным Путем либо с его несколько более массивной соседкой.

При сопоставлении теории с наблюдательными данными следует иметь в виду *широкий разброс* в распределении фоновой энтропии (и соответствующих внутренних профилей плотности)

¹⁾ Следует отметить, что при медленном накоплении барионов касп ТМ, если бы он существовал, не смог бы разрушиться, а только усилился. В этом смысле он *адиабатически* устойчив. Остается вопрос, какие плотности ТМ достигаются в центре гало. От этого, в частности, зависят ожидания экспериментаторов по обнаружению продуктов аннигиляции ТМ (гамма-квантов, позитронов, антипротонов и других частиц). Для ответа на этот вопрос необходимы дополнительные исследования. Идеальным объектом для изучения центральных распределений ТМ являются карликовые галактики с низкой поверхностной яркостью (рис. 3.15), в которых процесс звездообразования подавлен, а газ поддерживается в ротационном равновесии и не падает в центр гало. Исследования спиральных галактик также пока не указывают на существование центральных каспов. Все более значимую роль в восстановлении поверхностной плотности ТМ играют методы слабого гравитационного линзирования (рис. 2.12).

релаксированных гало, который реально наблюдается в галактиках (рис. 3.9) и на который прямо указывает теория космологической стандартной модели.

Действительно, для гауссовых начальных возмущений скорости функция распределения вероятности фоновой энтропии имеет вид:

$$dp(f_b) = \frac{\exp(-f_b/2)}{\sqrt{2\pi f_b}} df_b, \quad \text{где} \quad f_b \equiv \frac{E_b}{\langle E_b \rangle} \simeq \frac{s^2}{\sigma_s^2}. \quad (3.5.23)$$

Мы видим, что фоновая энтропия гало имеет широкое квазилинейное распределение вероятности и ее вариации вокруг средней величины (3.5.20) велики:

$$\langle f_b \rangle = 1, \quad \langle f_b^2 \rangle = 3 \langle f_b \rangle = 3.$$

Эти большие отклонения

$$\text{var } E_b > \langle E_b \rangle^2$$

подразумевают *значительные* отклонения от средних значений степенных индексов центральных профилей плотности в гало разных галактик.

С внутренней структурой равновесных гало тесно связана проблема кривых вращения, физические аспекты которой мы обсудим в следующем разделе.

3.6. Кривые вращения галактик

В разделе 3.3 мы рассматривали плоские кривые вращения звездных и газовых дисков, наблюдаемые во многих галактиках и свидетельствующие об изотермическом распределении ТМ внутри гало в диапазоне радиусов (масс) ~ 10 . Соответствующие степенные индексы профилей плотности и энтропии в таких гравитационно-связанных системах имеют вид ($r > r_0$, см. (3.3.9)):

$$\alpha \simeq 2, \quad \beta \simeq \frac{4}{3}.$$

Похожие распределения генерируются в аналитических моделях сферически-симметричного или квазисферического коллапса пылевой гравитирующей материи с гладким начальным распределением плотности, в котором фоновые мелкомасштабные неоднородности заведомо отсутствуют. В ходе нелинейного сжатия такого облака в его центральной области возникают каустики (самопересечения) и многопоточковые движения материи, индуцирующие процессы бурной бесстолкновительной релаксации.

В результате гравитационного перераспределения энергии между сферическими оболочками материи в центральной части гало формируются степенные профили плотности и энтропии с соответствующими интервалами индексов:

$$1,6 < \alpha_g < 2, \quad 1 < \beta_g < 1,3. \quad (3.6.1)$$

Как видно из этих соотношений, бурная релаксация приводит к распределению энтропии, которое подразумевает формирование каспа. Однако указанные индексы относятся к сравнительно большим радиусам/массам. В центральной части гало генерированная энтропия (3.6.1) не может конкурировать с фоновой и оказывается пренебрежимо малой по сравнению с ней (табл. 3.1). Следовательно, действие фоновой энтропии ограничивает центральную плотность и делает невозможным образование каспоподобных профилей гало.

Тот же вывод справедлив и для численного эксперимента, моделирующего наиболее сложные этапы формирования гало, не поддающиеся аналитическому расчету. Бурная релаксация, основанная на доминировании радиальных траекторий, — это далеко не единственный и не главный процесс релаксации бесстолкновительных систем, отсоединившихся от фридмановского потока в ходе космологического расширения.

Главная роль здесь принадлежит анизотропному коллапсу и эффектам слияния сгустков разных масс, которые образуются еще до возникновения центрального массивного образования. Все эти процессы реально наблюдаются в нелинейных релаксирующих гравитационно-связанных системах при больших красных смещениях и носят название *иерархического сгущивания* — это приливные «парные» («тройные» и пр.) взаимодействия сгустков материи, захваченных общим гравитационным полем, их рассеяние и слияние в более массивные образования с последующим повторением цикла.

С развитием методов численного моделирования с участием большого числа N взаимодействующих между собой гравитационным образом тел-частиц появилось понятие *универсального профиля плотности* гало Наварро–Френка–Уайта (НФУ), возникающего при усреднении приведенных профилей плотности сотен сгустков разных масс и размеров, образующихся в ходе слияний. Начальные распределения частиц в задаче N тел приближены к космологическим условиям, однако из-за конечности числа N динамический диапазон моделируемых масштабов *ограничен*.

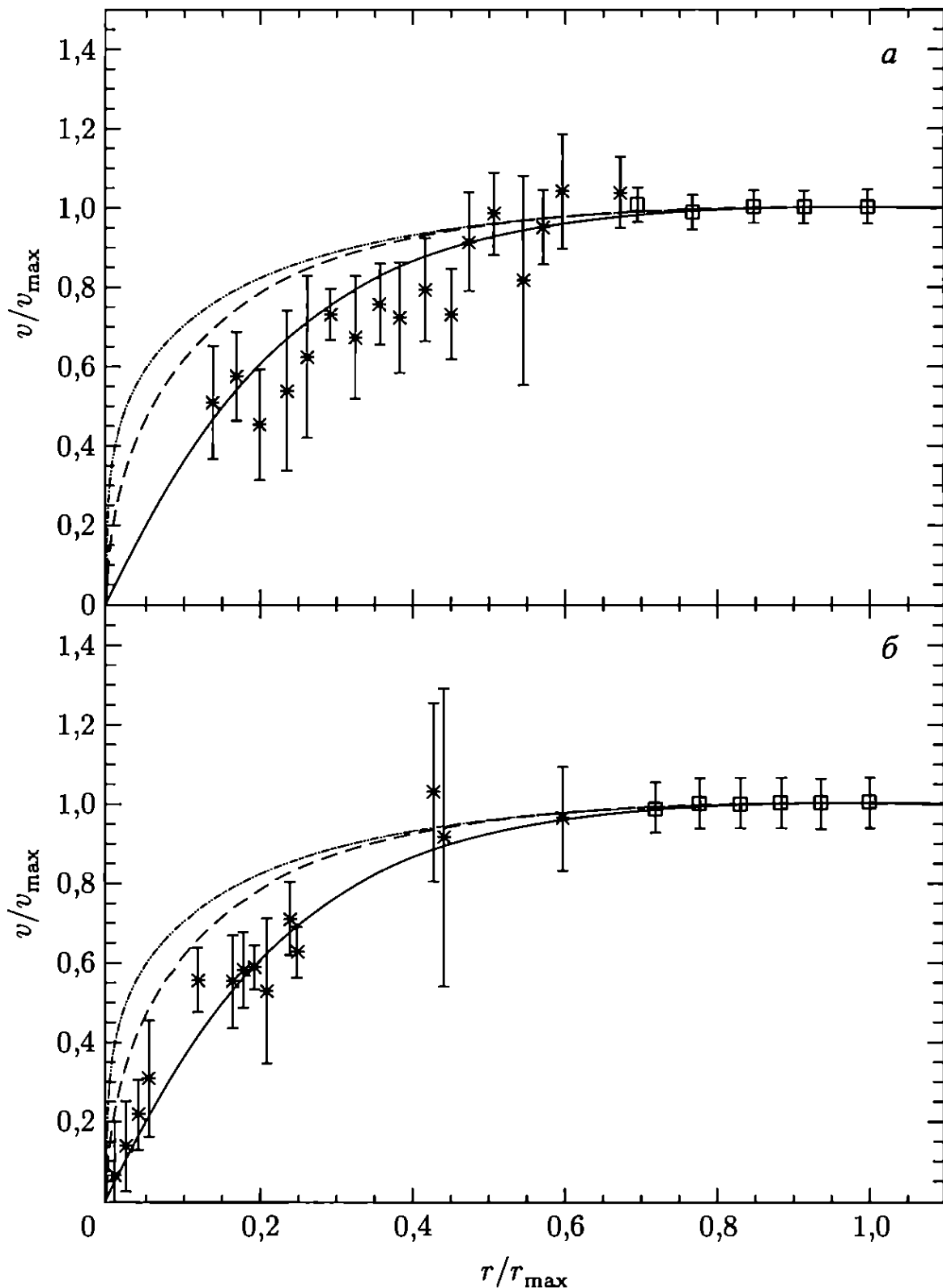


Рис. 3.15. Зависимость циркулярной скорости от радиуса в галактиках с низкой поверхностной яркостью LSB F571-8 (а) и LSB F583-1 (б) (по данным [Marchesini et al., 2002]); штриховая линия — приближение НФУ, сплошная — профиль Бюркерта

Структура профиля НФУ определяется одним параметром — характеристическим радиусом r_s , масштабирующим массу гало, и имеет вид:

$$\rho \propto \frac{1}{r(r+r_s)^2}. \quad (3.6.2)$$

При $r < r_s$ поведение профиля плотности асимптотически приближается к каспоподобному степенному закону с индексами:

$$\alpha_g = 1, \quad \beta_g = \frac{5}{6}.$$

По-видимому, при численном моделировании задачи N тел происходит недооценка фоновой энтропии мелкомасштабных возмущений в силу ограниченности пространственного разрешения метода. Можно надеяться, что с увеличением мощности численных методов эффекты фоновой энтропии смогут быть адекватно учтены. Уже сейчас появляются некоторые указания на то, что с увеличением числа N индекс α падает и становится меньше единицы, однако из-за мелкомасштабного обрезания модельного спектра мощности эффект начальной энтропии подавляется по сравнению с аналитическими оценками, приведенными в табл. 3.1.

Бесспорно, влияние фоновой энтропии на внутреннюю структуру гало и формирование центрального ядра с конечными давлением и плотностью ТМ должны быть подтверждены численным экспериментом. Тем не менее, мы уже сегодня можем проверять предсказания КСМ, измеряя скорости газа и звезд в галактиках и скоплениях.

Зависимость циркулярной скорости от радиуса связана с распределением плотности в гало. На рис. 3.15 приведены кривые вращения галактик, которые никак не описываются формулой (3.6.2), но хорошо аппроксимируются более мягким профилем:

$$\rho \propto \frac{1}{(1+x^2)(r+r_s)}, \quad x = \frac{r}{r_0}, \quad (3.6.3)$$

с $r_0 \sim r_s$. При точном равенстве $r_0 = r_s$ этот феноменологический профиль называется профилем Бюркерта. Подобные примеры типичны для галактик. В центральной области при $r < r_s$ наблюдаемая аппроксимация напоминает приближение изотермической сферы (3.3.14), однако при больших радиусах плотность падает обратно пропорционально r^3 и циркулярные скорости медленно затухают. Максимум циркулярной скорости достигается в точке локального профиля плотности с индексом $\alpha = -2$.

Существует и другая проблема, связанная с численными профилями НФУ: они не описывают плоские кривые вращения галактик в наблюдаемом диапазоне масштабов. Тем не менее, эти профили являются хорошей аппроксимацией распределения массы в богатых скоплениях галактик (рис. 3.16, 3.17).

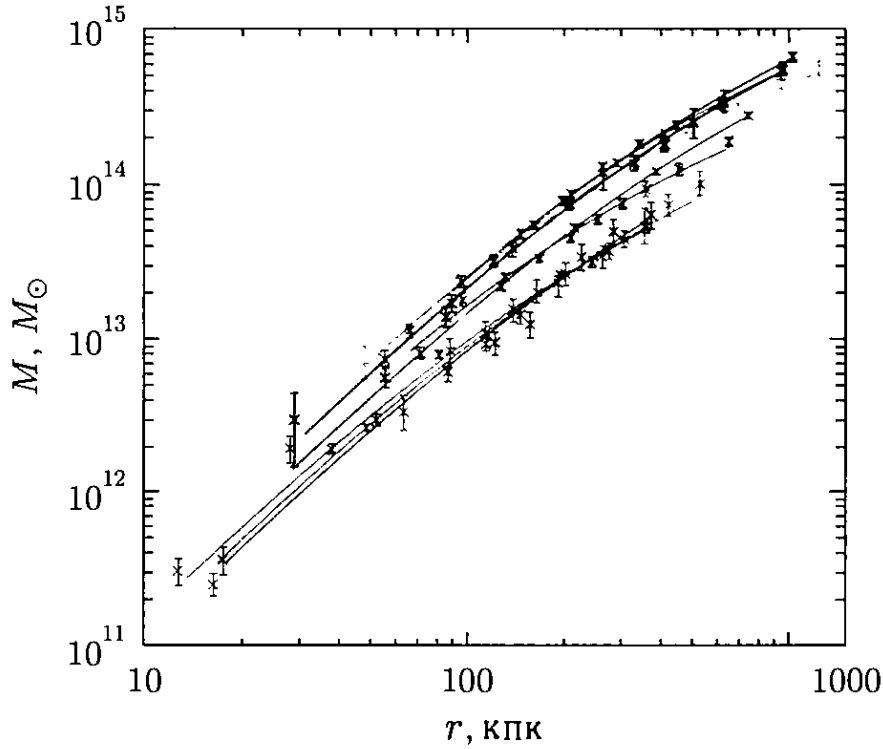


Рис. 3.16. Распределение массы темной материи внутри скоплений галактик (по данным [Pointecouteau, 2005])

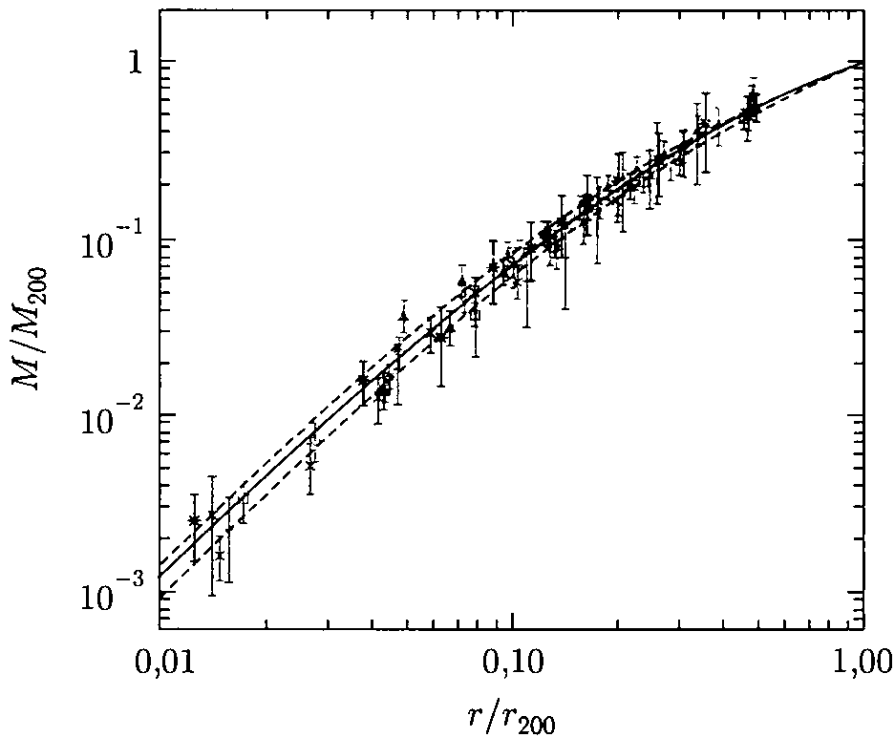


Рис. 3.17. Кривые, представленные на рис. 3.16, в переменных, нормированных на массу и радиус сферы, внутри которой средняя плотность в 200 раз превышает плотность материи во Вселенной. Сплошная линия — приближение НФУ

Чтобы понять физический смысл и многообразие наблюдаемых кривых вращения, следует принять во внимание начальную фоновую энтропию ТМ в нелинейных гало. Форма полных профилей распределений плотности и энтропии в равновесных гало

определяется суммарным действием фоновой и сгенерированной энтропий частиц ТМ.

Для пояснения физики этого воздействия рассмотрим простейшую модель объединенной энтропийной функции масс в следующем виде:

$$E(M) = \sqrt{E_b^2 + E_g^2}, \quad f_b = \frac{E_b}{E} \Big|_{r_{\max}}, \quad (3.6.4)$$

где константа $f_b \in (0, 1)$ задает относительный вклад фоновой энтропии на фиксированном внутреннем радиусе гало $r = r_{\max}$, соответствующем максимальной циркулярной скорости или локальному индексу наклона плотности $\alpha = -2$. Функции $E_{b,g}(M)$ имеют степенной вид с соответствующими индексами наклонов в следующих интервалах значений:

$$\beta_b \in \left(\frac{1}{3}, \frac{2}{3}\right), \quad \beta_g \in \left(\frac{5}{6}, \frac{4}{3}\right). \quad (3.6.5)$$

Теперь мы можем аналитически рассчитать кривые вращения частиц в равновесных гало ТМ и сравнить их с данными наблюдений и численного моделирования.

Результаты представлены на рис. 3.18. Скорости вращений $v_{\text{tot}} = v(r)$ как функции внутреннего радиуса r получены для моделей иерархической ($\beta_g = 5/6$, рис. 3.18а) и бурной ($\beta_g = 4/3$, рис. 3.18б) релаксаций для двух предельных значений $f_b \ll 1$ и $f_b \simeq 1$ и трех значений индексов фоновой энтропии. Скорости и радиусы нормированы на соответствующие значения максимальной скорости и радиуса r_{\max} , на котором она достигается: $v_{\max} \equiv v(r_{\max}) \geq v(r)$. Как видно из рисунка, для гало с $\beta_b < 0,5$ и любых параметров β_g и f_b в отведенных им интервалах изменений, индуцированные кривые вращения полностью перекрывают область между аппроксимациями НФУ и Бюркerta. При $f_b \ll 1$ кривые сдвигаются в область профилей НФУ, а при увеличении f_b — в сторону наблюдаемых профилей Бюркerta.

Таким образом, наблюдаемые кривые вращения успешно воспроизводятся простыми моделями (3.6.4) с соответствующим подбором параметров f_b и β_b . Разброс распределений фоновой (3.5.23) и генерированной (3.6.5) энтропий в гало способен объяснить требуемое разнообразие форм наблюдаемых кривых вращения. При увеличении индекса $\beta_b > 0,5$ кривые вращения концентрируются ближе к профилям НФУ. По-видимому, это указывает на то, что в карликовых галактиках с $M < 10^8 M_{\odot}$

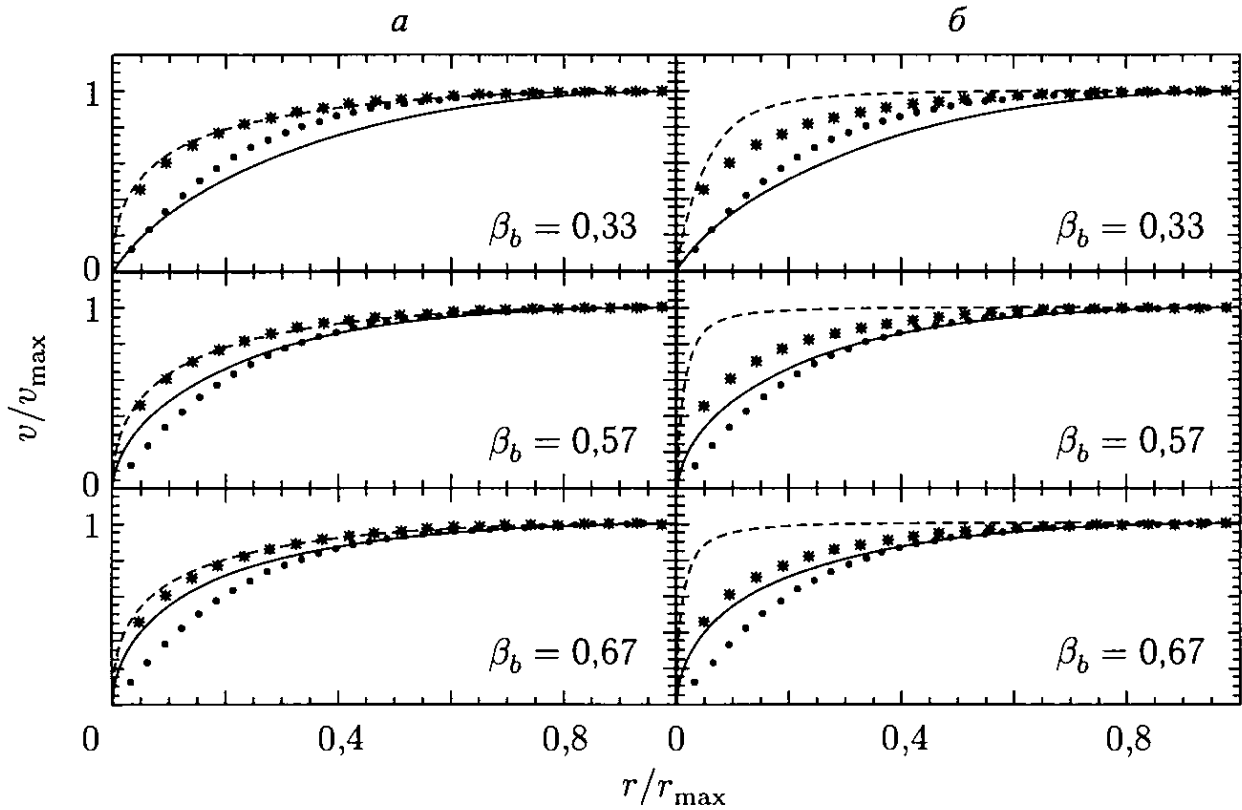


Рис. 3.18. Нормированные кривые вращений $v(r)$ для моделей иерархической ($\beta_g = 5/6$, *a*) и бурной ($\beta_g = 4/3$, *б*) релаксаций при $f_b \ll 1$ (штриховая линия) и $f_b \simeq 1$ (сплошная линия). Аппроксимации НФУ и Бюркerta обозначены соответственно звездочками и точками

чаще могут реализовываться каспоподобные распределения ТМ (см. табл. 3.1).

Рассмотренные примеры доказывают следующее.

- Фоновая энтропия препятствует образованию центральных каспов ТМ в гало с массами в интервале $10^8 - 10^{14} M_\odot$. Для гало меньших и бóльших масс влияние фоновой энтропии ослаблено.
- Учет влияния фоновой энтропии на распределение плотности внутри гало позволяет воспроизвести наблюдаемые кривые вращения в широком диапазоне масштабов и способен решить проблему каспов в рамках КСМ.

В заключение заметим, что рассмотренные нами модели основаны на простых предположениях о распределении частиц ТМ в гало в пренебрежении барионной компонентой вещества. Полная задача об эволюции галактических профилей плотности, включающая ТМ и барионы и учитывающая влияние различных диссипативных процессов на перенос углового момента из центральной части галактик и др., безусловно, очень сложна и требует дальнейших исследований. В том числе и с помощью численного моделирования.

КОСМОЛОГИЧЕСКИЕ МОДЕЛИ

Следующий пример гравитирующей системы, который мы рассмотрим, — это сама Вселенная. Окружающий нас мир существует конечное время. Он возник в результате катастрофического гравитационного явления — Большого взрыва — и находится в процессе релаксации, эволюционируя во времени. Эта система неустойчива, она переходит из одного состояния в другое, и корни этой неустойчивости — в природе гравитации.

В этой главе мы рассмотрим космологические модели, обладающие пространственной симметрией (группой движений), и выясним, что ранняя Вселенная описывается одной из таких простейших геометрий со слабонарушенной группой Фридмана.

4.1. Фридмановская геометрия

Источником геометрии Фридмана является однородный и изотропный поток материи. Для того чтобы организовать столь симметричное движение без вращения и сдвиговой деформации, необходимо стартовать со специальных начальных условий для пространственного распределения материи и поля скоростей (задача Коши). В силу заданной симметрии материальная среда движется свободно вдоль геодезического пучка параллельных прямых и пространственные градиенты давления не возникают. Динамика течения зависит от состояния материи (давления и плотности энергии) и геометрического параметра k , задающего соотношение между кинетической энергией потока и его гравитационной энергией связи.

Пусть u^α — касательный вектор вдоль линий тока среды, и $u^\alpha u_\alpha = 1$. Тогда аффинным параметром является собственное время t , а сечения $t = \text{const}$ представляют собой изотропное пространство:

$$ds^2 = dt^2 - a^2 dl^2, \quad u_\alpha = t_{,\alpha} = (1, 0, 0, 0), \quad (4.1.1)$$

где $d\ell$ описывает статическое (сопутствующее) пространство постоянной кривизны ¹⁾ $\kappa = \text{const}$:

$$\begin{aligned} d\ell^2 &= dx^2 + \kappa^{-1}dX^2 = d\chi^2 + r^2d\Omega = \\ &= \frac{dr^2}{1 - \kappa r^2} + r^2d\Omega = \frac{d\bar{r}^2 + \bar{r}^2d\Omega}{(1 + \kappa\bar{r}^2/4)^2}, \end{aligned} \quad (4.1.2)$$

где

$$r = \frac{\sin(\sqrt{\kappa}\chi)}{\sqrt{\kappa}} = \frac{\bar{r}}{1 + \kappa\bar{r}^2/4},$$

$d\Omega \equiv de^2 = d\theta^2 + \sin^2\theta d\varphi^2$ — квадрат интервала на поверхности единичной сферы ($|e| = 1$, ср. (2.2.1)), координаты (t, \mathbf{x}) называются фридмановскими, а (t, r, θ, φ) — координатами Робертсона–Уокера. В геометрии (4.1.1) нарушена лоренцева симметрия, однако метрика инвариантна относительно шестипараметрической группы пространственных движений (4.1.2).

Однородность статического пространства очевидна в координатах \mathbf{x} , представляющих интервал (4.1.2) на изотропной сферической ($\kappa > 0$) или псевдосферической ($\kappa < 0$) гиперповерхности в некотором вспомогательном евклидовом 4-пространстве (\mathbf{x}, X) :

$$\kappa\mathbf{x}^2 + X^2 = 1.$$

От декартовых координат \mathbf{x} можно перейти к сферическим с центром в произвольной точке \mathbf{x}_0 (совмещенной с нулем) и единичным вектором \mathbf{e} с помощью линейного соотношения:

$$\mathbf{x} = r\mathbf{e}. \quad (4.1.3)$$

Эффективный тензор энергии-импульса в геометрии (4.1.1) определяется из уравнений Эйнштейна (1.8.13):

$$T_{\alpha}^{\beta} = (\rho + p)u_{\alpha}u^{\beta} - p\delta_{\alpha}^{\beta} = \text{diag}(\rho, -p, -p, -p), \quad (4.1.4)$$

где функции времени ρ и p соответствуют сопутствующим плотности энергии и эффективному давлению материи. Условие (4.1.4) никак не ограничивает состав и физические свойства

¹⁾ Поскольку физической величиной является отношение κ/a^2 , то формально важен только знак κ , что позволяет ограничиться рассмотрением значений $\kappa = 0, 1, -1$. Однако в этом случае a имеет смысл радиуса кривизны пространства (4.1.1), и мы лишены возможности выбора произвольной нормировки для a . Более удобно считать, что a есть масштабный фактор, подлежащий отдельной нормировке. Обычно принимают, что в современной Вселенной физические и сопутствующие масштабы совпадают, то есть $a(t_0) = 1$. Тогда пространственная кривизна равна κa^{-2} , а $|\kappa|^{-1/2}$ есть радиус кривизны.

материи: оно говорит лишь о том, что все компоненты среды распределены однородно и изотропно в пространстве и движутся вместе с общим потоком. Тождества Бианки обеспечивают сохранение энергии при движении гравитирующей материи:

$$\dot{\rho} + 3H(\rho + p) = 0, \quad H = \frac{1}{3} u^\alpha{}_{;\alpha} = \frac{\dot{a}}{a}, \quad (4.1.5)$$

а динамика потока описывается уравнением Фридмана (G_0^0 -компонента (1.8.13)), которое связывает темп объемного расширения материи (функцию Хаббла $H = H(t)$) с ее плотностью энергии:

$$H^2 = \frac{8\pi G}{3} \rho - \frac{\kappa}{a^2}. \quad (4.1.6)$$

Отсюда получаем для близких частиц среды, находящихся на расстоянии ax друг от друга ($r \ll |\kappa|^{-1/2}$), знакомое выражение для ньютоновской энергии связи:

$$\frac{V_H^2}{2} - \frac{GM}{ar} = -\frac{\kappa r^2}{2} = \text{const}(t), \quad (4.1.7)$$

где локальная хаббловская скорость потока

$$V_H = \frac{\partial}{\partial t} (ax) = \dot{a}x = \bar{H}x, \quad (4.1.8)$$

$$\bar{H} \equiv aH = \dot{a} = \frac{a'}{a}, \quad M = \frac{4\pi}{3} \rho a^3 r^3$$

— масса в сфере радиуса r . Из (4.1.7) следует, что интеграл движения κ имеет смысл удельной энергии связи частиц в потоке гравитирующей среды.

Мы получили закон Хаббла (4.1.8) из условий однородности и изотропии пространства. Верно и обратное: расширение материи по закону Хаббла свидетельствует о пространственной однородности и изотропии, то есть об отсутствии центра расширения и независимости скорости разлета от направления. Поэтому однородный и изотропный поток материи называют *хаббловским потоком* (подробнее см. раздел 10.3).

Зависимость V_H от времени прямо связана со знаком второй производной масштабного фактора или, эквивалентно, с величинами безразмерных факторов $\bar{\gamma}$ или γ по отношению соответ-

ственно к нулю или единице:

$$\begin{aligned}\bar{\gamma} &\equiv -\frac{\bar{H}'}{\bar{H}^2} = -\frac{d \ln \bar{H}}{d \ln a} = -\frac{a\ddot{a}}{\dot{a}^2} = \frac{\Omega_\rho}{2} \left(1 + \frac{3p}{\rho}\right), \\ \gamma &\equiv -\frac{\dot{H}}{H^2} = -\frac{d \ln H}{d \ln a} = 1 + \bar{\gamma}, \\ \gamma_\rho &\equiv \frac{3}{2} \left(1 + \frac{p}{\rho}\right) = 1 + \frac{\bar{\gamma}}{\Omega_\rho}, \quad \Omega_\rho \equiv \frac{8\pi G\rho}{3H^2}.\end{aligned}\tag{4.1.9}$$

Хаббловская скорость возрастает в ускоряющейся (*инфляционной*) Вселенной ($\ddot{a} > 0$ или $\bar{\gamma} < 0$, $\gamma < 1$, $\gamma_\rho < 1$) и уменьшается в замедляющейся Вселенной ($\ddot{a} < 0$ или $\bar{\gamma} > 0$, $\gamma > 1$, $\gamma_\rho > 1$). С помощью параметра состояния материи γ_ρ уравнение (4.1.5) можно переписать в более лаконичной форме:

$$\frac{\dot{\rho}}{\rho} + 2H\gamma_\rho = 0.\tag{4.1.10}$$

Мы видим, что гравитация обладает в равной мере и свойством притяжения, и свойством отталкивания. Для однородного потока материи все зависит от соотношения между ее эффективным давлением и плотностью энергии. Наше представление о гравитации как о всемирном *тяготении* основано на опыте с нерелятивистскими материальными источниками ($|p| \ll \rho$). Когда же давление (пространственная компонента тензора энергии-импульса материи) сопоставимо по величине с плотностью энергии и отрицательно по знаку ($p < -\rho/3$), то гравитационное взаимодействие проявляется как всемирное *отталкивание*. В этом нет никакой экзотики, так устроена гравитация.

Уравнение Фридмана имеет еще один интересный аспект. В окрестности произвольного хаббловского наблюдателя, движущегося вместе с потоком в точке $\mathbf{x} = 0$ своих координат \mathbf{x} , уравнение (4.1.6) можно переписать в форме уравнения Пуассона:

$$\frac{\Delta\Phi}{a^2} = 4\pi G\rho,\tag{4.1.11}$$

$$\Phi = \frac{1}{4} \left(\bar{H}^2 + \kappa \right) \mathbf{x}^2 = \frac{2\pi G}{3} \rho a^2 \mathbf{x}^2,$$

где $\Delta \equiv \delta^{ik} \frac{\partial^2}{\partial x^i \partial x^k}$ — пространственный лапласиан в окрестности наблюдателя. Данное уравнение и по форме, и по содержанию

является релятивистским обобщением уравнения Пуассона для изотропной среды ¹⁾.

Для определения физического смысла потенциала Φ построим локально-инерциальную систему отсчета в окрестности мировой линии хаббловского наблюдателя. Это всегда можно сделать, поскольку материя движется свободно в хаббловском потоке. Связь между эйлеровыми (τ, y) и фридмановскими (t, \mathbf{x}) координатами находим в виде разложения по степеням \mathbf{x}^2 . В главном порядке имеем:

$$\tau = t + \frac{a\dot{a}}{2} \mathbf{x}^2, \quad y = a\mathbf{x}(1 + \Phi), \quad (4.1.12)$$

$$(1 + 2\Psi) d\tau^2 - (1 - 2\Phi) dy^2 = dt^2 - a^2 \left(d\mathbf{x}^2 + \kappa (\mathbf{x}d\mathbf{x})^2 \right),$$

где

$$\Psi = \frac{\bar{\gamma}}{2} \bar{H}^2 \mathbf{x}^2 = \left(1 + \frac{3p}{\rho} \right) \Phi$$

— гравитационный потенциал локально-инерциального наблюдателя, характеризующий гравитационное воздействие на его окрестность однородного хаббловского потока.

Таким образом, Ψ и Φ — это гравитационные потенциалы в окрестности хаббловского наблюдателя. В ньютоновском пределе ($|p| \ll \rho$) получаем $\Psi = \Phi$ в соответствии с (1.10.4). Сохраняя преемственность с ньютоновским гравитационным потенциалом Φ , мы оставляем эту букву для обозначения *пространственного* потенциала релятивистской метрики (4.1.12), поскольку именно Φ , а не временной потенциал Ψ , входит в обобщенное уравнение Пуассона (4.1.11).

4.2. Инфляция и симметрия де Ситтера

Рассмотрим подробнее, как выглядит хаббловский поток с точки зрения локально-инерциальных наблюдателей.

Инерциальные лоренцевы наблюдатели фиксируют геометрию потока, измеряя его физические характеристики, например плотность или давление. Плотность материи, измеренная в собственной системе отчета наблюдателя, равна (см. (4.4.2)):

$$\rho_{\Gamma} \equiv v^{\alpha} v^{\beta} T_{\alpha\beta} = \rho \left[1 + \frac{2}{3} \gamma_{\rho} (\Gamma^2 - 1) \right], \quad (4.2.1)$$

¹⁾ Ср. (1.8.1). В точном виде лапласиан в уравнении (4.1.11) следует писать в ковариантной форме: $\Phi_{;i}^i = 4\pi G\rho a^2$, где ковариантные производные берутся в метрике (4.1.2). Это уравнение тождественно уравнению Фридмана.

где v^α — скорость наблюдателя, а $\Gamma \equiv v_\alpha u^\alpha \geq 1$ есть Γ -фактор его движения относительно хаббловского потока.

В замедляющейся Вселенной ($\gamma > 1$) лоренцевы наблюдатели измеряют разную плотность в зависимости от Γ и таким образом могут идентифицировать гиперповерхности однородности. В инфляционной Вселенной ($\gamma < 1$) происходит частичное восстановление лоренцевой симметрии: наблюдатели, находящиеся в конусе скоростей

$$\Gamma^2 \ll 1 + \gamma_\rho^{-1}, \quad (4.2.2)$$

обнаруживают одну и ту же плотность потока $\rho_\Gamma = \rho$. Можно сказать, что в конусе (4.2.2) происходит размывание фридмановской геометрии. В предельном случае $\gamma_\rho \rightarrow 0$ плотность материи неизменна (см. (4.1.5), (4.1.9)) и лоренцева симметрия полностью восстанавливается:

$$\rho = -p \equiv \frac{3\lambda}{8\pi G} = \text{const}, \quad a = a_\kappa(t) = \frac{e^{\sqrt{\lambda}t} + \kappa e^{-\sqrt{\lambda}t}}{2\sqrt{\lambda}}. \quad (4.2.3)$$

Координаты (4.1.1), (4.2.3) представляют собой одно из разбиений геометрии де Ситтера — пространства-времени постоянной положительной кривизны $\lambda \equiv \Lambda/3 > 0$:

$$ds^2 = (1 - \lambda R^2) dT^2 - \frac{dR^2}{1 - \lambda R^2} - R^2 d\Omega = \quad (4.2.4)$$

$$= dt^2 - a^2 \left(\frac{dr^2}{1 - \kappa r^2} + r^2 d\Omega \right), \quad (4.2.5)$$

$$R = ar, \quad T = \frac{1}{2\sqrt{\lambda}} \ln \left[\kappa \text{cth} \left(\frac{1}{2} \ln \sqrt{\frac{1 - \kappa r^2}{1 - \frac{\kappa}{\lambda a^2}}} \right) \right].$$

Из сравнения с (4.1.2), (4.1.12) получаем:

$$\Psi = -2\Phi = -\frac{\lambda}{2} R^2.$$

Пространство-время де Ситтера инвариантно относительно 10-параметрической группы движений (аналог группы Пуанкаре для $\lambda \neq 0$). Это становится очевидным после представления $T_\alpha^\beta = \rho \delta_\alpha^\beta$, а также интервала (4.2.4) в виде

$$ds^2 = dT^2 - dR^2 - d\mathcal{R}^2 \quad (4.2.6)$$

в некотором игрушечном псевдоевклидовом 5-мерном пространстве на изотропной псевдосферической гиперповерхности

$$R^2 + \mathcal{R}^2 - T^2 = \lambda^{-1}.$$

Обратный переход к координатам T и R ($< \lambda^{-1}$) осуществляется с помощью преобразований

$$R = Re, \quad \frac{T}{\mathcal{R}} = \text{th}(\sqrt{\lambda}T).$$

Повороты в фиктивном 5-мерном пространстве (4.2.6) задают 10 параметров преобразований координат (T, R) , сохраняющих интервал (4.2.4). Мы видим, что геометрия де Ситтера определяется только плотностью энергии λ . Разные фридмановские разбиения (t, \mathbf{x}) , отвечающие разным $\varkappa = \text{const}$ и функциям $a_{\varkappa}(t)$ своего координатного времени, определяют одно и то же пространство-время постоянной положительной кривизны λ .

Очевидно, ни о какой инфляции и вообще эволюции в пространстве де Ситтера речи быть не может, так как все наблюдатели находятся в одинаковых условиях и движения среды (4.2.3) неизмеримы¹⁾. Говорить о временной эволюции можно лишь в отношении геометрий с выделенной (по внутренней симметрии, начальным и граничным условиям) системой наблюдателей со своим пространственно-временным разбиением, которые измеряют характеристики пространственных сечений $t = \text{const}$ и их изменения вдоль мировых линий собственного времени.

¹⁾ В геометрии (4.2.4) имеются как расширяющиеся, так и сжимающиеся однородные пространства (4.2.5). Так, перенормируя пространственные переменные $\tilde{r} = \varkappa r$, $\tilde{a} = \varkappa^{-1} a$ и $\tilde{\varkappa} = \varkappa^{-1}$, получаем «обратную эволюцию» сечений $t = \text{const}$: $\tilde{a} = a_{\tilde{\varkappa}}(-t)$, $\tilde{H} = -H$. Можно показать, что вырождаются не только однородные, но и неоднородные расщепления (движения). Например, при $\gamma_\rho \rightarrow 0$ произвольные малые скалярные деформации плоской модели Фридмана исчезают в эйлеровом представлении ($\Phi = \Psi \rightarrow 0$, см. раздел 5.1):

$$ds^2 = \left(1 - \frac{2}{\sqrt{\lambda}} \dot{q}\right) dt^2 - a^2 (\delta_{ij} - 2\mathcal{B}_{ij}) dx^i dx^j = dt_E^2 - a^2 dy^2,$$

где $t_E = t - \lambda^{-1/2} q$, $\mathbf{y} = \mathbf{x} - \nabla \mathcal{B}$, $\mathcal{B} = \lambda^{-1/2} \int a^{-2} q dt$, $a = a_0(t_E) = a_0(t) \cdot (1 - q)$, $H = \sqrt{\lambda}$, $q = q$ — произвольная малая функция пространственно-временных координат (при выборе $q = \varkappa x^2/2$ получаем однородное расщепление (4.2.5)). Также при $\gamma_\rho = 0$ вырождаются и вихревые деформации: при конечных v_i получаем из (4.5.9) $\mathcal{G}_i = 0$. Вырождение полей q и v_i в модели де Ситтера объясняется невозможностью различить движение среды в состоянии (4.2.3). Тем не менее, отбрасывая и не рассматривая эти поля, можно прямо определять геометрические отклонения от метрики де Ситтера во всех трех модах (Φ , \mathcal{G}_i и \mathcal{G}_{ij}).

Именно в этом смысле мы представляем космологическую эволюцию хаббловских потоков с выделенной фридмановской симметрией. Количественной мерой эволюции является параметр γ . Однако при $\gamma < 1$ возникает *де-ситтеровский конус скоростей* (4.2.2): хаббловские и другие наблюдатели, движущиеся внутри этого конуса, фиксируют восстановление де-ситтеровской группы и исчезновение самого процесса инфляции как эволюционного процесса, способного привести к каким бы то ни было изменениям.

При каком значении параметра γ происходит утрата фридмановской симметрии и исчезновение инфляции?

При ответе на этот вопрос хотелось бы иметь идеальные инструменты, имеющиеся в распоряжении наблюдателей для измерения плотности и других космологических параметров. Однако точность измерения всегда будет ограничена квантовым пределом в силу принципа неопределенности Гейзенберга. Таким образом, при ответе на поставленный вопрос надо учитывать квантовую природу материи и эффекты квантовой гравитации. К счастью, это всегда можно сделать в квазиклассическом пределе при $H \ll M_P$.

Из формул (4.2.4) следует, что статическая R -область пространства де Ситтера вокруг центрального наблюдателя простирается до его горизонта событий $R = \lambda^{-1/2}$, а далее лежит T -область — зона метрики, нестационарная по отношению к свободному наблюдателю, находящемуся в точке $R = 0$. Относительно него частицы в далекой T -области имеют отрицательную полную энергию из-за гравитационной энергии связи, что вызывает квантово-гравитационный параметрический эффект рождения полей из вакуума, похожий на тот, что имеет место в черных дырах.

Однако, в отличие от черных дыр, сгенерированная гравитацией энергия физических полей не может в этом случае уйти в R -области на бесконечность и реализовать себя в виде частиц в волновой зоне, поскольку размер самой R -зоны конечен, а типичная длина волны возмущений порядка этого размера. В результате окрестность наблюдателя заполнена флуктуационными полями с характерной фоновой энергией возмущений и плотностью энергии соответственно:

$$T_\lambda = \frac{\sqrt{\lambda}}{2\pi} \simeq \frac{H}{2\pi}, \quad \rho_\lambda = g_\lambda \lambda^2 \simeq g_\lambda H^4, \quad (4.2.7)$$

где значение префактора $g_\lambda \gtrsim 1$ определяется эффективным числом степеней свободы частиц, слабо зависящим от величины H .

В приближении метрики (4.2.4) этот эффект стационарен и называется эффектом Гиббонса–Хокинга: спектр флуктуаций имеет равновесный вид с температурой $T_\lambda = \text{const}$ (термостат).

Эффект Гиббонса–Хокинга ограничивает снизу эволюционный параметр γ фридмановской геометрии:

$$\gamma \geq \gamma_\lambda \simeq \frac{\rho_\lambda}{\rho} \sim \left(\frac{H}{M_P} \right)^2. \quad (4.2.8)$$

При $\gamma < \gamma_\lambda$ фридмановская симметрия утеряна (необнаружима) и мы имеем дело с флуктуирующей моделью де Ситтера. Для современной Вселенной этот эффект пренебрежимо мал: $\gamma \simeq 0,4$, тогда как $\gamma_\lambda \sim 10^{-122}$. Однако в период Большого взрыва температура T_λ могла быть сопоставима с энергией Великого объединения, где фридмановская группа может быть размыта.

Для примера рассмотрим фридмановскую эволюцию скалярного поля φ с массой m . На инфляционной сепаратрисе имеем (см. разделы 7.4, 9.3):

$$H \simeq \frac{2m\varphi}{M_P}, \quad \gamma \simeq \left(\frac{M_P}{4\varphi} \right)^2, \quad \gamma_\lambda \simeq \left(\frac{m\varphi}{5M_P^2} \right)^2.$$

Сравнивая параметры γ (см. (4.2.8)), получаем область значений поля φ для космологической инфляции:

$$1 \lesssim \frac{\varphi}{M_P} \lesssim \sqrt{\frac{M_P}{m}}.$$

При больших полях $\varphi > M_P^{3/2} m^{-1/2}$ мы имеем геометрию постоянной положительной кривизны с квантовыми флуктуациями поля φ . При $m \sim 10^{-6} M_P$ получаем оценку температуры на нижнем пределе де-ситтеровской кривизны:

$$T_\lambda \sim \sqrt{mM_P} \sim 10^{16} \text{ ГэВ}.$$

Этот пример показывает, что процесс инфляционной эволюции в геометрии Фридмана имеет *промежуточный* характер, ограниченный по величине параметра γ как сверху ($\gamma < 1$ по определению инфляции), так и снизу ($\gamma > \gamma_\lambda$).

Пока неизвестно, возникла ли наша Вселенная в результате нарушения де-ситтеровской симметрии и если это так, то каким образом этот переход мог произойти. Более естественным нам видится прямо противоположный процесс, когда квазифридмановская Вселенная создается в ходе локально-инфляционного

расширения из произвольно распределенной материи, а не является осколком более симметричной структуры (например, де-ситтеровской). Процесс построения упорядоченного хаббловского потока мы рассмотрим в последующих главах, а сейчас вернемся к возможности перехода от де-ситтеровской симметрии к фридмановской ($dS \rightarrow F$).

Если такой переход с понижением симметрии и имел место в прошлом, то в нем, по-видимому, центральную роль сыграло изменение состояния материи под воздействием квантовых флуктуаций и свойства самой гравитации. Можно проследить, как в этом случае эволюционировали средние космологические параметры.

С учетом (4.2.7) при $R_{\kappa} > \bar{H}^{-1}$ уравнение Фридмана модифицируется:

$$H^4 - 2m_P^2 H^2 + \bar{\rho} = 0,$$

где введены вспомогательные обозначения $m_P = 0,24 \cdot M_P / g_{\lambda} \sim \sim 10^{18}$ ГэВ и $\bar{\rho} = \rho / g_{\lambda}$, а $R_{\kappa} \equiv |\kappa|^{-1/2}$ — радиус пространственной кривизны. Отсюда получаем искомую ветвь решения с ограниченными плотностью и кривизной $H > 0$:

$$m_P^2 - H^2 = \sqrt{m_P^4 - \bar{\rho}} > 0. \quad (4.2.9)$$

Рассмотрим параметр $\gamma_{\rho} \equiv 3/2 (1 + p/\rho)$ как меру отклонения состояния материи от де-ситтеровского (см. (4.1.9)). При $\gamma_{\rho} \simeq 0$ мы имеем фоновое пространство-время постоянной положительной кривизны (ср. (4.2.3)). Пусть в некоторый момент t_1 состояние материи изменилось и параметр γ_{ρ} стал положительным. Тогда при $t \geq t_1$ из (4.1.10) и (4.2.9) получаем:

$$\dot{\rho} = -2H\gamma_{\rho}\rho < 0, \quad \dot{H} = -\frac{\gamma_{\rho}\bar{\rho}}{2\sqrt{m_P^4 - \bar{\rho}}} < 0, \quad (4.2.10)$$

то есть состояние де Ситтера *неустойчиво* по отношению к необратимому фридмановскому расширению с уменьшением H и ρ . Чтобы такой переход действительно произошел, необходимо, чтобы параметр γ_{ρ} превысил пороговое значение γ_{λ} , иначе следующая флуктуация вернет развивающееся решение (4.2.10) в стационарное де-ситтеровское состояние.

Как геометрически описать фазовый переход $dS \rightarrow F$?

Продемонстрируем это на примере плоской геометрии Фридмана. При $\kappa = 0$ из (4.2.5) получаем следующее соотношение

между временем центрального наблюдателя и фридмановскими координатами:

$$\tau \equiv -2e^{-2\sqrt{\lambda}T} = \frac{1}{2} \left(r^2 - \frac{1}{\lambda a^2} \right).$$

С помощью новой переменной τ метрику (4.2.4) можно переписать в следующем виде:

$$ds^2 = \frac{\lambda a^4 d\tau^2 - dR^2}{1 - \lambda R^2} - R^2 d\Omega, \quad (4.2.11)$$

где масштабный фактор a остается функцией времени t (см. (4.2.3)). В аналогичной форме можно представить и метрику Фридмана:

$$\frac{(a\bar{H}d\tau)^2 - dR^2}{1 - H^2 R^2} - R^2 d\Omega = a^2 (d\eta^2 - dr^2 - r^2 d\Omega), \quad (4.2.12)$$

где преобразование к новым координатам имеет вид:

$$\tau = \frac{r^2}{2} + \int \frac{d\eta}{H}, \quad R = ar.$$

Таким образом, непрерывный переход с нарушением dS-симметрии описывается с помощью метрики (4.2.12) при трансформации постоянной кривизны $H = \sqrt{\lambda} = \text{const}$ в функцию времени $H(t)$, при этом масштабный фактор $a = a(t)$ определяется из уравнения Фридмана.

Заметим, что, несмотря на внешнее сходство между выражениями (4.2.11) и (4.2.12), эффект Гиббонса–Хокинга неприменим при $H \neq \text{const}$, поскольку отсутствует понятие термостата T_λ . Для описания квантово-гравитационного рождения возмущений в модели Фридмана необходимо пользоваться нестационарным формализмом (см. главы 5–8).

4.3. Однородные пространства

Модели Фридмана высокосимметричны, они допускают 6-параметрическую группу движений: 3 трансляции и 3 поворота в пространстве. Это свойство позволяет представить все скаляры модели (a, H, ρ, p, \dots) в виде функций одной координаты — собственного времени t . На самом деле последнее утверждение связано только с однородностью, предполагающей 3-параметрическую группу трансляций. В этом разделе мы

рассмотрим общие космологические модели с однородными пространствами.

Однородность предполагает существование группы движений — как минимум 3-параметрической в силу трехмерности пространства, — совмещающих пространство с самим собой. Если в плоском пространстве 3 трансляции оставляют неизменными 3 дифференциала dx^1 , dx^2 и dx^3 , то в искривленном пространстве инвариантными остаются 3 дифференциальных формы $e_i^p dx^i$, где $p = 1, 2, 3$. Это означает, что существуют преобразования пространственных координат $x'^i = x'^i(x^j)$, сохраняющих эти формы:

$$e_i^p(x^j) dx^i = e_i^p(x'^j) dx'^i, \quad (4.3.1)$$

причем трансляционные векторные поля e_i^p в обеих частях равенства — одни и те же функции своих координат.

Метрический тензор всегда можно представить в тетрадном виде:

$$g_{\alpha\beta} = e_\alpha^a e_\beta^b \eta_{ab}, \quad (4.3.2)$$

где $a = 0, p$. Для модели Фридмана удобно выбрать:

$$e_\alpha^0 = u_\alpha = \delta_\alpha^0, \quad e_\alpha^p = a \delta_\alpha^p.$$

Для пространственно однородных космологий можно считать $e_\alpha^0 = \delta_\alpha^0$ и ограничиться нахождением оставшейся триады векторов e_α^p . После определения репера e_i^p из уравнений (4.3.1) имеем:

$$ds^2 = dt^2 - \gamma_{pq} (e_i^p dx^i) (e_j^q dx^j) = dt^2 - \gamma_{pq} e_i^p e_j^q dx^i dx^j, \quad (4.3.3)$$

где $\gamma_{pq} = \gamma_{pq}(t)$ — симметричная матрица масштабных факторов.

Для построения обратного метрического тензора введем тройку обратных векторов e_p^i и обратную матрицу γ^{pq} :

$$e_p^i e_j^p = \delta_j^i, \quad e_p^i e_i^q = \gamma_{pr} \gamma^{rq} = \delta_p^q. \quad (4.3.4)$$

Все пространственные геометрические объекты в однородных космологических моделях также факторизуются на произведения реперов e_i^p и соответствующих функций времени. Например:

$$u_i = e_i^p u_p(t), \quad T_i^0 = e_i^p T_p^0(t), \quad T_i^j = e_i^p e_j^q T_p^q(t).$$

Вернемся к решениям уравнений (4.3.1).

Условия интегрируемости (4.3.1) относительно функций $x'^i = x'^i(x^k)$ приводят к тому, что следующие комбинации должны сводиться к постоянным числам:

$$C_{pq}^r \equiv e_p^i e_q^j \left(\frac{\partial e_i^r}{\partial x^j} - \frac{\partial e_j^r}{\partial x^i} \right) = \text{const}, \quad (4.3.5)$$

называемым *структурными постоянными* группы движений. Эти условия интегрируемости и есть условия существования однородных пространств. Всего известны девять не приводимых друг к другу типов однородных пространств по классификации Бианки. В дальнейшем нам понадобятся только первые два из них.

Космологическая модель первого типа обладает плоским пространством ($C_{pq}^r = 0$):

$$ds^2 = dt^2 - a^2 dx^2 - b^2 dy^2 - c^2 dz^2, \quad (4.3.6)$$

где a , b и c — функции времени. Эта модель обобщает плоскую модель Фридмана ($a = b = c$) на анизотропный случай.

Осесимметричная модель второго типа с однородным искривленным пространством представляет собой топологическое произведение двумерной сферы на прямую (модель Кантовского–Сакса):

$$ds^2 = dt^2 - a^2 dx^2 - b^2 (d\theta^2 + \sin^2\theta d\phi^2). \quad (4.3.7)$$

В вакуумном случае она описывает T -область геометрии Шварцшильда (рис. 2.4).

4.4. Квази Хаббловские течения

В предыдущих разделах мы познакомились с однородными космологическими моделями, эволюция которых протекает одинаково во всех точках пространства. Полная группа всех пространственных движений включает три трансляции и три поворота (группа Фридмана). При нарушении фридмановской симметрии пространство остается однородным, если преобразования группы сохраняют три трансляции.

В этом разделе мы рассмотрим неоднородные космологические модели с полностью нарушенной симметрией Фридмана, но все же близкие к ней: существуют малые геометрические объекты (*малые функции*), задающие отклонения данной геометрии от пространственно-плоской модели Фридмана (см (4.1.1), (4.1.2), $\kappa = 0$). Мы будем называть такие геометрические модели *квазифридмановскими*, порождающие их потоки материи — *квази-*

хаббловскими, а малые функции — космологическими возмущениями или просто *возмущениями*. Наличие возмущений приводит к тому, что эволюция в этих моделях протекает по-разному в разных точках пространства. В этом случае говорят о существовании пространственной *космологической структуры*.

Данные наблюдательной космологии, экстраполированные в прошлое в соответствии с ОТО, указывают на то, что геометрия ранней Вселенной была именно квазифридмановской с амплитудой возмущений $\sim 10^{-5}$. Столь простые начальные условия привели к развитию сложных нелинейных структур, которые мы сегодня наблюдаем и частью которых являемся сами. В последующих главах мы рассмотрим процессы рождения квазифридмановских геометрий и начальных космологических неоднородностей, приведших к образованию структуры, а сейчас обратимся к описанию свойств квазихаббловских течений материи.

Важной геометрической характеристикой пространства-времени является тензор Эйнштейна или полный тензор энергии-импульса материи T_{μ}^{ν} . Энергия и импульс переносятся в пространстве-времени вдоль времениподобных мировых линий — траекторий потока. Касательный к ним вектор u^{μ} есть собственный вектор тензора T_{μ}^{ν} :

$$T_{\mu}^{\nu} u^{\mu} = \varepsilon u^{\nu}, \quad (4.4.1)$$

а собственным значением является скаляр полной сопутствующей плотности энергии ¹⁾:

$$\varepsilon = \rho c^2.$$

Единичный вектор u^{μ} ($u_{\mu} u^{\mu} = 1$) описывает скорость переноса полной энергии потока. По определению,

$$p_{\alpha\nu} u^{\mu} T_{\mu}^{\nu} = 0,$$

где $p_{\mu}^{\nu} \equiv \delta_{\mu}^{\nu} - u_{\mu} u^{\nu}$ — проекционный пространственный тензор, ортогональный к u^{μ} . Соответственно симметричный тензор $p_{\alpha\nu} T_{\beta}^{\nu}$ называется тензором давлений.

Рассмотрим решения уравнений (4.4.1).

Во фридмановской Вселенной решение (4.4.1) единственно для всех видов материи: компоненты среды движутся коллине-

¹⁾ В принятых нами единицах измерений плотность массы и энергии совпадают ($c = 1$). Их разные обозначения, соответственно ρ и ε , важны только для восстановления размерности и предельных переходов. Например, $G\rho$ не содержит константу скорости света и имеет такую же размерность, как и H^2 .

арно вектору u^μ (в координатах Фридмана $u^\mu = \delta_0^\mu$), и общий тензор энергии-импульса материи имеет универсальную форму:

$$T_\mu^\nu = (\varepsilon + p) u_\mu u^\nu - p \delta_\mu^\nu = \varepsilon u_\mu u^\nu - p p_\mu^\nu, \quad (4.4.2)$$

которая содержит единственный «новый» скаляр $p = p(t)$ — эффективное давление среды. Если материя состоит из нескольких свободных компонент, то ε и p равны суммам парциальных плотностей и давлений компонент среды. Тождества Бианки (4.1.5) обеспечивают выполнение закона сохранения энергии (уравнение непрерывности).

В отличие от плотности энергии, давление может иметь скачки и разрывы первого рода, определяемые физическими свойствами самой среды. Расходимость потока материи изотропна и описывается функцией Хаббла:

$$u_{\mu;\nu} = H p_{\mu\nu}, \quad H = H(t) = \frac{1}{3} u^\nu_{;\nu} = \frac{\dot{a}}{a}. \quad (4.4.3)$$

При взятии ковариантной производной мы учли связности фридмановской системы отсчета:

$$\Gamma_{\alpha\beta}^\gamma = H (u_\alpha p_\beta^\gamma + u_\beta p_\alpha^\gamma - u^\gamma p_{\alpha\beta}). \quad (4.4.4)$$

При слабом нарушении группы Фридмана разные компоненты среды могут двигаться в разных направлениях, однако отклонения их скоростей от средней u^μ малы. При этом условии решение (4.4.1) остается единственным, и мы получаем *общий* вид тензора энергии-импульса материи в квазифридмановской геометрии:

$$T_\mu^\nu = (\varepsilon + p) u_\mu u^\nu - p \delta_\mu^\nu - s_\mu^\nu = \varepsilon u_\mu u^\nu - (p p_\mu^\nu + s_\mu^\nu), \quad (4.4.5)$$

содержащий дополнительный по отношению к (4.4.2) малый тензор натяжений s_μ^ν , ортогональный потоку u^μ ($u^\mu s_\mu^\nu = 0$). Как видим, полный тензор давлений квази Хаббловского потока имеет два слагаемых (члены в скобках правой части (4.4.5)): диагональный тензор p и анизотропный тензор натяжений s_μ^ν . Полное давление потока можно определить по следу тензора давлений:

$$p_{\text{tot}} = \frac{1}{3} p^{\mu\nu} T_{\mu\nu} = p + \frac{1}{3} s_\nu^\nu. \quad (4.4.6)$$

В квазифридмановской геометрии, как и в изотропной фридмановской, выполняются правила линейного сложения взаимо-

действующих лишь гравитационным образом компонент материи:

$$\varepsilon = \sum_m \varepsilon_m, \quad p = \sum_m p_m, \quad u^\mu = \sum_m f_m u_m^\mu, \quad (4.4.7)$$

где скаляры f_m относительных вкладов компонент материи имеют вид:

$$f_m \equiv \frac{\varepsilon_m + p_m}{\varepsilon + p}, \quad \sum_m f_m = 1.$$

Столь простые линейные суперпозиции скоростей компонент среды и парциальных сопутствующих плотностей энергии и давлений справедливы только для членов нулевого + первого порядков малости по отклонениям скоростей u_m^μ от средней скорости u^μ .

Анизотропия давления связана с присутствием в веществе слабозаимодействующих частиц (см. также раздел 4.5). Материальные среды, для которых $s_\mu^\nu = 0$, мы будем называть *идеальными*. Идеальные среды имеют паскалево (изотропное) давление и охватывают широкий класс материальных полей. В частности, потенциальные течения идеальных сред тождественны движениям скалярных полей (подробнее см. раздел 6.4).

Проецируя тождества Бианки $T_{\mu;\nu}^\nu = 0$ на касательное и ортогональные к u^μ направления, получаем уравнения сохранения энергии и импульса квазихаббловского потока, включающие нулевой и первый порядки по малым отклонениям геометрии от фридмановской:

$$u^\nu \varepsilon_{;\nu} + u^\nu_{;\nu} (\varepsilon + p_{\text{tot}}) = 0, \quad (4.4.8)$$

$$a_\mu \equiv u^\nu u_{\mu;\nu} = u^\nu (u_{\mu;\nu} - u_{\nu;\mu}) = \frac{p_\mu^\nu p_{;\nu} + s_{\mu;\nu}^\nu}{\varepsilon + p}, \quad (4.4.9)$$

где a_μ — ускорение потока ($a_\nu u^\nu = 0$). Мы также использовали вспомогательное выражение для ковариантной производной малого тензора натяжений (см. (4.4.4)):

$$s_{\mu;\nu}^\nu = s_{\mu;\nu}^\nu - H s_\nu^\nu u_\mu.$$

Для однозначного разбиения скалярной части давления на изотропную и анизотропную части мы доопределим далее скалярную компоненту тензора натяжений условием ее нулевой расходимости (см. (4.5.2)). Тогда ускорение квазихаббловского потока в релятивистском уравнении Эйлера (4.4.9) будет регулироваться градиентами изотропной скалярной и анизотропной вихревой частей полного давления.

4.5. Космологические возмущения

Поскольку отклонения от изотропной геометрии малы, возмущения метрического тензора и тензора энергии-импульса можно разложить по неприводимым представлениям метрики Фридмана $g_{\mu\nu}^{(F)} = \text{diag}(1, -a^2\delta_{ij})$:

$$g_{\mu\nu} = g_{\mu\nu}^{(F)} + h_{\mu\nu}, \quad (4.5.1)$$

$$\frac{1}{2}h_{\mu\nu} = \begin{pmatrix} D & \frac{1}{2}(C_{,i} + a^{-1}\mathfrak{S}_i + a^2\dot{\xi}_i) \\ \text{sym} & a^2(A\delta_{ij} + B_{,ij} + \mathfrak{S}_{ij} + \xi_{(i,j)}) \end{pmatrix},$$

$$\varepsilon = \varepsilon^{(F)} + \delta\varepsilon, \quad p = p^{(F)} + \delta p, \quad (4.5.2)$$

$$s_{\mu}^{\nu} = \text{diag}\left(0, \frac{\Delta S \delta_{ij} - S_{,ij}}{8\pi G a^2} + \frac{\dot{S}_{(i,j)}}{8\pi G a^3} + S_{ij}\right),$$

$$u_{\mu} = (1 + D, v_{,i} + v_i), \quad s_{\nu}^{\nu} = \frac{\Delta S}{4\pi G a^2}, \quad s_{\mu,\nu}^{\nu} = \left(0, \frac{\Delta \dot{S}_i}{16\pi G a^3}\right),$$

где условия неприводимости означают, что

$$\mathfrak{S}_{,i}^i = v_{,i}^i = S_{,i}^i = \xi_{,i}^i = \mathfrak{S}_i^i = \mathfrak{S}_{i,j}^j = S_i^i = S_{i,j}^j = 0, \quad (4.5.3)$$

все манипуляции с пространственными индексами выполняются с помощью единичного тензора δ_{ij} , а круглые скобки в индексах означают симметризацию.

Координатные расщепления (t, x^i) , соответствующие представлениям ¹⁾ (4.5.1) и (4.5.2), мы будем называть *квазифрид-*

¹⁾ Разбиение геометрического объекта на *фон* и *возмущение* неоднозначно: при малом преобразовании координат $x^{\mu} \rightarrow x^{\mu} - \xi^{\mu}$ мы получим новый фон (те же фоновые функции, но другого времени t) и новое возмущение, хотя сам объект не меняется. Раскладывая малый произвольный вектор $\xi_{\mu} = (X, a^2(Y_{,i} + \xi_i))$ на неприводимый вихревой вектор ξ и два потенциала X и Y , получаем следующие калибровочные преобразования переменных:

$$\begin{aligned} h_{\mu\nu} &\rightarrow h_{\mu\nu} + \xi_{\mu;\nu} + \xi_{\nu;\mu}, & u_{\mu} &\rightarrow u_{\mu} + X_{,\mu}, \\ A &\rightarrow A - HX, & B &\rightarrow B + Y, & C &\rightarrow C + X + a^2\dot{Y}, \\ D &\rightarrow D + \dot{X}, & v &\rightarrow v + X, & t &\rightarrow t - X, \end{aligned}$$

функции v_i , \mathfrak{S}_i и \mathfrak{S}_{ij} калибровочно-инвариантны.

мановскими. Из уравнений Эйлера (4.4.9) получаем следующие связи:

$$\delta_p \equiv \frac{\delta p - \dot{p}v}{\varepsilon + p} = \dot{v} - D, \quad (4.5.4)$$

$$\Delta S_i = 16\pi G a^3 (\varepsilon + p) v_i. \quad (4.5.5)$$

Фоновые функции времени a , H , γ , $\varepsilon^{(F)}$ и $p^{(F)}$ удовлетворяют уравнениям Фридмана (индекс « F » по возможности опускаем):

$$\frac{3}{2} H^2 = 4\pi G \varepsilon^{(F)}, \quad \gamma \equiv -\frac{\dot{H}}{H^2} = \frac{3}{2} \left(1 + \frac{p}{\varepsilon}\right). \quad (4.5.6)$$

Остальные функции малы и зависят от всех четырех координат, они связаны между собой линеаризованными уравнениями Эйнштейна:

$$\delta G_{\mu}^{\nu} = 8\pi G \delta T_{\mu}^{\nu}. \quad (4.5.7)$$

Как видно, все космологические неоднородности разбиваются на три независимые друг от друга в линейном порядке моды возмущений (классификация Лифшица):

- скалярная S (потенциалы A , B , C , D , v , $\delta\varepsilon$, δp и S);
- векторная, или вихревая, V (векторы \mathfrak{S}_i , v_i , S_i и ξ_i);
- тензорная T (тензоры \mathfrak{S}_{ij} и S_{ij}).

Наиболее просто устроены моды T и V .

Тензорные поля \mathfrak{S}_{ij} и S_{ij} калибровочно-инвариантны и связаны между собой волновым уравнением:

$$\ddot{\mathfrak{S}}_{ij} + 3H\dot{\mathfrak{S}}_{ij} - a^{-2}\Delta\mathfrak{S}_{ij} = 8\pi G S_{ij}. \quad (4.5.8)$$

Векторные поля \mathfrak{S}_i , S_i и v_i также калибровочно-инвариантны и связаны следующими соотношениями:

$$\mathfrak{S}_i = S_i, \quad \Delta\mathfrak{S}_i = 16\pi G a^3 (\varepsilon + p) v_i, \quad (4.5.9)$$

а ξ_i — произвольный вектор, определяющий выбор калибровки, см. также (4.5.5). Примером векторной моды могут быть космологические магнитные поля. В идеальной среде $\dot{S}_i = 0$ (см. (4.5.2)), и уравнения (4.5.5) и (4.5.9) описывают закон сохранения момента импульса:

$$a^3 (\varepsilon + p) v_i = \text{const}(t). \quad (4.5.10)$$

Анизотропные давления в правых частях уравнений (4.5.8) и (4.5.9) возникают, если в веществе присутствуют слабозаимо-

действующие частицы. Для примера возьмем свободные частицы с некоторой функцией распределения по скоростям (например, изотропной). Поскольку частицы движутся по геодезическим, то их физические скорости убывают обратно пропорционально масштабному фактору. При приливном гравитационном воздействии (например, в поле гравитационной волны) локальная деформация среды описывается тремя масштабными факторами в трех пространственных направлениях (первый тип Бианки, см. (4.3.6)). В результате импульсы частиц, движущихся в тех направлениях, где расширение замедленно, оказываются с течением времени несколько больше, чем у частиц, которые распространяются в ортогональных направлениях с более быстрым расширением. Так возникает анизотропия тензора давлений, которая, в свою очередь, обратно воздействует на эволюцию возмущений метрики согласно уравнениям (4.5.7). Этот эффект универсален для всех мод возмущений и исчезает в пределе малых длин пробега частиц. В идеальной среде ($\dot{S}_i = S_{ij} = 0$) уравнения (4.5.8) и (4.5.9) описывают соответственно свободные гравитационные волны и сохранение момента импульса (4.5.10).

Наибольший интерес представляет собой скалярная мода, поскольку именно она связана с возмущениями плотности, эволюция которых приводит к образованию галактик. По этой причине скалярную моду часто называют модой возмущений плотности. Она полностью описывается четырьмя гравитационными (A, B, C, D) и четырьмя материальными ($v, \delta\varepsilon, \delta p, S$) потенциалами. Скаляр S анизотропного давления калибровочно-инвариантен, а остальные семь функций нет.

Введем калибровочно-инвариантные переменные для потенциала скорости, возмущений плотности и давления материи:

$$q = A + H v, \quad (4.5.11)$$

$$\delta\varepsilon_c \equiv \delta\varepsilon - \dot{\varepsilon} v = \varepsilon - \varepsilon_c, \quad (4.5.12)$$

$$\delta p_c \equiv \delta p - \dot{p} v = p - p_c = p_{\text{tot}} - p_v, \quad (4.5.13)$$

где $\varepsilon_c \equiv \varepsilon^{(F)}(t_c)$, $p_c \equiv p^{(F)}(t_c)$. Буква «с» в индексе означает *сопутствие*: сопутствующие переменные тождественны лагранжевым переменным и строятся на пространственноподобных гиперповерхностях постоянного времени $t_c = t + v = \text{inv}$ (см. сноску на с. 147).

В отличие от полной плотности энергии, давление квазиабловского потока многозначно. Мы ввели дополнительный скаляр

сопутствующего *объемного* давления (ср. (4.4.6)):

$$p_v = p_c + \frac{1}{3} s_\nu^\nu, \quad \varepsilon + p_v = (\varepsilon + p_{\text{tot}})(1 - \delta_p). \quad (4.5.14)$$

Скаляр q безразмерный. Определим безразмерные возмущения сопутствующих плотности и давления среды:

$$\delta_\varepsilon \equiv \frac{\delta\varepsilon_c}{\varepsilon + p}, \quad \delta_p \equiv \frac{\delta p_c}{\varepsilon + p}. \quad (4.5.15)$$

Итак, мы построили четыре калибровочно-инвариантных скаляра q , δ_ε , δ_p и S . Оставшиеся четыре метрических потенциала A , B , C и D калибровочно-неинвариантны: любые два из них можно выбрать произвольным образом с помощью подбора функций X и Y (см. сноску на с. 147). Таким образом, у нас имеется всего шесть независимых скаляров, описывающих возмущения плотности в модели Фридмана.

Центральную роль в описании моды S играет скалярное поле $q = q(t, \mathbf{x})$. Из уравнений Эйнштейна низшего порядка следует связь исходных потенциалов S -моды с полем q :

$$\frac{\Delta\Phi}{a^2} = 4\pi G \delta\varepsilon_c, \quad \Phi = \frac{H}{a} \int (\gamma q - S) \frac{da}{H}, \quad (4.5.16)$$

$$v - C + a^2 \dot{B} = \frac{q}{H}, \quad \dot{v} - D = \frac{\dot{q}}{H}, \quad (4.5.17)$$

$$\delta_p = \frac{\dot{q}}{H}, \quad (4.5.18)$$

где $q \equiv q - \Phi$ — скалярный потенциал пекулярной скорости материи (см. (5.1.15)). Эти преобразования вытекают соответственно из временного, пространственно-временного и пространственно-бесследового уравнений (4.5.7) и (4.5.4).

Мы видим, что измеряемые физические переменные зависят от производных q : градиент и временная производная q определяют пекулярную скорость и возмущение давления, а вторые производные — компоненты пространственной кривизны. Скаляр q имеет двойной физический смысл. С одной стороны, это безразмерный потенциал полной скорости материи Hv (лабораторная интерпретация). С другой стороны, это относительное возмущение масштабного фактора $A \equiv \delta a/a$ (космологическая интерпретация). В силу сложившихся исторических причин скалярным потенциалом гравитационного поля мы называем поле Φ . Поэтому мы будем называть поле q скалярным потенциалом кривизны или просто *скаляром кривизны*.

Ключевое уравнение (4.5.16) является релятивистским уравнением Пуассона: оно связывает лапласиан гравитационного по-

тенциала Φ и лагранжево возмущение плотности $\delta\varepsilon_c$. Поле Φ (аналог ньютоновского потенциала) получается интегрированием q и S вдоль мировой линии потока. Уравнение (4.5.18) является релятивистским уравнением Эйлера или вторым законом Ньютона: оно связывает ускорение потока \dot{q} с действующим на него градиентом давления δp_c . Следует подчеркнуть, что на ускорение потока влияет только та часть давления, которую мы обозначили буквой p — сюда не входит тензор натяжений s_{μ}^{ν} , имеющий, по определению, нулевую расходимость (см. (4.5.2)).

Уравнения (4.5.16)–(4.5.18) полностью решают гравитационную проблему скалярных возмущений в космологии Фридмана¹⁾. Они задают четыре связи на шесть потенциалов S -моды. Очевидно, одних только уравнений гравитации недостаточно для описания динамики скалярных возмущений. Для этого необходима информация о физическом состоянии материи в виде двух недостающих связей (уравнений состояния). В главе 6 мы выясним, какие именно характеристики материи потребуются, а сейчас вернемся к общему анализу скалярной моды.

4.6. Анизотропная Вселенная

Мы рассмотрели малые геометрические возмущения плоской модели Фридмана. Данные наблюдательной космологии не дают пока серьезных оснований усомниться в том, что пространство видимой Вселенной изотропно и плоско. Однако, если такие свойства и будут обнаружены в будущем, отклонения глобальной симметрии будут описываться *малыми* анизотропией и искривлением пространства в целом и должны содержаться в классе рассмотренных возмущений. Этот подкласс возмущений большого размера представляет для нас интерес, поскольку он может быть использован в будущем для возможной коррекции (перенормировки) фона.

Из полученных уравнений (4.5.16)–(4.5.18) следует, что задающими скалярами возмущений плотности в плоской модели Фридмана являются q и S . Они калибровочно-инвариантны, однако в их определении еще остаются степени свободы.

¹⁾ Подчеркнем, что мы использовали только уравнения Эйнштейна первого порядка по времени (гравитация не увеличивает число степеней свободы скалярных и вихревых возмущений материи). Уравнения второго порядка содержат новую динамическую информацию лишь о гравитационно-волновой моде (см. (1.8.16), (4.5.8)).

Обращаясь к исходным потенциалам, можно заметить, что потенциал A и, следовательно, скаляр q (см. (4.5.11)) определены с точностью до аддитивной постоянной из-за возможности изотропных пространственных растяжений. В то же время потенциалы B , C , v , S и Φ вводились в виде комбинаций их пространственных производных. Поэтому они определены с точностью до постоянных, а B , S и Φ — еще и линейных по x^i членов.

Это означает, что первое интересное для нас нетривиальное решение уравнений (4.5.16)–(4.5.18) имеет вид $q = S = \delta_\epsilon = 0$. Поскольку поле Φ связано с q и S с помощью временного интеграла, то даже в этом случае существует ненулевое решение для Φ :

$$\Phi = \frac{H}{a} P(\mathbf{x}), \quad \Delta P = 0. \quad (4.6.1)$$

Докажем, что это решение переводит изотропную модель Фридмана в однородную модель первого типа по классификации Бианки (4.3.6).

Действительно, мы всегда можем выбрать калибровку $A = C = 0$, в которой единственным ненулевым метрическим потенциалом решения (4.6.1) оказывается

$$B = -P \int \frac{da}{a^3 H}.$$

В классе равномерно ограниченных в пространстве функций $P_{,ij}$ нетривиальным физическим решением гармонического уравнения (4.6.1) является билинейная форма

$$P(\mathbf{x}) = \frac{1}{2} c_{ij} x^i x^j,$$

где $c_{ij} = c$ — матрица констант с равном нулю следом $\text{Sp}(c) = 0$. Эта форма задает линейный по малому параметру c член разложения анизотропной модели Бианки первого типа с плоским сопутствующим пространством в окрестности изотропной модели Фридмана:

$$\gamma_{ij} = a^2 \exp\left(2c \int \frac{da}{a^3 H}\right), \quad H^2 = \frac{8\pi G}{3} \epsilon + \frac{\text{Sp}(c^2)}{6a^6}. \quad (4.6.2)$$

Решения (4.6.1), (4.6.2) с $q = 0$ описывают возмущения бесконечного масштаба, не искажающие кривизну. Можно сказать, что они перенормируют плоский фон. Нетрудно убедиться в том, что и для мод V и T возмущения бесконечного масштаба также

приводят к модели (4.6.2). В практическом плане речь может идти о масштабах анизотропии много больше космологического горизонта: в этом пределе все решения будут соответствовать модели Бианки первого типа, а не плоской модели Фридмана.

Что касается неоднородностей конечной, но большой длины волны, то они возмущают кривизну и индуцируются ненулевым скаляром q . Таких решений много. В качестве примера приведем подкласс искривленных моделей Фридмана (на фоне плоского). Легко проверить, что уравнения S-моды удовлетворены при

$$q = \Phi = A = -S = \frac{\kappa}{4} \mathbf{x}^2,$$

что описывает первый член разложения метрики (4.1.2) по малой константе κ : $C = B = D = v = 0$, $\bar{r} = |\mathbf{x}|$.

ВОЗНИКНОВЕНИЕ ХАББЛОВСКОГО ПОТОКА

В данной главе мы исследуем физические свойства квази-фридмановской геометрии и покажем, каким образом и почему она выделена среди решений ОТО. Мы построим геометрические переменные и действие для слабонеоднородной Вселенной, рассмотрим ее взаимосвязь с квантовой гравитацией и получим действия трех мод возмущений квазиклассического хаббловского потока.

5.1. Геометрические переменные и лагранжевы системы

В предыдущей главе мы определили сектор малых возмущений геометрии Фридмана с помощью калибровочно-инвариантных функций q , \mathfrak{S}_i и \mathfrak{S}_{ij} . Для описания *полной* геометрии слабо нарушенных хаббловских течений нам понадобятся калибровочно-инвариантные переменные, включающие *оба* порядка теории возмущений (нулевой + первый). Такими переменными являются скаляры полной плотности ε и давления p среды, однако нам нужны еще масштабные и хаббловские скалярные факторы (фоновые функции a и H не удовлетворяют этим условиям, поскольку они калибровочно-неинвариантны).

Определим две группы геометрических скаляров:

$$\begin{aligned} a &\equiv a \cdot (1 - A) = a_c \cdot (1 - q), \\ a_c &= a(t_c) = a + \dot{a}v, \end{aligned} \quad (5.1.1)$$

$$\begin{aligned} H_c &= H(t_c) \equiv \frac{\dot{a}_c}{a_c} = u^\mu (\ln a)_{,\mu}, \\ a_v &\equiv a \det^{1/3}(\delta_{ij} - \mathcal{B}_{,ij}) = b \cdot (1 - q), \\ b &\equiv a_c \cdot (1 - \Delta \mathcal{B}/3), \end{aligned} \quad (5.1.2)$$

$$H_v \equiv \frac{\dot{b}}{b} = \frac{1}{3} u^\mu{}_{;\mu} = u^\mu (\ln a_v)_{,\mu}.$$

Здесь точка над переменной обозначает производную по *сопутствующему* времени ¹⁾ t_c , буква «v» в индексе означает *объем*: функции a_v и H_v соответствуют локальным объемным масштабному фактору и темпу расширения материи.

Скалярный масштабный фактор $a = a(x^\mu)$ (5.1.1) определен для *любых* квазифридмановских систем отсчета (4.5.1). Для более глубокого понимания физического смысла введенных инвариантных переменных удобно обращаться к координатным системам *частного* вида. Рассмотрим наиболее интересные из них.

Лагранжевы ортогональные координаты $x_c^\mu = (t_c, \mathbf{x}_c)$ определяются условиями *сопутствия* S-моды возмущений и *ортогональностью* квазифридмановского разбиения (4.5.1):

$$v = C = \mathfrak{S}_i + a^3 \dot{\xi}_i = 0. \quad (5.1.3)$$

Метрика в этих координатах имеет следующий вид:

$$ds^2 = (1 - 2\delta_p) dt_c^2 - a^2 (\delta_{ij} - 2B_{,ij} - 2\bar{\mathfrak{S}}_{ij}) dx_c^i dx_c^j, \quad (5.1.4)$$

$$u_\mu = (1 - \delta_p, v_i), \quad \delta_p = \frac{\dot{q}}{H},$$

$$B \equiv \int q \frac{d\eta}{H}, \quad q \equiv q - \Phi,$$

$$\bar{\mathfrak{S}}_{ij} = \mathfrak{S}_{ij} - \int \mathfrak{S}_{(i,j)} \frac{dt}{a^3}.$$

Как видно, собственное время τ в точке \mathbf{x}_c связано с сопутствующим временем t_c условием:

$$d\tau = (1 - \delta_p) dt_c = \left(1 - \frac{\dot{q}}{H}\right) dt_c, \quad (5.1.5)$$

где $ds = d\tau$ на мировой линии элемента потока $\mathbf{x}_c = \text{const}$.

Рассмотрим подробнее структуру пространственного расширения квазихаббловского потока материи. Собственное расстояние между двумя близкими элементами среды, разделенными координатами δx_c^i , равно:

$$\delta r_i = a (\delta_{ij} - B_{,ij} - \bar{\mathfrak{S}}_{ij}) \delta x_c^j, \quad (5.1.6)$$

$$\det \left(\frac{\delta r_i}{\delta x_c^j} \right) = a_v^3.$$

¹⁾ Напомним, что фоновый масштабный фактор $a = a(t)$ есть функция своего времени в любой координатной сетке (t, x^i) .

Дифференцируя δr_i по собственному времени τ , получаем поле парных скоростей движения материи:

$$\begin{aligned}\delta V_i &\equiv \frac{\partial \delta r_i}{\partial \tau} = H_{ij} \delta r^j, \\ H_{ij} &\equiv H_c \delta_{ij} - h_{ij}, \quad H_i^i = 3H_v, \\ h_{ij} &= \frac{q_{,ij}}{a\dot{H}} + \dot{\mathfrak{S}}_{ij} - a^{-3} \mathfrak{S}_{(i,j)},\end{aligned}\tag{5.1.7}$$

где H_{ij} — матрица калибровочно-инвариантных хаббловских функций. В общем случае хаббловскую матрицу можно получить, проецируя тензор деформации скорости $u_{\mu;\nu}$ на направления $p_{\mu\nu}$, ортогональные потоку энергии u^μ :

$$u_{\mu;\nu} = a_\mu u_\nu + H_c p_{\mu\nu} + \text{diag}(0, v_{i,j} - h_{ij}),\tag{5.1.8}$$

где ускорение $a_\mu \equiv u^\nu u_{\mu;\nu} = (0, \delta_{p,i} + \dot{v}_i)$ (см. (4.4.9)).

Таким образом, пространственный разлет материи описывается матрицей H_{ij} . О хаббловском расширении здесь напоминает лишь одно свойство: относительная скорость разбегания точек среды пропорциональна расстоянию между ними. Однако эти скорости анизотропны и зависят от положения в пространстве. Векторная и тензорная моды не возмущают объемное расширение среды, поэтому зависимость H_v от координат определяется только скалярной модой возмущений. В следующем разделе мы выведем уравнения для H_v , аналогичные фридмановским, а сейчас вернемся к системам отсчета.

Переход к *произвольным* квазифридмановским координатам (t, x^i) осуществляется с помощью преобразований:

$$\begin{aligned}t_c &= t + v, \\ x_c^i &= x^i + \int (v - C)_{,i} \frac{dt}{a^2} - \int \mathfrak{S}_i \frac{dt}{a^3} - \xi_i.\end{aligned}\tag{5.1.9}$$

Среди других координатных систем выделяется эйлерова ортогональная система отсчета (t_E, y) , где $y^i \equiv x_E^i$, которая фиксируется отсутствием сдвиговой скалярной деформации пространства и ортогональностью времени:

$$B = C = \mathfrak{S}_i + a^3 \dot{\xi}_i = 0.\tag{5.1.10}$$

Эйлеровы координаты являются аналогом абсолютных ньютоновских пространственных и временных координат, именно в них понятным образом определяется *пекулярная* скорость материи.

Из (5.1.9) получаем преобразование систем отсчета и потенциал пекулярной скорости:

$$\begin{aligned} t_E &= t_c - v, & \mathbf{y} &= \mathbf{x}_c - \nabla B, \\ v_E &= v \equiv \frac{q}{H}, & B &\equiv \int v \frac{dt}{a^2}. \end{aligned} \quad (5.1.11)$$

По определению, лагранжево и эйлерово представления отличаются только для S-моды (ср. (5.1.3)), представления же мод V и T тождественны в любых квазифридмановских системах отсчета с ортогональным расщеплением $g_{0i} = 0$.

Метрический тензор в эйлеровом представлении имеет следующий вид:

$$ds^2 = (1 + 2\Psi) dt_E^2 - a^2 (\delta_{ij} - 2\bar{\mathcal{E}}_{ij}) dy^i dy^j. \quad (5.1.12)$$

Гравитационный потенциал

$$\Psi = \Phi - S \quad (5.1.13)$$

учитывает локальное изменение хода времени в зависимости от плотности энергии и пекулярной скорости:

$$\frac{\delta \epsilon_E}{\epsilon + p} = \delta_\epsilon - 3Hv, \quad (5.1.14)$$

$$\mathbf{v}_{\text{pec}} \equiv -\frac{\nabla v}{a} = -\frac{\nabla q}{H}. \quad (5.1.15)$$

Мы видим, что скалярный масштабный фактор

$$a = a(t_E) \cdot (1 - \Phi)$$

полностью описывает локально изотропное эйлерово пространство наблюдателя, заполненное неоднородной материей. Метрика (5.1.12) является релятивистским пределом слабого поля в модели Фридмана (ср. (1.10.7)). Для нерелятивистской среды ($|S/\Phi| \ll 1$) скалярная часть метрики (5.1.12) соответствует ньютоновскому пределу (1.10.7): $\Psi = \Phi$.

Итак, лагранжевы координаты определяются пространственным сопутствием потенциальным движениям материи ($v = 0$), а эйлеровы — локально-изотропной деформацией пространства ($B = 0$). Для однозначной фиксации систем отсчета мы использовали условие ортогональности линий времени соответствующим пространственным сечениям ($C = 0$). Последнее условие продиктовано исключительно удобством ортогональных расщеплений, и от него можно отказаться. В этом случае появляется уникальная возможность введения единых «лагранжево-эйлеровых» координат (t_c, \mathbf{y}) при требовании одновременного выполнения условий сопутствия и отсутствия анизотропии деформации

S-моды:

$$v = B = 0. \quad (5.1.16)$$

Сохраняя ортогональность по векторной моде:

$$\mathfrak{S}_i + a^3 \dot{\xi}_i = 0, \quad (5.1.17)$$

получаем метрику в следующем виде:

$$ds^2 = (1 - 2\delta_p) dt_c^2 - 2v_{,i} dy^i dt_c - a^2 (\delta_{ij} - 2\overline{\mathfrak{S}}_{ij}) dy^i dy^j,$$

Как видно, перекрестная компонента чисто потенциальная, тогда как пространство искривлено только V- и T-модами возмущений.

Заметим, что при нарушении ортогональности S-моды мы можем отказаться и от условия (5.1.17). Среди таких разбиений (5.1.16) наибольший интерес представляет выбор $\xi_i = 0$:

$$ds^2 = (1 - 2\delta_p) dt_c^2 - 2V_i dy_E^i dt_c - a^2 (\delta_{ij} - 2\mathfrak{S}_{ij}) dy_E^i dy_E^j, \\ V_i = v_{,i} - a^{-1} \mathfrak{S}_i, \quad (5.1.18)$$

где $y_E^i = y^i + \int \mathfrak{S}_i a^{-3} dt$. В этом представлении векторная мода приписана вектору V_i , а локальная изотропия пространственных сечений нарушена *только* присутствием гравитационных волн \mathfrak{S}_{ij} . Обе системы отсчета (5.1.16)–(5.1.17) и (5.1.18) фиксированы однозначным образом.

Рассмотренные системы отсчета (5.1.4), (5.1.12), (5.1.18) определены соответственно *двойными* условиями (5.1.3), (5.1.10), (5.1.16) для скалярной моды. В дальнейшем мы сохраним по умолчанию обозначение (t_c, \mathbf{y}) для любого лагранжево-эйлерового разбиения по S-моде (5.1.16). Нам также понадобится наиболее общий класс «лагранжевых неортогональных» систем отсчета (t_c, \mathbf{x}) , определенных *единственным* условием сопутствия в форме

$$v = 0. \quad (5.1.19)$$

Очевидно, системы отсчета (t_c, \mathbf{y}) являются подклассом систем (t_c, \mathbf{x}) . Эти общие лагранжевы (по S-моде) системы отсчета (5.1.19) играют такую же роль в квазифридмановской геометрии, как и фридмановская система отсчета (4.1.1) в геометрии Фридмана (см. далее разделы 5.2, 5.3). Рассмотрим этот класс разбиений (5.1.19) более детально.

Представим метрику в *нелинейном* АДМ-представлении (см. раздел 1.9):

$$ds^2 = (N^2 - N_i N^i) dt_c^2 - 2N_i dt_c dx^i - \gamma_{ij} dx^i dx^j = \\ = (N^2 - V_i V^i) dt_c^2 - 2V_i dt_c dy^i - a^2 E_{ij} dy^i dy^j = \\ = N^2 dt_c^2 - \gamma_{ij}^{(c)} dx_c^i dx_c^j, \quad (5.1.20)$$

допускающем свободу преобразований

$$t_c = t = \text{inv}, \quad \mathbf{x} = \mathbf{x}(\mathbf{y}, t) \quad (5.1.21)$$

или $\mathbf{x}_c = \mathbf{x}_c(\mathbf{x}, t)$, которые сохраняют условие (см. (1.9.5))

$$N t_{, \alpha} = n_\alpha = (N, 0, 0, 0) \quad (5.1.22)$$

(индекс «с» у времени t_c по возможности опускаем).

Преобразования (5.1.21) содержат три произвольных функции координат, которые можно обращать: $\mathbf{y} \leftrightarrow \mathbf{x} \leftrightarrow \mathbf{x}_c$, считая t параметром. По отношению к этим преобразованиям N является инвариантом (скаляром), V_i – 3-вектором:

$$V^i = N^j \frac{\partial y^i}{\partial x^j} - \dot{y}^i|_{\mathbf{x}}, \quad V^j \frac{\partial x^i}{\partial y^j} = N^i + \dot{x}^i|_{\mathbf{y}} \quad (5.1.23)$$

(также $N^i = -\dot{x}^i|_{\mathbf{x}_c}$), а γ_{ij} – метрический 3-тензор; точка над переменной означает производную по t при фиксированной координате, указанной при вертикальной черте.

В нулевом и первом порядках малости из (5.1.18) получаем:

$$V_i = v_{,i} - a^{-1} \mathfrak{S}_i, \quad E_{ij} = \delta_{ij} - 2\mathfrak{S}_{ij}. \quad (5.1.24)$$

Формулы (5.1.20)–(5.1.23) являются *точным* решением ОТО во всех порядках теории возмущений¹⁾. Предполагая их использование для разложений по малым отклонениям, перепишем (5.1.21) в виде, удобном для итераций:

$$B^i = B^i(\mathbf{x}, t) \equiv x^i - y^i \equiv - \int_0^t \dot{y}^i|_{\mathbf{x}} dt = \int_0^t \dot{x}^i|_{\mathbf{y}} dt \quad (5.1.25)$$

(интегралы берутся соответственно вдоль линий $\mathbf{x} = \text{const}$ и $\mathbf{y} = \text{const}$). Отсюда получаем представления 3-тензоров:

$$\begin{aligned} \gamma_{ij} &= a^2 E_{kl} \left(\delta_i^k - \frac{\partial B^k}{\partial x^i} \right) \left(\delta_j^l - \frac{\partial B^l}{\partial x^j} \right), \\ a^2 E_{ij} &= \gamma_{kl} \left(\delta_i^k + \frac{\partial B^k}{\partial y^i} \right) \left(\delta_j^l + \frac{\partial B^l}{\partial y^j} \right), \end{aligned}$$

¹⁾ Для построения нелинейных космологических решений необходимо задавать связи между вектором n^μ и скоростью (или скоростями) материи, что определяется индивидуально в зависимости от постановки задачи. Так, точное равенство $u_\alpha = n_\alpha$ выполняется для однокомпонентной среды при отсутствии вращения. Это условие можно нарушить, рассматривая, например, более общие (несопутствующие) системы отсчета (см. (4.5.17)), что несколько усложнит полученные нами нелинейные соотношения. Напротив, линейное приближение (5.1.19) верно для любых сред.

где в последнем равенстве предполагается $B = B(\mathbf{x}(\mathbf{y}, t), t)$, а также соотношения, связывающие потенциалы N^i и B^i (не являющиеся векторами, см. (5.1.23)) с вектором V^i :

$$\dot{B}^i|_{\mathbf{x}} + N^j \frac{\partial y^i}{\partial x^j} = V^i, \quad (5.1.26)$$

$$\dot{B}^i|_{\mathbf{y}} + N^i = V^j \frac{\partial x^i}{\partial y^j}. \quad (5.1.27)$$

В линейном порядке по V^i эти формулы переходят в соотношения (4.5.17) для квазифридмановских сопутствующих (по S-моду) систем отсчета (5.1.19).

5.2. Неоднородные уравнения Фридмана

Теперь у нас есть необходимые геометрические переменные для описания динамики квазихаббловского потока.

Суммируя уравнения (4.5.6) и (4.5.16), получаем искомые уравнения эволюции для масштабного фактора объемного расширения материи $b = b(t, \mathbf{x})$ в лагранжевых координатах, сопутствующих полной плотности энергии ¹⁾:

$$H_v^2 = \frac{8\pi G}{3} \varepsilon - \frac{\varkappa}{b^2}, \quad (5.2.1)$$

$$\dot{\varepsilon} + 3H_v(\varepsilon + p_v) = 0, \quad (5.2.2)$$

где

$$H_v = \frac{\dot{b}}{b} = \frac{1}{3} \mathcal{K}$$

— локальный хаббловский фактор объемного расширения (5.1.2) (внешняя кривизна, ср. (1.9.14), (5.1.22)),

$$\varkappa = \varkappa(t, \mathbf{x}) \equiv \frac{2}{3} \Delta q \quad (5.2.3)$$

— локальная (внутренняя) кривизна пространства. Дифференцируя (5.2.1) и исключая $\dot{\varepsilon}$ с помощью (5.2.2), получаем уравнение для производной хаббловской функции:

$$-\dot{H}_v = 4\pi G(\varepsilon + p_v) - \frac{\varkappa}{b^2} + \frac{\dot{\varkappa}}{2bH}. \quad (5.2.4)$$

Полученные уравнения напоминают по форме динамические уравнения Фридмана, однако входящие в них скаляры являют-

¹⁾ Буква «с» в индексе опущена (см. (5.1.21)). Уравнение (5.2.2) следует из (4.4.8), (4.5.14).

ся функциями всех четырех координат, а точка означает частную производную по сопутствующему времени t_c при фиксированных \mathbf{x}_c ($\dot{x} \equiv \partial x / \partial t_c = (1 - \delta_p) u^\mu x_{,\mu}$, см. (5.1.5)). Скаляр $\varepsilon = \varepsilon(t, \mathbf{x})$ представляет собой полную плотность энергии квази-хаббловского потока (4.4.1), тогда как p_v — это лишь та часть давления, у которой зависимость от пространственных координат на сечениях $t_c = \text{const}$ определяется следом анизотропного тензора натяжений s_i^i (см. (4.5.14)).

Квазифридмановские уравнения (5.2.1) и (5.2.2) справедливы для малой пространственной кривизны $|\varkappa| \ll 1$, и с этим связан их простой и ясный физический смысл. В частности, из (5.2.3) следует, что скаляр q является потенциалом пространственной кривизны в лагранжевой системе отсчета. Уравнения (5.2.1) и (5.2.2) определяют эволюцию квазифридмановской Вселенной в геометрических инвариантных переменных, описывающих нулевой + первый порядок по отклонениям квазифридмановской геометрии от фридмановской.

Эти уравнения имеют скалярный вид, хотя описываемая ими геометрия анизотропна (см. (5.1.4)). Моды V и T не влияют на объемное расширение. Чтобы решить квазифридмановское уравнение (5.2.1) для объемного фактора b , надо знать малую функцию q — задающий скаляр S -моды возмущений, лапласиан которого определяет пространственную кривизну \varkappa . Предполагая, что функция q нам известна, ищем решение (5.2.1) в виде:

$$\begin{aligned} H_v &= H(t_c) + \frac{\Delta\Phi - \Delta q}{3aH}, \\ \varepsilon &= \varepsilon(t_c) + \frac{\Delta\Phi}{4\pi G a^2}, \quad p_v = p(t_c) + \frac{\Delta S}{12\pi G a^2}, \end{aligned} \quad (5.2.5)$$

где Φ и S — произвольные малые функции. Подставляя (5.2.5) в (5.2.1), (5.2.2), получаем *правильную* связь между Φ и q :

$$\Phi = \frac{H}{a} \int a (\gamma q - S) dt, \quad (5.2.6)$$

причем S остается произвольной функцией. Мы уже подчеркивали, что S не определяется из уравнений Эйнштейна, а задается физическим состоянием материи.

5.3. Действие неоднородной Вселенной

Мы получили все ключевые уравнения ОТО, описывающие квазифридмановскую Вселенную с помощью геометрических

переменных. Построим гравитационное действие этой геометрии в виде функции инвариантных полей.

В качестве независимых метрических потенциалов $g_{\mu\nu}$ выберем скалярный масштабный фактор a и тензорное поле \mathfrak{S}_{ij} , а также функции хода N и сдвига V_i (соответственно пространственная, временная и пространственно-временная части метрики). Последние две функции являются геометрическими объектами (скаляр и вектор соответственно) в классе лагранжевых сопутствующих систем отсчета (5.1.18)–(5.1.26). Для других координатных систем эти функции следует заменить на их калибровочно-инвариантные аналоги (см. (4.5.17)), однако для получения лагранжиана квазифридмановской Вселенной знание этих преобразований в явном виде не требуется (см. сноску на с. 162).

Итак, наша задача состоит в том, чтобы заменить АДМ-переменные N , N^i и γ^{ij} на их инвариантные аналоги

$$N, \quad V^i, \quad a, \quad \mathfrak{S}_i^j \quad (5.3.1)$$

и выразить через них действие (1.9.19). Напомним, что N и γ^{ij} — это геометрические объекты в классе (5.1.21). Поэтому, по сути, речь идет только о замене N^i на V^i , так как N^i и B^i не являются векторами (см. (5.1.26), (5.1.27)). Вихревой вектор является частью полного вектора V^i , а потенциал $\xi_i = 0$.

Для решения этой задачи достаточно произвести данную замену в любой из координатных сеток (t, \mathbf{x}) (5.1.21) и далее обобщить полученный лагранжиан на весь класс преобразований (5.1.21), воспользовавшись инвариантностью ряда (5.3.1)¹⁾. Наиболее подходящей для этих целей является лагранжево-эйлерова система отсчета (5.1.18), поскольку для нее $N_i = V_i$ (см. (5.1.20)).

Выделим из общих полей вихревую и тензорную моды возмущений согласно (5.1.24):

$$V^i = V_E^i - a^{-3}\mathfrak{S}^i, \quad \gamma_{ij} = a^2\left(\delta_i^j - 2\mathfrak{S}_i^j\right), \quad (5.3.2)$$

где V_E^i — потенциальный вектор,

$$V_{Ei,j} - V_{Ej,i} = \mathfrak{S}_{|i}^i = \mathfrak{S}_i^i = \mathfrak{S}_{i,j}^j = 0,$$

¹⁾ Подобным же образом лагранжиан (5.3.3) обобщается на *любые* системы отсчета.

запятая и вертикальная черта означают соответственно пространственную и ковариантную производные в полной метрике γ_{ij} .

Эти соотношения будем считать определением мод. Подставляя (5.3.2) в формулы гравитационного действия (1.9.19) и раскладывая его до второго порядка по V^i и \mathfrak{S}_i^j (1), получаем лагранжеву плотность квазихаббловского потока в виде функции пяти инвариантных гравитационных (N , V_E^i , a , \mathfrak{S}^i и \mathfrak{S}_i^j) и материальных (m) полей:

$$S = S \left[N, V_E^i, a, \mathfrak{S}^i, \mathfrak{S}_i^j, m \right] = \int (\mathcal{L}_{gS} + \mathcal{L}_{gV} + \mathcal{L}_{gT} + N \mathcal{L}_m) a^3 dt dy, \quad (5.3.3)$$

где

$$\begin{aligned} \sqrt{\gamma_n} (\mathcal{L}_g + N \mathcal{L}_m) &= a^3 (\mathcal{L}_{gS} + \mathcal{L}_{gV} + \mathcal{L}_{gT} + N \mathcal{L}_m) + \frac{\mathcal{E}^{\mu}_{,\mu}}{8\pi G}, \\ 8\pi G \mathcal{L}_{gS} &= \frac{\mathcal{H}(2\mathcal{D} - 3\mathcal{H})}{N} + \frac{\nabla Q \nabla(Q - 2N)}{a^2} = \\ &= -\frac{3\hat{\mathcal{H}}^2}{N} + \frac{\mathcal{D}^2}{3} + \left(\frac{\nabla \hat{Q}}{a} \right)^2 - \left(\frac{\nabla N}{a} \right)^2, \end{aligned} \quad (5.3.4)$$

¹⁾ Для этого вычислим кривизны лагранжево-эйлеровых сечений $t_c = \text{const}$. Для внутренней кривизны имеем (знак «п» у детерминанта опускаем):

$$\begin{aligned} \sqrt{\gamma} NR^{(3)} &= \sqrt{\gamma} N \gamma^{ij} (\Gamma_{il}^k \Gamma_{jk}^l - \Gamma_{ij}^l \Gamma_{lk}^k) - N \left(\frac{(\gamma \gamma^{ij})_{,j}}{\sqrt{\gamma}} \right)_{,i} = \\ &= 2a \nabla Q \nabla(Q - 2N) - a \mathfrak{S}_{i,k}^j \mathfrak{S}_j^{i,k} + \mathcal{E}_{1,i}^i, \quad \mathcal{E}_1^i = 2a \mathfrak{S}_j^k \mathfrak{S}_{i,k}^j - \frac{N(\gamma \gamma^{ij})_{,j}}{\sqrt{\gamma}}. \end{aligned}$$

Для внешней кривизны из (1.9.10) получаем:

$$\begin{aligned} \kappa &= \frac{(\sqrt{\gamma})'}{\sqrt{\gamma}} - \mathcal{D}, \quad \kappa_{ij} \kappa^{ij} - \kappa^2 = \\ &= -\frac{1}{4} \left(\dot{\gamma}_{ij} \dot{\gamma}^{ij} + \left(\frac{\dot{\gamma}}{\gamma} \right)^2 \right) + \mathcal{D} \frac{\dot{\gamma}}{\gamma} + \dot{\gamma}^{ij} V_{i|j} + \frac{1}{2} V^{i|j} (V_{i|j} - V_{j|i}) + \mathcal{E}_{2|i}^i, \end{aligned}$$

где $\mathcal{E}_2^i = V^i{}_{|j} V^j - \mathcal{D} V^i$. Вычисляя отдельно слагаемые

$$\begin{aligned} -\frac{1}{4} \left(\dot{\gamma}_{ij} \dot{\gamma}^{ij} + \left(\frac{\dot{\gamma}}{\gamma} \right)^2 \right) &= -6\mathcal{H}^2 + 4\mathcal{H} (\mathfrak{S}_i^j \mathfrak{S}_j^i)' + \dot{\mathfrak{S}}_i^j \dot{\mathfrak{S}}_j^i, \\ \dot{\gamma}^{ij} V_{i|j} &= -2\mathcal{H}\mathcal{D} + 2 (V^i \dot{\mathfrak{S}}_i^j)_{,j}, \\ a^6 V^{i|j} (V_{i|j} - V_{j|i}) &= \mathfrak{S}^{i|j} \mathfrak{S}_{i|j} - (\mathfrak{S}^j \mathfrak{S}_{i|j}^i)_{,i} \end{aligned}$$

и учитывая, что $\sqrt{\gamma} = a^3 (1 - \mathfrak{S}_i^j \mathfrak{S}_j^i)$, получаем (5.3.3)–(5.3.7).

$$\mathcal{L}_{gV} = \frac{\delta_{ij} \nabla \mathfrak{S}^i \nabla \mathfrak{S}^j}{32\pi G a^6}, \quad (5.3.5)$$

$$\mathcal{L}_{gT} = \frac{1}{16\pi G} \left(\mathfrak{S}_{i,\mu}^j \mathfrak{S}_j^{i,\mu} - m_G^2 \mathfrak{S}_i^j \mathfrak{S}_j^i \right), \quad (5.3.6)$$

$$\mathcal{E}^0 = 2a^3 H \mathfrak{S}_i^j \mathfrak{S}_j^i, \quad (5.3.7)$$

$$\begin{aligned} \mathcal{E}^i = & 2a \mathfrak{S}_j^k \mathfrak{S}_{k,i}^j - a \mathfrak{S}_j^k \mathfrak{S}_{i,k}^j + \\ & + a^3 \left(2N Q_{,i} - \frac{1}{2} \mathcal{D} V^i + \mathfrak{S}_j^i V^j + \frac{1}{2} V_{,j}^i V^j - \frac{1}{4} \mathfrak{S}_{,j}^i \mathfrak{S}^j \right). \end{aligned}$$

Здесь введены следующие вспомогательные обозначения:

$$Q \equiv -\ln |a|, \quad \hat{Q} \equiv Q - N,$$

$$\mathcal{H} \equiv \frac{\dot{a}}{a} = -\dot{Q}, \quad \hat{\mathcal{H}} \equiv \mathcal{H} - \frac{\mathcal{D}}{3},$$

$$\mathcal{D} \equiv V_{|i}^i = \frac{(a^3 V_E^i)_{,i}}{a^3},$$

$$m_G^2 \equiv 4\dot{H} + 6H^2 + 16\pi G \mathcal{L}_m.$$

Полученное действие справедливо вплоть до членов второго порядка малости. В выражении для \mathcal{D} учтено, что квадратичный член в детерминанте $\sqrt{\gamma_n} = a^3 \left(1 - \mathfrak{S}_i^j \mathfrak{S}_j^i \right)$ является превышением точности, поскольку величина \mathcal{D} отсутствует в нулевом порядке. По этой же причине при умножении N на члены второго порядка полагаем $N = 1$.

Каждое из слагаемых в (5.3.3) является скаляром по отношению к полной группе преобразований лагранжевых систем отсчета (5.1.20), (5.1.21). Полная производная в действии не отражается на уравнениях эволюции. Лагранжианы (5.3.5) и (5.3.6) существуют только во втором порядке возмущений, тогда как скалярная плотность (5.3.4) описывает полное течение хаббловского потока. Приравнивая вариации (5.3.3) по \mathfrak{S}^i и \mathfrak{S}_i^j к нулю, получаем соответствующие уравнения движения векторной (4.5.9) и тензорной (4.5.8) мод возмущений¹⁾.

Наиболее интересен блок скалярных движений действия (5.3.3). Варьируя по N , V_E^i и a , получаем классические урав-

¹⁾ Пространственное уравнение вихревой моды $\mathfrak{S}_i = S_i$ получается варьированием действия по переменной ξ_i , которая в данном классе преобразований равна нулю, оно также следует из тождества Бианки (4.5.5). Заметим, что масса гравитона зануляется на классическом фоне: $m_G = 0$.

нения скалярного течения с учетом нулевого и первого порядков (ср. (1.9.20), (1.9.21), (5.2.1)):

$$\frac{1}{6} \mathcal{K}^2 + \frac{\Delta Q}{a^2} = 4\pi G \varepsilon, \quad (5.3.8)$$

$$- \left(\frac{\mathcal{H}}{N} \right)_{,i} = 4\pi G u_\mu T_i^\mu, \quad (5.3.9)$$

$$\frac{1}{2} \mathcal{K}^2 + \frac{\dot{\mathcal{K}}}{N} + \frac{\Delta \hat{Q}}{a^2} = -12\pi G p_{\text{tot}}, \quad (5.3.10)$$

где $\mathcal{K} = u^\mu_{;\mu} = 3\hat{\mathcal{H}}/N$ — внешняя кривизна, $\varepsilon = u_\mu u_\nu T^{\mu\nu}$, $p_{\text{tot}} = \gamma^{\mu\nu} T_{\mu\nu}/3$ (см. (1.9.14), для S-моды $u_\mu = n_\mu$).

Разница (5.3.8) и (5.3.10) с исключением члена \mathcal{K}^2 дает уравнение в форме второго уравнения Фридмана:

$$- \frac{\dot{\mathcal{K}}}{N} + \frac{\Delta(2Q + N)}{a^2} = 12\pi G (\varepsilon + p_{\text{tot}}), \quad (5.3.11)$$

а дифференцируя (5.3.8) и исключая $\dot{\mathcal{K}}$ с помощью (5.3.11), получаем уравнение сохранения энергии (ср. (5.2.2)):

$$\dot{\varepsilon} + 3\hat{\mathcal{H}}(\varepsilon + p_{\text{tot}}) = -\frac{1}{4\pi G a^2} \Delta \left(\frac{\mathcal{H}}{N} \right). \quad (5.3.12)$$

Заметим, что для скалярной моды правая часть (5.3.9) равна нулю (см. (5.1.19)): $u_\mu T_i^\mu = 0$. Если считать, что условие связи (5.3.9) выполнено, то отсюда получаем совпадение уравнений (5.3.8) и (5.2.1) ($\mathcal{K} = 3H_v$), (5.3.11) и (5.2.4) ($\Delta \hat{Q} = -H\Delta N$, $\varkappa = 2\Delta Q/3$), а также закон сохранения энергии (5.2.2) с равной нулю правой частью (5.3.12).

5.4. Квантовая Вселенная и квазиклассический предел

Представление гравитационного действия \mathcal{L}_g неоднородной Вселенной в виде функционала (5.3.3) скалярного масштабного фактора \mathbf{a} или заменяющего его экспоненциального фактора Q :

$$\mathbf{a} \equiv e^{-Q},$$

открывает возможность для квантования квазиаббловского потока с помощью метрических полей N , V_i , Q и материальных

степеней свободы m , ограниченных функциональными связями (5.3.8) и (5.3.9) ¹⁾.

Моды V и T являются пробными полями хаббловского потока, и их квантование выполняется стандартным образом с помощью полученных действий ²⁾ (5.3.5) и (5.3.6). Эти моды можно отнести к материальным степеням свободы и рассматривать как дополнительные поля с ненулевым спином, поскольку их гравитационное воздействие на квазифридмановскую геометрию ничтожно: уравнения (5.2.1) и (5.2.2) имеют существенно скалярный вид.

Скалярная мода задает структуру всего квазихаббловского потока и определяет его эволюцию в сопутствующем времени φ , которое с геометрической точки зрения есть скалярный потенциал полной скорости потока:

$$u_\mu = \frac{\varphi_{,\mu}}{w}, \quad (5.4.1)$$

где $w^2 = g^{\mu\nu} \varphi_{,\mu} \varphi_{,\nu}$ — нормировочный скаляр, который можно считать кинетическим членом скалярного поля φ . Формула (5.4.1) правильно описывает нулевой и первый порядки скалярной моды и никак не ограничивает природу и состояние материальных полей m , из которых состоит сам поток ³⁾. Поле φ можно считать общей кумулятивной переменной всех материаль-

¹⁾ Каноническое квантование квазифридмановских геометрий (5.3.3)–(5.3.7) выполняется путем формального обобщения метода де Витта (развитого для строго фридмановских моделей). Для этого следует перейти к операторам полей $\phi_a = (N, V_E^i, Q, \varphi, \dots)$ и канонических импульсов полей $\pi_a \equiv \delta L / \delta \dot{\phi}_a$, действующим на функционал состояний $\Psi[\phi_a]$ и удовлетворяющим коммутационным соотношениям $[\phi_a(t, \mathbf{y}) \pi_b(t, \mathbf{y}')] = i \delta_{ab} \cdot \delta^{(3)}(\mathbf{y} - \mathbf{y}')$, где $L[\phi_a] = \int (\mathcal{L}_g + N \mathcal{L}_m) a^3 dy$ — лагранжиан системы, $\delta^{(3)}(\mathbf{y})$ — трехмерная дельта-функция Дирака. В результате получаем уравнения Шредингера для операторов полной энергии (гамильтониана) и импульсов системы, которые в квазиклассическом пределе соответствуют ограничениям (5.3.8) и (5.3.9). Этот подход лишен недостатков оригинального формализма де Витта, поскольку здесь квантуются *все* степени свободы (компоненты) метрического тензора, а не только та их часть, которая описывает однородную фридмановскую моду. Подробности общего метода можно найти в цитируемой литературе, а в книге мы детальнее рассмотрим квантовые свойства сектора возмущений, ответственного за генерацию наблюдаемой структуры Вселенной (см. главы 8–10).

²⁾ Действие для V -моды следует дополнить материальным членом, зависящим от v_i , который также имеет вид пробного поля, поскольку хаббловский поток обладает чисто скалярной симметрией (напомним, что v_i — переменная первого порядка малости, см. (4.5.2)).

³⁾ Во втором и последующих порядках следует принимать во внимание физическое состояние среды и взаимодействия между модами возмущений. Здесь мы эти вопросы не рассматриваем.

ных полей $m = 1, 2, \dots, M$, остальные $M - 1$ переменных задают изометрические степени свободы (см. раздел 6.2).

Таким образом, в роли основных динамических переменных ранней Вселенной выступают локальный фактор объемного расширения Q и скалярный потенциал скорости материи φ . Любая из этих переменных (или их комбинация) может играть роль времени в квазиклассическом пределе.

Для того чтобы понять, как в ранней Вселенной могла возникнуть геометрия с квазифридмановской структурой, ограничим максимальным образом число независимых переменных ($M = 1$) и опустим вихревую моду.

Запишем метрику квазихаббловского потока в лагранжево-эйлеровой системе отсчета, где роль координат $y^a = (y^0, y^p)$ выполняют четыре скалярных поля $y^0 \equiv \varphi$ и $y^p \equiv \mathbf{y}$ (см. раздел (5.1)):

$$\begin{aligned} ds^2 &= \frac{d\varphi^2}{W^2} - 2(\nabla v d\mathbf{y}) \frac{d\varphi}{W} - e^{-2Q} dl^2 = \\ &= \left(\frac{d\varphi}{W} + \mathbf{a} v d\mathbf{y} \right)^2 - \mathbf{a}^2 \left(dl^2 + (v d\mathbf{y})^2 \right) = \\ &= \frac{d\varphi^2}{w^2} - E_{pq} \left(\mathbf{a} dy^p - v^p \frac{d\varphi}{W} \right) \left(\mathbf{a} dy^q - v^q \frac{d\varphi}{W} \right), \end{aligned} \quad (5.4.2)$$

где

$$W \equiv w \sqrt{1 + \mathbf{v}^2}, \quad (5.4.3)$$

$$\mathbf{v}^2 = v_p v^p = e^{2Q} E^{pq} v_{,p} v_{,q}, \quad (5.4.4)$$

v — потенциал пекулярной скорости потока ($\mathbf{v}_{\text{pec}} \equiv \mathbf{v} = v_p$):

$$v_p = -e^Q v_{,p}, \quad v^p = E^{pq} v_{,q}, \quad \nabla \equiv \frac{\partial}{\partial y^a}$$

$\mathbf{E} = E_{pq}$ — метрика внутреннего пространства:

$$dl^2 = E_{pq} dy^p dy^q, \quad E_{ps} E^{sq} = \delta_p^q.$$

Выпишем также компоненты полной скорости материи и контравариантные компоненты метрики в координатах (5.4.2):

$$\begin{aligned} u_a &= (w^{-1}, 0, 0, 0), \quad u^a = \left(w, e^Q \frac{v^p}{\sqrt{1 + \mathbf{v}^2}} \right), \\ g^{ab} &= \left(\begin{array}{cc} w^2 & \text{sym} \\ we^Q \frac{v^p}{\sqrt{1 + \mathbf{v}^2}} & e^{2Q} \left(\frac{v^p v^q}{1 + \mathbf{v}^2} - E^{pq} \right) \end{array} \right). \end{aligned}$$

Как видно в этих *геометрических* координатах y^a , квадрат импульса поля φ есть нулевая компонента метрического потенциала: $w^2 = g^{\varphi\varphi}$.

В пренебрежении изометрическими степенями свободы решение (5.4.1), (5.4.2) является точным со стандартным действием, полученным в разделе 1.9. Здесь мы имеем три задающих скаляра w , v и Q и метрику внутреннего пространства E_{pq} , содержащую тензорные степени свободы. Из трех скаляров динамическим является только Q , так как переменные w и v не имеют собственных импульсов в действии. Можно сказать по-другому: коэффициенты перед их временными производными оказались пренебрежимо малыми.

Это могло произойти в силу того, что в ходе релаксации в очень ранней Вселенной временная координата была подвержена «растяжению» (*временной инфляции*), за счет быстрого изменения потенциала $g_{\varphi\varphi}$, что привело к подавлению импульсов \dot{w} и \dot{v} в начальном лагранжиане и к возникновению *дальнодействующей* гравитации в том виде, в котором она нам известна (ОТО). По этой причине уравнения (5.3.8) и (5.3.9) *уже являются* квазиклассическим пределом квантовой гравитации, и их повторная процедура канонического и вторичного квантования малоинформативна (см. сноску ²) на с. 166).

По конечной метрике ранней Вселенной сложно восстановить исходную геометрическую симметрию и механизм ее нарушения. Мы рассмотрели лишь простейший пример возможного решения ¹⁾. Однако нельзя не увидеть общие закономерности подобного сценария.

¹⁾ Из решения (5.4.2) естественно возникает метрика слабонеоднородной Вселенной. Перепишем уравнения (1.9.9) и (5.1.20) в терминах переменных y^a , заметив, что сопутствующее время t_c есть функция φ : $t_c = t_c(\varphi)$. Отсюда восстанавливаем связь с введенными ранее АДМ-переменными: $N = \dot{\varphi}/w$, $\gamma_{ij} = aE$, $V_i = \dot{\varphi}/W \nabla v$ (или в линейном порядке $V_i = \nabla v$, см. (5.1.18)).

Из (5.3.9) и (5.1.1) в нулевом и первом порядках получаем

$$-w Q_{,\varphi} = f(\varphi) \equiv H_c = \frac{\mathcal{H}}{N},$$

что напрямую связывает переменные w и $Q_{,\varphi}$. Отсюда также следует сохранение энергии на классических траекториях движения материи (см. (5.3.12)) и следующее выражение для внешней кривизны:

$$\mathcal{K} = 3H_c - \mathcal{D}, \quad \mathcal{D} = \frac{\Delta v}{a^2}.$$

Из (5.3.8) получаем связь между Q и плотностью энергии $\epsilon = u_\mu u_\nu T^{\mu\nu}$:

$$\frac{3}{2}H_c^2 + \frac{\Delta\phi}{a^2} = 4\pi G\epsilon, \quad \phi \equiv Q - H_c v.$$

Можно предположить, что в геометрическом описании физического мира отсутствовали исходные (фундаментальные) *материальные* поля, а имелись только фундаментальные *геометрические координаты* y^a и *метрические потенциалы* g^{ab} (индексы a и b и координаты с сигнатурой времени типа φ могли быть многомерны). Понятия материи, пространства и времени возникли взаимосвязанно и как *вторичные* переменные (поля), удобные для описания нарушенной геометрии. Результат нарушения воспринимается нами как эволюция (движение или релаксация) полей, удовлетворяющих ограничениям в виде «законов сохранения энергии и импульса».

В простейшем подходе роль основных материальных полей играют сами «фундаментальные координаты» y^a , а квадраты импульсов материальных полей и их взаимодействия описываются «геометрическими потенциалами» g^{ab} . Роль времени и пространства играют *классические ожидания* ряда этих переменных («растянутых» инфляцией как во времени, так и в пространстве), которые мы обозначим как координаты классического четырехмерного многообразия x^μ с метрическим тензором $g^{\mu\nu}$. В результате получаем (как минимум) четыре скалярных поля φ, y^1, y^2, y^3 и десять динамических потенциалов g^{ab} , инвариантных относительно *любых* координатных преобразований $x^\mu = (t, x^i)$:

$$y^a = y^a(x^\mu), \quad g^{ab} = g^{\mu\nu} y^a_{,\mu} y^b_{,\nu}. \quad (5.4.5)$$

Первое из полей (φ) могло стать «инфлатоном», растянувшим три пространственные координаты на ранних стадиях расширения Вселенной. Эффективным потенциалом, обеспечившим инфляционную динамику, могла быть одна из аддитивных частей фундаментального лагранжиана, зависящая только от y^a и g^{ab} , но независящая от производных $g^a_{,c}$. Эту часть мы интерпретируем как «лагранжиан физического поля φ »: $\mathcal{L}(\varphi, w)$ (в w^2 не входят производные метрических потенциалов (см. (5.4.1)), и $g^{\varphi\varphi} = w^2$ (см. (5.4.2))).

В более общем подходе полей типа φ может быть несколько (обозначим их $\varphi_m, m = 1, 2, \dots$). Тогда лагранжиан является произвольной функцией самих полей φ_m и их дифференциальных билинейных форм:

$$w^2_{mn} = g^{\mu\nu} \varphi_{m,\mu} \varphi_{n,\nu}. \quad (5.4.6)$$

Этот метод можно обобщить и на поля ненулевых спинов.

На этом пути мы встречаем пока больше вопросов, чем получаем ответов. В ситуации, когда неизвестна исходная геометрическая структура, но известен конечный «продукт» в форме квази-хаббловского потока (5.4.2), мы должны исходить из реально наблюдаемого факта *больших* (протяженных) масштабов изотропной гиперповерхности u и относительно *малых* (наблюдаемых) масштабов ее структуры. Это позволяет считать пространственную гиперповерхность ранней Вселенной большим *классическим* объектом, на котором существуют *квантовые* метрические и материальные возмущения, малые по амплитуде и масштабу, что и определяет квазиклассический предел квантовой теории гравитации.

Для определения полного лагранжиана скалярной геометрической моды необходимо обратиться к материальной части действия $\mathcal{L}_m = \mathcal{L}(m)$. Преобразуем материальные степени свободы m , выделив коллективные адиабатические скаляры φ и $w = \dot{\varphi}/N$ (здесь время сопутствующее: $t = t_c(\varphi)$, см. (5.3.3)), а оставшиеся изометрические переменные, включающие поля и их производные, обозначим через \hat{m} (подробнее см. раздел 6.2). Предполагая, что задающий лагранжиан зависит от исходных полей и билинейных форм типа (5.4.5), (5.4.6), мы в результате такого преобразования получаем лагранжеву функцию $\mathcal{L}(m)$, зависящую от φ и $w = \dot{\varphi}/N$, а также от совокупности переменных \hat{m} с кинетическими членами, которые могут содержать разные компоненты метрики в зависимости от внутренних симметрий полей.

Тогда общее действие скалярной части имеет следующий вид (см. (5.3.3), (5.3.4)):

$$S = S[N, V^i, Q, \varphi, \hat{m}] = \int [\mathcal{L}_{\text{GS}} + N\mathcal{L}(\varphi, \dot{\varphi}/N, \hat{m})] a^3 dt dy. \quad (5.4.7)$$

Покажем, что в этом действии можно выделить аддитивные фридмановскую моду $S[a, \varphi] = \int L(a, \varphi) a^3 dt dy$ и кумулятивную (адиабатическую) моду скалярных возмущений $S[q] = \int L(q) a^3 dt dy$. Для этого представим исходные переменные в терминах отклонений от фридмановского фона:

$$N = 1 + \delta N, \quad Q = Q_c + q, \quad (5.4.8)$$

где $Q_c \equiv -\ln a_c$, $\varphi = \varphi(t)$ в силу выбора времени, вектор V^i сам является отклонением, изометрические переменные в явном виде не выписываем (их учет даст дополнительное слагаемое $S[\hat{m}]$ в действии (5.4.11)). Полагая (5.4.8) определением функций δN

и q , разложим материальную часть действия по переменным δN , V^i и q до величин второго порядка включительно:

$$\begin{aligned} N\mathcal{L}(\varphi, \dot{\varphi}/N, \hat{m}) &= \mathcal{L}_1 - \varepsilon_1 \delta N - \frac{1}{2} \frac{\partial \varepsilon}{\partial N} \delta N^2 + \dots = \\ &= p_c - \varepsilon_c \delta N + (\varepsilon + p) \beta^{-2} \left(\hat{\delta}_p \delta N + \frac{1}{2} \delta N^2 \right) + \dots, \end{aligned}$$

где индексом «1» помечены функции \mathcal{L} и $\varepsilon \equiv -\mathcal{L} - \partial \mathcal{L} / \partial \ln N$ при $N = 1$, β^{-2} — «средний» обратный квадрат скорости звука. Функция первого порядка малости $\hat{\delta}_p$ учитывает кумулятивное действие изометрических мод возмущений:

$$\varepsilon_1 = \varepsilon_c + (\varepsilon + p) \hat{\delta} + \dots, \quad \hat{\delta} \equiv -\beta^{-2} \hat{\delta}_p \quad (5.4.9)$$

(здесь разложение идет до членов первого порядка включительно). Заметим, что член второго порядка $\propto V^i$ в данном разложении отсутствует, поскольку он умножается на нулевой вектор $u_\mu T_i^\mu = 0$ (условие сопутствия S-моды). Также отсутствуют вариации по пространственным компонентам метрики, поскольку они содержатся в членах, начиная с третьего порядка малости.

Введем вспомогательные обозначения:

$$F \equiv 3H^2 - 8\pi G \varepsilon_c, \quad \Pi = H\delta N + \dot{q}, \quad (5.4.10)$$

и разложим с учетом (5.4.8) член в гравитационном действии (индекс «с» у a и H опущен):

$$\frac{\mathcal{H}^2}{N} = H^2 - 2H\dot{q} - H^2 \delta N + \Pi^2.$$

Подставляя эти разложения в (5.4.7), получаем скалярную компоненту действия в следующей форме:

$$\begin{aligned} S[a, \varphi, q] &= \int dt dy \left[a^3 L(a, \varphi) + a^3 L(q) + \right. \\ &\left. + \frac{a^3}{8\pi G} \left(F\delta_1 + \frac{1}{2} m_G^2 \delta_2 + \Pi \delta_3 + 2\Pi_{,i} \delta^i \right) + \frac{1}{4\pi G} S_{,\mu}^\mu \right], \quad (5.4.11) \end{aligned}$$

где

$$L(a, \varphi) = p_c - \frac{3H^2}{8\pi G}, \quad (5.4.12)$$

$$L(q) = \frac{\gamma}{8\pi G} \left[\beta^{-2} \dot{q} \left(\dot{q} - 2H\hat{\delta}_p \right) - \left(\frac{\nabla q}{a} \right)^2 \right], \quad (5.4.13)$$

$$\delta_1 \equiv e^{-3q} \delta N - \frac{\dot{q} \left(\dot{q} - 2H\hat{\delta}_p \right)}{2H^2 \beta^2},$$

$$\begin{aligned}\delta_2 &\equiv e^{-3q} - 1 + \frac{\dot{q} (\dot{q} - 2H\widehat{\delta}_p)}{2H^2\beta^2}, \\ \delta_3 &\equiv -3\Pi + \gamma\beta^{-2} (H\delta N - \dot{q} + 2H\widehat{\delta}_p), \\ \delta^i &\equiv V^i - H^{-1}q_{,i}, \\ S^0 &\equiv a^3H (1 - e^{-3q}) + \frac{a}{2H} (\nabla q)^2, \\ S^i &\equiv \frac{a^3\mathcal{H}}{N} V^i.\end{aligned}$$

Формулы (5.4.12) и (5.4.13) описывают соответственно лагранжевы плотности фридмановской геометрии и S-моды космологических возмущений. Вектор S^μ не оказывает влияния на уравнения движения. Если считать классическим фон ($F = m_G = 0$) и выполненным условие гравитационного дальнего действия ($\Pi = 0$, см. (5.3.9)), то остальные члены действия (5.4.11) оказываются равными нулю.

5.5. Образование и разрушение хаббловских потоков

Рассмотрим асимптотическое поведение скалярного поля q в разных частотных пределах.

В коротковолновом пределе ($k > \bar{H}$) важную роль играют градиенты поля, определяющие неоднородную структуру возмущений: пекулярную скорость ($\propto \nabla q$) и кривизну или возмущение плотности ($\propto \Delta q$). В этой асимптотике функция q характеризует степень разрушения хаббловских потоков.

В длинноволновом пределе пространственные градиенты пренебрежимо малы: скорость, плотность и давление материи невозмущены. В этой асимптотике скаляр q представляет собой «вмороженное» гравитационное поле, зависящее только от пространственных координат:

$$q \rightarrow q_0 = q_0(\mathbf{x}).$$

Синхронно-сопутствующая квазиизотропная метрика

$$ds^2 = dt^2 - a^2(t)\bar{a}^2(\mathbf{x}) d\mathbf{x}^2 \quad (5.5.1)$$

описывает главный нулевой член разложения геометрии Вселенной по производным функции $\bar{a} = \bar{a}(\mathbf{x})$. Решение (5.5.1) нелинейно и предполагает разложение по малому параметру

$$\frac{|\nabla \ln \bar{a}|}{\bar{H}} \ll 1.$$

Для малых q_0 справедливо равенство

$$\ln \bar{a} = -q_0.$$

С точки зрения полной геометрии q_0 есть начальная крупномасштабная функция q на гиперповерхности Коши $t = \text{const}$.

Как мы видим, реальная фоновая геометрия (5.5.1) квазиизотропная, а не фридмановская. В малой окрестности любой точки \mathbf{x} геометрию (5.5.1) можно считать фридмановской с масштабным фактором $a(t)$, поскольку значение $q_0(\mathbf{x})$ локально неотлично от константы и его всегда можно положить равным нулю путем линейного растяжения пространственных координат. Но в конечной области пространства метрика неоднородная с плавно изменяющимся в пространстве метрическим потенциалом. Подчеркнем, что эта неоднородность физическая и не может быть устранена никакими преобразованиями координат.

Таким образом, представление о том, что в качестве фонового пространства можно использовать глобально фридмановскую метрику, строго говоря, неверно. Пока возмущения малы, мы можем сделать так технически, поскольку теория возмущений допускает перенормировку фона¹⁾ (см. раздел 4.6). Однако, по сути дела, фридмановского фона у нас нет: есть только локальный фридмановский фон, а глобальная геометрия квазиизотропна и описывается функцией $\bar{a}(\mathbf{x})$. Последнее утверждение очевидно для нелинейных функций $\bar{a}(\mathbf{x})$, так как никакие перенормировки в этом случае невозможны.

Тем не менее, в длинноволновом пределе в главном порядке по q хаббловские потоки локально не искажены (напомним, что теперь q вовсе не является мерой их искажения, как это было в коротковолновом пределе):

$$\delta \mathbf{V} \equiv \delta \dot{\mathbf{r}} = H(t) \delta \mathbf{r}, \quad H(t) = \frac{\dot{a}}{a}, \quad (5.5.2)$$

где $\delta \mathbf{r} = a \bar{a} \delta \mathbf{x}$. Надо понимать, что к метрике (5.5.1) имеются малые поправки $\sim \nabla q_0 / \bar{H}$, и в этом порядке хаббловские потоки слабо искажены. Принципиальное значение имеет ответ на вопрос, растут ли эти искажения во времени или затухают? В первом случае эволюция ведет к разрушению имеющихся хаббловских потоков, а во втором — к их построению.

Итак, мы имеем две асимптотики геометрии, которые эволюционно связаны друг с другом и описывают два противо-

¹⁾ Поэтому теорию возмущений можно корректно применять в конечных областях пространства с ограниченной вариацией $\bar{a}(\mathbf{x})$: $|q_0(\mathbf{x})| \ll 1$.

положных процесса: хаббловские потоки (длинноволновый предел) и структуру (коротковолновый предел). Для количественной оценки динамики производства структуры воспользуемся решением (5.5.1). В первом порядке по градиентам q_0 из уравнений (4.5.16)–(4.5.18) получаем¹⁾:

$$\begin{aligned} q &= q_0 + \int \delta_p \frac{da}{a}, & \Phi &= \phi(a)q_0, \\ v_{\text{pec}} &= -\nu(a) \nabla q_0, & \delta_\varepsilon &= g(a) \Delta q_0, \end{aligned} \quad (5.5.3)$$

где

$$\nu(a) = \frac{1}{a^2} \int \frac{da}{H}, \quad \phi(a) = \gamma \bar{H}^2 g(a) = 1 - \frac{H}{a} \int \frac{da}{H} \quad (5.5.4)$$

— факторы роста пекулярной скорости (5.1.15), гравитационного потенциала (4.5.16) и возмущения плотности (4.5.15).

Критическим режимом для функций роста является линейное космологическое расширение, при котором

$$\gamma = 1, \quad a \propto t.$$

В этом случае ν и g постоянны во времени и упомянутые асимптотики эволюционно разделены: хаббловские потоки сохраняются там, где они есть, и генерация новой структуры не идет.

При замедленном расширении ($\gamma > 1$) функции ν и g монотонно растут со временем: возмущения нарастают, разрушаются изначальные хаббловские потоки и развивается неоднородная структура. Со временем длинноволновые возмущения эволюционируют в коротковолновую область (функция $\bar{H} = \dot{a}$ монотонно уменьшается).

В случае инфляции ($\gamma < 1$) функции ν и g затухают со временем: новая структура не производится, напротив, идет строительство хаббловских потоков все в больших и больших масштабах. При ускоренном расширении первоначально короткие волны оказываются в длинноволновой области (функция \bar{H} монотонно растет).

Мы можем заключить, что гравитация в равной степени обладает двумя динамическими свойствами: *отталкиванием* (генерация хаббловских потоков) и *притяжением* (генерация структуры). Какая из этих двух сторон гравитации будет реализована, зависит от уравнения состояния материи (см. (4.1.9)): инфляция (отталкивание) реализуется при $(\varepsilon + 3p) < 0$, а замедленное

¹⁾ Скаляр анизотропного давления S , не имеющий отношения к гравитационной неустойчивости, положен равным нулю.

расширение (притяжение) — в противоположном случае, при $(\epsilon + 3p) > 0$.

5.6. Антиколлапс и энтропия

Остановимся подробнее на квазиизотропном решении (5.5.1), которое описывает локальный хаббловский поток (5.5.2) с функцией $H = H(t)$, не зависящей от \mathbf{x} . Столь высокосимметричное пространство-время, близкое по свойствам к идеальной геометрии Фридмана, имеет перед ней громадное преимущество. В отличие от геометрии Фридмана, его не надо специально готовить: квазиизотропная геометрия возникает естественным образом динамически в ходе инфляционного расширения ($\ddot{a} > 0$).

Здесь мы сталкиваемся с удивительным свойством гравитации: ускоренное *локальное* объемное расширение хаотически распределенной материи формирует регулярный *коррелированный* хаббловский поток в причинно-несвязанной области пространства $\ell \equiv k^{-1} > \dot{a}^{-1}$. Поскольку функция \dot{a}^{-1} уменьшается со временем, то все новые и новые расширяющиеся части материи пересекают и уходят за хаббловский радиус (все меньшие и меньшие лагранжевы размеры ℓ попадают в причинно-несвязанную область). В ходе этого процесса начальные складки и неоднородности пространственной геометрии в масштабах, уходящих за хаббловский размер, «выглаживаются», идеально стыкуясь друг с другом в причинно-несвязанных областях. Процесс выглаживания сопутствующих масштабов продолжается во времени, в результате чего асимптотически выстраивается идеальная изотропная пространственноподобная гиперповерхность с *однородным* распределением материи¹⁾.

Этот процесс напоминает коллапс наоборот. В процессе коллапса также происходит «выглаживание» и упорядочение гравитационного поля.

Действительно, начальное распределение сжимающейся материи, будь то компактное облако или звезда, достаточно хаотично, а гравитационное поле имеет кулоновский вид лишь на асимптотически больших расстояниях. Однако метрика черной дыры, образовавшейся при коллапсе, строго регулярна на

¹⁾ С точки зрения теории возмущений выглаживается и асимптотически исчезает падающая ветвь возмущений. Растущая ветвь вне хаббловского размера совместима с изотропным и однородным распределением материи, а ее пространственная кривизна застывает: $q = q(\mathbf{x})$.

любых радиусах и зависит лишь от двух параметров: массы и углового момента. В ходе коллапса происходит гравитационное «сгребание» (сгущивание) вещества в область под горизонт событий. В ходе же антиколлапса (инфляции) идет прямо противоположный процесс: вещество рассеивается («растягивается», расширяется) за область космологического горизонта, где его пространственное распределение стремится к однородному и изотропному.

Ответ на вопрос о том, какой процесс запускается — коллапс или инфляция, — зависит от свойств материального источника гравитационного поля: эффективного давления и пространственного распределения материи.

Столь замечательное свойство гравитации создавать упорядоченные системы из начально хаотических распределений материи обусловлено дальнодействием и релятивизмом гравитационного взаимодействия. Последнее особенно важно при ответе на вопрос, куда же девается энтропия, ведь она может только возрасти в замкнутой системе.

Конечно, полная энтропия возрастает как при коллапсе, так и при инфляционном расширении, однако необходимо также учитывать ее гравитационную часть, связанную с релятивистскими компактными поверхностями, отделяющими области пространства-времени, откуда частицы не могут попасть к наблюдателю. Информация об этих недоступных для конкретного наблюдателя областях для него утеряна, что выражается присутствием и ростом энтропии на границах областей (голографический принцип) со всеми вытекающими отсюда последствиями (например, тепловые потоки от границ вследствие квантово-гравитационного рождения пар частиц, см. разделы 2.6, 4.2). Эти границы являются горизонтами событий частиц для наблюдателя, и источником их является гравитирующая масса, охваченная горизонтом.

Поскольку площадь горизонта событий всегда конечна¹⁾, то и полная энтропия релятивистской системы остается конечной, но она перераспределяется от материальных к гравитационным степеням свободы как при коллапсе, так и при инфляционном расширении. Например, энтропия компактной гравитирующей

¹⁾ Для шварцшильдовской геометрии это $4\pi r_g^2$, где $r_g = 2GM$. Для пространственно-плоской фридмановской геометрии вместо r_g следует писать хаббловский радиус $H^{-1} = 2GM = 8\pi G\rho H^{-3}/3$ (уравнение Фридмана). В рамках ОТО энтропия равна числу планковских площадок, уместяющихся на поверхности горизонта событий (см. раздел 2.6).

системы заданной массы максимальна для состояния черной дыры. Аналогично, энтропия ускоренно расширяющейся области материи при заданной средней плотности максимальна для состояния де Ситтера.

Мы видим, как общая относительность связывает воедино термодинамику, гравитационное дальнее действие и квантовую механику. Измеряемые термодинамические и геометрические характеристики потока связаны с наблюдателем в точке x . Однако любые два наблюдателя, оказывающиеся с течением времени в причинно не связанных областях из-за ускоренного расширения материи, независимы и равноценны (принцип относительности) и должны измерять одни и те же локальные параметры инфляционного потока. Этим свойством обладает только квази-изотропное решение с функцией $H(t)$, не зависящей от x .

ЧТО НУЖНО ЗНАТЬ О МАТЕРИИ

Геометрические возмущения развиваются в гравитирующей среде, которая является источником всего квази Хаббловского потока. Для фридмановской эволюции необходимо знать плотность и давление материи. Какие дополнительные сведения о веществе требуются для описания динамики космологических неоднородностей? Существуют ли достаточно общие характеристики среды, не предполагающие детализацию ее микроскопического устройства? В данной главе мы получим ответы на эти и связанные с ними вопросы, а также рассмотрим в качестве примеров эволюцию возмущений плотности в до- и послерекомбинационной Вселенной.

6.1. Геометрические характеристики среды

Каждая из трех геометрических мод возмущений фридмановской Вселенной эволюционирует во внешнем однородном гравитационном поле, описываемом функцией времени $a(t)$, а также чувствует анизотропные натяжения в среде (если таковые имеются), которые описываются неприводимой частью тензора давлений $s_{\mu\nu}$ (см. (4.5.2)). Полный тензор натяжений зависит от физического состояния среды и является ее внутренней характеристикой.

Что еще надо знать о среде для определения динамики космологических возмущений?

Для мод T и V больше ничего, так как они не возмущают ни давление, ни плотность. Что касается моды S , то надо знать, как соотносятся между собой возмущения давления и плотности материи. Остановимся на этом подробнее.

Неоднородная часть потенциальных течений любой гравитирующей среды в квазифридмановской геометрии — S -мода космологических возмущений — описывается скалярным полем $q = q(t, \mathbf{x})$, которое является суперпозицией неоднородных частей потенциала общей скорости материи и масштабного фактора

Вселенной (см. (4.5.11), (5.1.1), (5.1.2)):

$$q = Hv - \frac{\delta a}{a} = \frac{a_c - a}{a} = \frac{a_v - b}{a}. \quad (6.1.1)$$

С геометрической точки зрения скаляр q задает меру отклонения между скалярными масштабными факторами a и a_c (эквивалентно, b и a_v). Нас интересуют малые отклонения между скалярами a и a_c и, следовательно, линейный порядок малости по q .

Эволюция поля q подчиняется уравнениям S-сектора космологических возмущений (4.5.16)–(4.5.18), которые мы перепишем в следующем виде:

$$\delta_\epsilon = \frac{\Delta\Phi}{\gamma H^2}, \quad \delta_p = \frac{\dot{q}}{H}, \quad \Phi = \frac{H}{a} \int (\gamma q - S) \frac{da}{H}. \quad (6.1.2)$$

Чтобы вывести отсюда ключевое динамическое уравнение для скаляра q , надо задать две связи на три полевых переменных δ_ϵ , δ_p и S , то есть иметь минимальную информацию о физическом состоянии материи.

Свойства материи при энергиях, не превышающих энергию электрослабого перехода, нам хорошо известны. Они описываются стандартной моделью элементарных частиц и стандартной космологической моделью. О материи в более ранний период времени мы знаем мало, не говоря уже о том, что нет прямой информации о свойствах праматерии в период Большого взрыва. К счастью, существует достаточно общая и простая концепция *идеальных* материальных сред или полей, которая хорошо работает в применении к квазифридмановской геометрии и широко используется при рассмотрении динамики очень ранней Вселенной.

Ниже мы рассмотрим свойства идеальных сред, а затем вернемся к эволюции космологических возмущений.

6.2. Адиабатические и изометрические возмущения

У идеальных сред отсутствует анизотропная часть давления:

$$s_\mu^\nu = 0. \quad (6.2.1)$$

В этом случае $p = p_{\text{tot}}$ и тензор энергии-импульса квазиабловского потока (4.4.5) паскалев по форме:

$$T_{\mu\nu} = (\epsilon + p) u_\mu u_\nu - p g_{\mu\nu}. \quad (6.2.2)$$

Мы уже описали векторную и тензорную моды возмущений для идеальных гравитирующих сред (см. (4.5.8)–(4.5.10)). Остановимся подробнее на скалярных материальных средах, для которых условие идеальности означает

$$S = 0. \quad (6.2.3)$$

Оставшиеся уравнения, необходимые для замыкания линейной системы (6.1.2), задают связи между сопутствующими возмущениями плотности и давления материи. Здесь возможны варианты многокомпонентных сред и разных способов задания их состояний.

Рассмотрим ряд примеров.

Идеальные скалярные системы охватывают широкий класс материальных сред: от многокомпонентных гидродинамических до полевых, включая их различные сочетания и взаимодействия. Исходными переменными слабовозмущенной N -компонентной среды являются потенциалы пекулярных скоростей составляющих ее компонент $v_m = v_m(t, \mathbf{x})$:

$$u_\mu^{(m)} = (1 + D, \nabla v_m). \quad (6.2.4)$$

Здесь и далее индексы m и l нумеруют компоненты среды: $m, l = 1, \dots, N$.

Определение средней скорости материи связано с переносом полной плотности энергии и локального импульса (аналог скорости центра масс, см. (4.4.1)):

$$u_\mu = f_m u_\mu^{(m)} = (1 + D, \nabla v), \quad (6.2.5)$$

где потенциал полной скорости материи

$$v = \sum_m f_m v_m, \quad (6.2.6)$$

а весовые доли компонент среды:

$$f_m = f_m(t) \equiv \frac{W_m}{W}, \quad \Sigma f_m = 1. \quad (6.2.7)$$

Здесь и далее верхние индексы компонент заключены в скобки, поскольку никаких правил поднятия или опускания для них не предусмотрено; для совпадающих верхнего и нижнего индексов компонент предполагается суммирование (в этом случае знак суммы опускаем).

Описание многокомпонентной среды в форме (6.2.5) предполагает существование парциальных энтальпий W_m , сумма которых, по определению, составляет полную плотность энталь-

пии материи $W = \varepsilon + p$. Понятия *парциальных* давлений или плотностей энергий не требуется. Отдельно эти функции нигде в уравнениях не фигурируют и их определение неоднозначно из-за нелинейного ¹⁾ взаимодействия между компонентами среды: в уравнения входят лишь *полные* p и ε . Вся информация о физическом состоянии материи содержится в нелинейной функции полного эффективного давления среды p .

По аналогии со скаляром q (4.5.11) введем калибровочно-инвариантные переменные для описания компонент среды:

$$q_m = H v_m + A, \quad (6.2.8)$$

$$q = \sum f_m q_m, \quad (6.2.9)$$

$$\hat{q}_m \equiv q_m - q = H \hat{v}_m, \quad (6.2.10)$$

где $\hat{v}_m \equiv v_m - v$. Из N скаляров \hat{q}_m независимыми являются только $N - 1$, так как наложена одна алгебраическая связь $\sum f_m \hat{q}_m = 0$.

Таким образом, N -компонентная среда наряду с *адиабатическим* скаляром кривизны q характеризуется дополнительно еще $N - 1$ *изометрическими* (или *изотермическими*) скалярами \hat{q}_m , которые описывают возмущения состава среды. В отличие от общего поля q , переменные \hat{q}_m , по определению, не связаны с возмущениями метрики и полной плотности, но взаимодействуют с q динамическим образом.

Малые возмущения полного давления и плотности энергии тоже можно разбить на парциальные амплитуды:

$$\delta p = f_m \bar{\delta}_p^{(m)}, \quad \delta \varepsilon = f_m \bar{\delta}_\varepsilon^{(m)}, \quad (6.2.11)$$

однако такие разбиения неоднозначны в силу нелинейности материальной системы, поскольку члены, описывающие взаимодействия, можно отнести к любой из взаимодействующих компонент среды. Для парциальных амплитуд (6.2.11) фиксированы лишь парные суммы:

$$\bar{\delta}_\varepsilon^{(m)} + \bar{\delta}_p^{(m)} = \frac{\delta W_m - \dot{W}_m v}{W_m}, \quad (6.2.12)$$

¹⁾ В нулевом и первом порядках имеем: $T_{\mu\nu} = W_m u_\mu^{(m)} u_\nu^{(m)} - p g_{\mu\nu} = W u_\mu u_\nu - p g_{\mu\nu}$. Нелинейность существенна в нулевом порядке теории возмущений. (Исключение представляют те компоненты среды, которые взаимодействуют с остальной материей только гравитационным образом, см. (4.4.7).) Малые переменные δp_m и $\delta \varepsilon_m$ всегда можно ввести, и они удовлетворяют правилам линейной суперпозиции, см. (6.2.11)–(6.2.14).

а в остальном выбор амплитуд диктуется состоянием материи и условиями задачи.

В каждой конкретной задаче можно построить собственные амплитуды компонент, не зависящие от их относительных движений.

Действительно, разбиения (6.2.11) очевидны в системе отсчета, сопутствующей общей скорости потока v . Амплитуды линейных возмущений пропорциональны как общим динамическим потенциалам потока q и \dot{q} , так и потенциалам относительных скоростей \hat{v}_m и ускорений \hat{v}_m компонент среды. Исключая из всех амплитуд слагаемые, пропорциональные самим полям \hat{v}_m , мы оставляем *собственные* возмущения давления и плотности:

$$\delta_p^{(m)} = \bar{\delta}_p^{(m)} - b^{(ml)} \hat{v}_l, \quad (6.2.13)$$

$$\delta_\varepsilon^{(m)} = \bar{\delta}_\varepsilon^{(m)} - c^{(ml)} \hat{v}_l. \quad (6.2.14)$$

В силу линейности это всегда можно сделать подбором соответствующих коэффициентов $b^{(ml)}$ и $c^{(ml)}$, являющихся функциями времени.

Оставшиеся амплитуды $\delta_p^{(m)}$ и $\delta_\varepsilon^{(m)}$ не зависят явно от вариации полей \hat{v}_l и, следовательно, характеризуют *внутреннее* состояние самих компонент. Для идеальных сред их можно связать алгебраически:

$$\delta_p^{(m)} = \beta_{ml}^2 \delta_\varepsilon^{(l)}, \quad (6.2.15)$$

с помощью матрицы звуковых скоростей $\beta_{ml}^2 = \beta_{ml}^2(t)$. Мы будем считать эту матрицу невырожденной, так что существует обратное преобразование:

$$\delta_\varepsilon^{(m)} = \beta_{ml}^{-2} \delta_p^{(l)}. \quad (6.2.16)$$

Соотношения (6.2.15) иногда связывают с гидродинамическим приближением, однако это не так: они охватывают широкий круг материальных сред и, в частности, полевые теории (см., напр., (6.4.1)). Правильнее классифицировать (6.2.15) как *второе* условие идеальности материи. Первое условие (6.2.3) явно нигде не использовалось при выводе (6.2.15). Оба они дополняют друг друга и достаточны для того, чтобы замкнуть систему уравнений S-моды (6.1.2).

Насколько уравнения (6.2.15) ограничивают свойства материи? Ответ на этот вопрос хорошо известен.

Вещество, состоящее из частиц, описывается с помощью кинетических уравнений. Условие (6.2.15) — это часто работающее приближение, в рамках которого эффекты неидеальности приводят к затуханию, диссипации, вязкости и другим процессам

(тогда надо учитывать $S \neq 0$). Для больших длин свободного пробега соотношения (6.2.3) и/или (6.2.15) нарушаются и надо пользоваться другими уравнениями.

Скалярные поля всегда удовлетворяют условиям идеальности (6.2.3) и (6.2.15) в квазифридмановской геометрии (см. раздел 6.4). Мы будем использовать эти условия далее при выводе динамических уравнений S-моды.

6.3. Уравнение движения поля q

Уравнения (6.2.3), (6.2.11) и (6.2.15) предоставляют необходимую информацию о свойствах материи и позволяют связать малые возмущения эффективного давления и полной плотности энергии материи:

$$\delta_p = \beta^2 \delta_\varepsilon + \widehat{\delta}_p, \quad (6.3.1)$$

где функция времени

$$\beta^2 = \sum_{m,l} f_m \beta_{ml}^2 \quad (6.3.2)$$

— это, по определению, «средний» (эффективный) квадрат скорости звука. Для него справедливы вспомогательные соотношения:

$$f_l \beta^2 + \bar{f}_l = \sum_m f_m \beta_{ml}^2, \quad \sum_m \bar{f}_m = 0. \quad (6.3.3)$$

Скаляр изометрического возмущения давления

$$\widehat{\delta}_p = \widehat{\delta}_p(t, \mathbf{x}) = \bar{f}_m \delta_\varepsilon^{(m)} + f_m \left(b^{(ml)} - \beta^2 c^{(ml)} \right) \widehat{v}_l \quad (6.3.4)$$

описывает ту часть полного давления среды, которая не связана с возмущением полной плотности энергии δ_ε .

Действительно, при $\delta_\varepsilon = 0$: $\delta_p = \widehat{\delta}_p$. С другой стороны, из определения адиабатической ветви возмущений¹⁾:

$$\delta_\varepsilon^{(m)} = \delta_\varepsilon, \quad \widehat{v}_m = 0 \quad \forall m, \quad (6.3.5)$$

следует, что $\widehat{\delta}_p = 0$. Полная динамическая связь $\widehat{\delta}_p$ с полевыми переменными \widehat{q}_m зависит от уравнения состояния материи (функции p).

¹⁾ Во избежание недоразумений отметим, что (6.3.5) не является решением динамических уравнений: адиабатическая и изометрические ветви взаимодействуют между собой в ходе эволюции. Формулы (6.3.5) определяют адиабатическую ветвь возмущений плотности в фиксированный момент времени, например в качестве начальных условий.

Вернемся к (6.1.2) и выведем динамическое уравнение для скаляра кривизны q .

Отсутствие анизотропного давления (6.2.3) позволяет непосредственно связать гравитационный потенциал Φ со скаляром q :

$$\Phi = \frac{H}{a} \int q \frac{\gamma da}{H}, \quad (6.3.6)$$

а из соотношений (6.1.2) и (6.3.1) следует:

$$\begin{aligned} \delta\varepsilon_c \equiv (\varepsilon + p) \delta\varepsilon &= \alpha^2 H \left(\dot{q} - H \widehat{\delta}_p \right), \\ \alpha^2 &= \frac{\varepsilon + p}{H^2 \beta^2} = \frac{\gamma}{4\pi G \beta^2} \end{aligned} \quad (6.3.7)$$

(индекс «S» в коэффициенте α_S по возможности опускаем). Подставляя эти формулы в релятивистское уравнение Пуассона (6.1.2), получаем

$$\alpha^2 a^3 \left(\dot{q} - H \widehat{\delta}_p \right) = \int \alpha^2 \beta^2 \Delta q \frac{da}{H}.$$

Дифференцирование этого соотношения дает нам искомое уравнение динамики скаляра кривизны:

$$\ddot{q} + \left(3H + 2 \frac{\dot{\alpha}}{\alpha} \right) \dot{q} - \beta^2 a^{-2} \Delta q = \frac{\left(\alpha^2 a^3 H \widehat{\delta}_p \right)'}{\alpha^2 a^3}. \quad (6.3.8)$$

В левой части мы имеем звуковой даламбертиан для скаляра q , описывающего общее адиабатическое возмущение плотности материи. Справа стоит комбинация кумулятивного изометрического скаляра $\widehat{\delta}_p$ и его производной, характеризующая динамическое воздействие изометрических возмущений плотности на эволюцию q .

Приведем еще одну удобную форму уравнения (6.3.8). Введем скаляр *изометрического возмущения плотности*

$$\widehat{\delta} \equiv -\frac{\widehat{\delta}_p}{\beta^2} = \delta_\varepsilon - \frac{q'}{\beta^2 H}. \quad (6.3.9)$$

С его помощью соотношение (6.3.1) можно переписать как

$$\delta_\varepsilon = \beta^{-2} \delta_p + \widehat{\delta}, \quad (6.3.10)$$

а уравнение движения адиабатического скаляра q в модели Фридмана приобретает следующий вид:

$$\frac{[\gamma a^2 (\beta^{-2} q' + \bar{H} \hat{\delta})]'}{\gamma a^2} - \Delta q = 0. \quad (6.3.11)$$

Здесь мы перешли к производным по конформному времени η , $\bar{\alpha} = \alpha a$, $\bar{H} = aH$. Заметим, что скаляры $\hat{\delta}$ и $\hat{\delta}_p$ совпадают с введенными ранее в разделе 5.4 (см. (5.4.9)) при условии выполнения фоновых уравнений движения (нулевого порядка) компонент материи.

Уравнения (6.3.8)–(6.3.11) справедливы для широкого класса сред. Информация о структуре материи нам не потребовалась. Единственная геометрическая характеристика среды, необходимая для вывода ключевого уравнения (6.3.8), — это β^2 — «средний» квадрат скорости распространения скалярных возмущений в среде (см. (6.3.2)).

Для вывода уравнений, описывающих эволюцию метрических скаляров \hat{q}_m с учетом влияния на них адиабатического скаляра q , следует конкретизировать состояние материи в части свойств и взаимодействий ее компонент. В качестве примера рассмотрим полевую идеальную среду.

6.4. Скалярные поля как идеальные среды

Один из способов задания уравнения состояния идеальной скалярной среды подсказывает теория поля.

В полевом подходе исходными переменными являются скалярные физические поля $\varphi_m = \varphi_m(t, \mathbf{x})$ с лагранжевой плотностью, зависящей от самих полей и их первых производных (знак материи «m» опускаем):

$$\mathcal{L} = \mathcal{L}(\varphi_1, \varphi_2, \dots, \varphi_N, w_1, w_2, \dots, w_N), \quad (6.4.1)$$

где

$$w_m^2 \equiv g^{\mu\nu} \varphi_{m,\mu} \varphi_{m,\nu} \quad (6.4.2)$$

— кинетические скаляры полей φ_m .

Тензор энергии-импульса материи получается прямым дифференцированием (6.4.1) по метрике:

$$T_{\mu\nu} \equiv \frac{2\partial(\sqrt{-g}\mathcal{L})}{\sqrt{-g}\partial g^{\mu\nu}} = \sum_m \frac{n_m}{w_m} \varphi_{m,\mu} \varphi_{m,\nu} - \mathcal{L} g_{\mu\nu}, \quad (6.4.3)$$

где

$$n_m \equiv \frac{\partial \mathcal{L}}{\partial w_m}. \quad (6.4.4)$$

Отсюда получаем выражение для следа:

$$T \equiv T_\mu^\mu = W - 4\mathcal{L}, \quad W = \sum_m n_m w_m. \quad (6.4.5)$$

Мы полагаем, что система N функций $n_l = n_l(w_m, \varphi_m)$ допускает обращение: $w_m = w_m(n_l, \varphi_l)$, при этом поля φ_l рассматриваются как параметры. Для такого обращения необходимо существование невырожденной матрицы преобразований (и обратной к ней):

$$\beta_{lm}^{-2} = \frac{\partial \ln |n_l|}{\partial \ln |w_m|}, \quad \beta_{ml}^2 = \frac{\partial \ln |w_m|}{\partial \ln |n_l|}. \quad (6.4.6)$$

Эти матрицы можно получить из исходной *симметричной* матрицы B_{lm} (и обратной к ней $B^{(lm)} B_{mn} = \delta_n^l$) с помощью прямого дифференцирования (6.4.1) по кинетическим скалярам:

$$B_{lm} = B_{ml} \equiv \frac{W \partial^2 \mathcal{L}}{n_l n_m \partial w_l \partial w_m}. \quad (6.4.7)$$

Требуемые соотношения имеют вид:

$$\beta_{lm}^{-2} = f_m B_{lm}, \quad f_m \beta_{ml}^2 = B^{(ml)}, \quad (6.4.8)$$

где $f_m = n_m w_m / W$. Методологически нет необходимости называть (6.4.7) квадратом звуковой матрицы β_{lm}^2 , поскольку положительная определенность нигде не требуется и не используется (нам не приходится извлекать квадратный корень из β_{lm}^2). В данном случае мы сохраняем традиционную форму записи, прибегая к понятию *скорости звука* при интерпретации компонент исходной матрицы B_{lm} .

Докажем, что тензор (6.4.3) описывает идеальные среды в квазифридмановской геометрии.

Прежде всего, введем касательные единичные векторы к мировым линиям движения полей φ_m :

$$u_\mu^{(m)} \equiv \frac{\varphi_{m, \mu}}{w_m}, \quad (6.4.9)$$

и общую скорость среды как линейную суперпозицию импульсов полей:

$$u_\mu = f_m u_\mu^{(m)}. \quad (6.4.10)$$

Считая относительные отклонения скоростей малыми, получаем тензор (6.4.3) в форме идеальной среды (6.2.2)–(6.2.5), где

$$p \equiv \mathcal{L}, \quad \varepsilon \equiv W - \mathcal{L}, \quad (6.4.11)$$

$$v = \sum_m f_m v_m, \quad v_m \equiv \frac{\delta \varphi_m}{\dot{\varphi}_m} \quad (6.4.12)$$

(напомним, что в нулевом порядке $w_m = \dot{\varphi}_m$).

Обратимся к амплитудам компонент. Из представлений (6.2.11) и (6.4.1) имеем:

$$\bar{\delta}_p^{(m)} = \delta_w^{(m)} + \frac{p, m}{n_m} \hat{v}_m, \quad (6.4.13)$$

$$\begin{aligned} \bar{\delta}_\varepsilon^{(m)} &= \delta_n^{(m)} - \frac{p, m}{n_m} \hat{v}_m = \\ &= \beta_{ml}^{-2} \delta_w^{(l)} - \frac{p, m}{n_m} \hat{v}_m + \frac{1}{n_m} \sum_l n_{m, l} \omega_l \hat{v}_l, \end{aligned} \quad (6.4.14)$$

где калибровочно-инвариантные переменные

$$\delta_w^{(m)} \equiv \frac{\delta w_m - \dot{w}_m v}{w_m} \quad \text{и} \quad \delta_n^{(m)} \equiv \frac{\delta n_m - \dot{n}_m v}{n_m}$$

соответствуют возмущениям парциальных скаляров (6.4.2) и (6.4.4) относительно общего потока v . Отсюда убеждаемся в тождественности матриц (6.4.6) и звуковых матриц идеальной среды (6.2.16) при выборе

$$\delta_\varepsilon^{(m)} \equiv \beta_{ml}^{-2} \delta_w^{(l)}, \quad \delta_p^{(m)} \equiv \delta_w^{(m)}, \quad (6.4.15)$$

а также получаем коэффициенты (6.2.13):

$$b^{(ml)} = \frac{p, m}{n_m} \delta_{ml}, \quad c^{(ml)} = \frac{n_{m, l} \omega_l - p, l \delta_{ml}}{n_m},$$

где δ_{ml} — символ Кронекера, $p, m \equiv \partial p / \partial \varphi_m$.

Для $c^{(ml)}$ выполняются правила суммирования:

$$f_m c^{(ml)} = \frac{\varepsilon, l \omega_l}{\varepsilon + p}.$$

Подставляя это выражение в (6.3.4), находим скаляр изометрического давления, воздействующего на эволюцию скаляра кривизны:

$$\hat{\delta}_p = \sum_m \left(\bar{f}_m \beta_{ml}^{-2} \delta_w^{(l)} + \frac{p, m - \beta^2 \varepsilon, m}{\varepsilon + p} w_m \hat{v}_m \right). \quad (6.4.16)$$

Таким образом, мы доказали, что поля φ_m с общим нелинейным лагранжианом (6.4.1) относятся к рассмотренному нами ранее более общему случаю идеальных сред (см. разделы 6.2 и 6.3). Форма (6.4.1) — это лишь один из способов задания уравнения состояния материи. Здесь мы имеем исчерпывающую информацию о среде.

Варьируя по φ_m действие (1.9.1) с лагранжевой плотностью $\mathcal{L}_m = \mathcal{L}$, получаем уравнения эволюции для изометрических скаляров \hat{q}_m :

$$(n_m u_m^\nu)_{;\nu} = p, m, \quad (6.4.17)$$

которые вместе с уже известным нам уравнением (6.3.8) для скаляра кривизны q составляют полную динамическую систему уравнений S-моды возмущений квази Хаббловского потока многокомпонентной скалярной среды. Полное действие такой системы получается путем разложения (1.9.1) до второго порядка по q_m с исключением высших производных (см. (1.9.4)).

Предоставив читателю инструмент для самостоятельных исследований, мы не пойдём по пути дальнейших теоретических обобщений и усложнений. Важнейшие физические и геометрические свойства ранней Вселенной, получившие экспериментальное подтверждение наблюдательной космологии, проявляются на уровне простейших гравитирующих систем, которым посвящена эта книга.

В качестве примеров рассмотрим динамику S-моды в многокомпонентных материальных средах, возникающих после появления в ранней Вселенной темной материи.

6.5. Гравитационная неустойчивость темной материи

После образования в ранней Вселенной бесстолкновительной нерелятивистской ТМ она взаимодействует с радиационно-барионной плазмой ($R = r + b$) только гравитационным образом. Влиянием темной энергии в эту эпоху ($z \gg 1$) можно пренебречь.

Из уравнения Фридмана получаем выражения для масштабного фактора и его производных при $z \gg 1$:

$$a = a_1 y, \quad t = t_1 \bar{\eta}^2 (3 + 2\bar{\eta}), \quad (6.5.1)$$

$$\bar{H} = k_c \frac{\sqrt{1 + 4y}}{y}, \quad \gamma = \frac{2(1 + 3y)}{1 + 4y}, \quad (6.5.2)$$

$$y \equiv \bar{\eta} (1 + \bar{\eta}), \quad \bar{\eta} \equiv \frac{\eta}{\eta_c},$$

где $a_1 = 4\Omega_r/\Omega_m \simeq 1,25 \cdot 10^{-3}$, $t_1 = a_1\eta_c/6 \simeq 4 \cdot 10^5$ лет, $H_0\eta_c = 4\sqrt{\Omega_r}/\Omega_m \simeq 0,13$, $k_c \equiv \eta_c^{-1} \simeq 1,8 \cdot 10^{-3}$ Мпк $^{-1}$, значения космологических параметров стандартной модели даны в списке обозначений.

Для весовых долей компонент среды (6.2.7) имеем:

$$f_r = \frac{1}{1+3y}, \quad f_m = \frac{3y}{1+3y}, \quad (6.5.3)$$

$$f_M = f_M f_m, \quad f_b = f_b f_m, \quad f_R = f_r + f_b,$$

где численные факторы $f_M = \Omega_M/\Omega_m \simeq 5/6$ и $f_b = \Omega_b/\Omega_m \simeq 1/6$ соответствуют массовым долям ТМ и барионов в нерелятивистской материи ($m = M + b$).

Из (6.5.1) следует соотношение между переменной y и космологическим красным смещением $z \gg 1$:

$$y \simeq \frac{800}{1+z}. \quad (6.5.4)$$

В момент равенства плотностей радиации и нерелятивистской материи $y_{eq} = 1/4$ и хаббловская частота равна

$$k_{eq} = \bar{H}_{eq} = 4\sqrt{2} k_c \simeq 10^{-2} \text{ Мпк}^{-1}. \quad (6.5.5)$$

Космологическая рекомбинация водорода ($z_{rec} \simeq 1100$) происходит несколько позже момента равенства плотностей ($y_{rec} \simeq 3/4$), но раньше смены темпа космологического расширения $\eta_c \simeq 2\eta_{rec}$. Определяя масштаб звукового горизонта на рекомбинации (см. (6.5.8)), получаем *удивительное совпадение* разных физических масштабов, которое выполняется в нашей Вселенной:

$$k_{rec} \equiv \frac{\bar{H}}{\beta_R} \Big|_{rec} \simeq k_{eq}. \quad (6.5.6)$$

Если соотношение между плотностями барионов и ТМ сохраняется в ходе времени ($\varepsilon_b/\varepsilon_M \simeq 1/5$), то вклад барионов в плотность энергии радиационно-доминированной плазмы растет и достигает одной трети к моменту рекомбинации, заданному только температурой излучения ($\varepsilon_b/\varepsilon_R \simeq 1/3$ при $\eta = \eta_{rec}$).

Эволюцию космологических возмущений дорекомбинационной материи можно исследовать в приближении двух идеальных сред — плазмы и ТМ. R-компонента обладает давлением и описывается уравнением состояния:

$$\mathcal{L}_R \equiv p = p(w) = \frac{\varepsilon_r}{3}, \quad \varepsilon_R = wp, \quad w - p = \varepsilon_r + \varepsilon_b. \quad (6.5.7)$$

М-компоненту считаем холодной пылевой средой с нулевым давлением: $p_M = 0$ ¹⁾. Определим квадрат скорости звука в плазме и среднюю звуковую скорость согласно (6.3.2):

$$\beta_R^2 = \frac{\dot{p}}{\dot{\epsilon}_R} = \frac{1}{3(1 + 3f_b y)} = \frac{f_r}{3f_R}, \quad (6.5.8)$$

$$\beta^2 = f_R \beta_R^2 = \frac{f_r}{3} = \frac{1}{3(1 + 3y)}. \quad (6.5.9)$$

Путем преобразования уравнений (6.4.17) для потенциала скорости плазмы и плотности ТМ получаем динамическую систему из двух уравнений для адиабатического скаляра q (6.3.8) и скаляра изометрического давления $\hat{\delta}_p$ (6.3.1):

$$H \left(a f_R \bar{\delta}' \right)' = -\Delta q', \quad (6.5.10)$$

где

$$\bar{\delta} \equiv -\frac{\hat{\delta}_p}{f_M y f_r^2} \equiv \frac{\hat{\delta}}{f_M} = \delta_M - \delta_R + \frac{3\bar{q}}{f_r}, \quad (6.5.11)$$

$$\bar{q} \equiv H \int_0^{\eta} \frac{q'}{H} d\eta,$$

δ_l — сопутствующие возмущения плотности компонент среды $l = R, M, b, \dots$. Очевидно, для плазмы выполняются равенства:

$$\delta_R = \delta_r = \delta_b, \quad v_R = v_r = v_b, \quad (6.5.12)$$

и возмущение общей плотности материи равно (см. (6.3.9)):

$$\delta_\epsilon = f_R \delta_R + f_M \delta_M = \frac{3\dot{q}}{f_r H} + f_M \bar{\delta}. \quad (6.5.13)$$

Обращая формулы (6.5.13) и (6.5.11), выводим соотношения для возмущений компонент материи в виде функций адиабатического и изометрического скаляров:

$$\delta_R = \frac{3f_R}{f_r} \left(\frac{v}{f_R} \right)', \quad (6.5.14)$$

$$\delta_M = 3y \left(\frac{v}{y f_r} \right)' + \bar{\delta}, \quad (6.5.15)$$

$$v = \frac{\bar{q}}{H} \equiv \int_0^{\eta} \frac{q'}{H} d\eta = \frac{q - A}{H}, \quad A = H \int_0^{\eta} \gamma q dt. \quad (6.5.16)$$

¹⁾ Рост неоднородностей ТМ обусловлен именно пренебрежимо малым давлением М-частиц, в то время как возмущения релятивистской материи не нарастают из-за эффективно высокого давления (см. (6.5.31)). По этой причине влиянием релятивистских бесстолкновительных нейтрино на эволюцию возмущений ТМ можно пренебречь. Однако мы учитываем вклад нейтрино в фоновую плотность г-компоненты.

Уравнение (6.3.8) упрощается при трансформации¹⁾ к конформной переменной $\tilde{q} = \mathcal{K}q$:

$$\tilde{q}'' + (\omega^2 - U) \tilde{q} = -\frac{\omega_r}{3} f_M H \left(\frac{\bar{\delta}}{H} \right)', \quad (6.5.17)$$

$$\mathcal{K} = \frac{\omega_r}{f_r \bar{H}} = \frac{\alpha \omega_r}{2\sqrt{\rho_c} \Omega_r}, \quad U = \frac{4k_c^2(1+6y)}{y(1+4y)}. \quad (6.5.18)$$

Интегрируя (6.5.10) и исключая $\bar{\delta}'$ из правой части (6.5.17), получаем следующую систему, пригодную для решения методом итераций:

$$\omega_R^{-2} \tilde{q}'' + (1 - \omega_R^{-2} U) \tilde{q} = f_M \left(3y f_r \tilde{A} - \frac{2f_R \nu \bar{\delta}}{3f_r(1+2\bar{\eta})} \right), \quad (6.5.19)$$

$$\bar{\delta}' = \frac{f_r}{f_R} \sqrt{3} k (\tilde{q} - \tilde{A}), \quad \tilde{A} \equiv \mathcal{K}A = \frac{2}{y f_r} \int_0^{\bar{\eta}} \frac{\tilde{q} d\bar{\eta}}{1+2\bar{\eta}}. \quad (6.5.20)$$

Обратимся к галактическим масштабам:

$$\nu \equiv \frac{3\sqrt{3} k_c}{k} \simeq \frac{k_{\text{eq}}}{k} \ll 1. \quad (6.5.21)$$

Для них потенциал U не существен, поскольку эти масштабы попадают внутрь хаббловского радиуса на радиационно-доминированной стадии $\bar{\eta} \ll 1$, когда фактор \sqrt{U} оказывается меньше частоты ($\sqrt{U} \ll \bar{H} \sim k$) и далее остается малым ($\sqrt{U} \ll \omega$). По этой же причине правая часть (6.5.19) подавлена малым фактором ν , что позволяет представить решение в виде ряда по степеням ν .

Предполагая начальные условия для растущей адиабатической ветви возмущений в виде:

$$\bar{H} \gg k: \quad q = q(\mathbf{x}), \quad \delta_l = 0, \quad (6.5.22)$$

получаем главное решение уравнения (6.5.19) в адиабатическом приближении:

$$\tilde{q} = \left(\frac{f_R}{f_r} \right)^{1/4} q_k \sin \kappa \simeq q_k \sin \kappa, \quad (6.5.23)$$

¹⁾ Мы перешли к фурье-переменным, индекс k у них опущен. По определению, $\omega = \beta k$, $\omega_R = \beta_R k$, $\omega_r = k/\sqrt{3}$. В этом и следующем разделах v и A — вспомогательные функции, определяемые равенствами (6.5.16). Они совпадают с потенциалами (4.5.1) при $C = D = v_M = 0$ (тогда $v = f_R v_R$).

где $q_k = q_k(\bar{H} \gg k)$ — спектральная амплитуда $q(\mathbf{x})$, фазовая функция определяется выражением

$$\kappa \equiv \int_0^{\eta} \omega_R d\eta \simeq \frac{k\eta}{\sqrt{3}},$$

отличием между f_R и f_r на стадии $y < 1$ можно пренебречь в силу малости f_b . Из (6.5.23) и (6.5.20) получаем лидирующие члены спектральных гармоник при $\bar{\eta} < 1/2$:

$$q \simeq q_k \frac{\sin \kappa}{\kappa}, \quad (6.5.24)$$

$$q + \frac{1}{3} \delta_b \simeq q_k \cos \kappa, \quad (6.5.25)$$

$$\frac{1}{3} \delta_b \simeq -\frac{\kappa^2}{2} \Phi \simeq q_k \left(\cos \kappa - \frac{\sin \kappa}{\kappa} \right), \quad (6.5.26)$$

$$\tilde{A} = \tilde{q} - \frac{k\nu}{\sqrt{3} a f_r} \simeq 2q_k \frac{1 - \cos \kappa}{\kappa}, \quad (6.5.27)$$

$$\bar{\delta} \simeq 3q_k (1 - j_1 - \cos \kappa), \quad (6.5.28)$$

где функции $j_{1,2}$ зависят от фазовой переменной κ :

$$j_1(\kappa) = 2 \int_0^{\kappa} \frac{1 - \cos x}{x} dx, \quad j_2(\kappa) = 2 \int_0^{\kappa} \frac{\sin x}{x} dx.$$

Подставляя (6.5.27) и (6.5.28) в правую часть уравнения (6.5.19), получаем решение для \tilde{q} после двух итераций:

$$\tilde{q} = q_k (\sin \kappa + 2\nu f_M Q), \quad (6.5.29)$$

где Q и $\bar{j}_{1,2}$ — функции фазовой переменной κ :

$$\begin{aligned} Q &= j_1 + \bar{j}_1 \cos \kappa - \bar{j}_2 \sin \kappa, \\ \bar{j}_n &= j_n(\kappa) - \frac{1}{2} j_n(2\kappa), \quad n = 1, 2. \end{aligned} \quad (6.5.30)$$

Мы видим, что $Q \simeq (2 + \cos \kappa) \ln \kappa$ при $1 < \kappa < \nu^{-1}$. Поправка $\propto \nu$ в (6.5.29) содержит плавно растущую функцию времени

$\sim \ln(k\eta/\sqrt{3})$ и остается меньше лидирующего осциллирующего члена в области своего определения ¹⁾).

Экстраполяция полученных решений в причинно-связанную область дает асимптотическое поведение возмущений плотности компонент материи в дорекомбинационной Вселенной (см. (6.5.14), (6.5.15)):

$$\delta_b \simeq 3q_k \cos \kappa, \quad \delta_M \simeq -6q_k \ln \kappa. \quad (6.5.31)$$

Мы видим, что звуковые волны в плазме распространяются с *постоянной* амплитудой, в то время как возмущения плотности ТМ *логарифмически нарастают*.

К моменту рекомбинации водорода $\bar{\eta} \simeq 1/2$ суммарный рост коротковолновых возмущений плотности ТМ по отношению к начальным возмущениям кривизны оценивается следующим фактором:

$$g_M \equiv -\frac{\delta_{M\text{rec}}}{q_k} \simeq 6 \ln \kappa_{\text{rec}}, \quad \kappa_{\text{rec}} \simeq \beta_R k \eta_{\text{rec}} \simeq \frac{3k}{2k_{\text{eq}}}. \quad (6.5.32)$$

¹⁾ При $\bar{\eta} \ll 1$ выводим уравнение для функции Q :

$$3k^{-2}Q'' + Q = j_1,$$

откуда при условии (6.5.22) следует (6.5.30). Для производных Q имеем:

$$\frac{dQ}{d\kappa} = -\bar{j}_1 \sin \kappa - \bar{j}_2 \cos \kappa, \quad \frac{d^2Q}{d\kappa^2} = -\bar{j}_1 \cos \kappa + \bar{j}_2 \sin \kappa,$$

что можно использовать для проверки решения. При $\kappa \ll 1$: $j_1 = -\bar{j}_1 = \kappa^2/2$, $\bar{j}_2 = \kappa^3/3$, $Q = -\kappa^4/12$. При $\kappa \gg 1$: $j_1/2 = \bar{j}_1 + \ln 2 = \ln \kappa + C$, $j_2 = 2\bar{j}_2 = \pi$, $Q \simeq (2 + \cos \kappa) \ln \kappa$, где $C = 0,577$ — константа Эйлера.

Для корректной экстраполяции к моменту рекомбинации необходимо учитывать точный вид функции $\mathcal{K} = \omega_r \eta (1 + \bar{\eta})(1 + 2\bar{\eta})\gamma/2$, при этом разницей в амплитудах между f_R и f_r можно пренебречь. В коротковолновом пределе адиабатическое приближение справедливо в области, включающей в себя окрестность $\bar{\eta} \sim 1/2$. В главном порядке по $\nu \ll 1$ имеем:

$$q = q_k \frac{\sin \kappa}{\mathcal{K}}, \quad \tilde{A} = q_k \frac{2\nu}{3y f_r} I_s, \quad \frac{\bar{\delta}}{3q_k} = 1 - J - \cos \kappa,$$

$$I_s = \int_0^{\kappa} \frac{\sin \kappa \, d\kappa}{1 + 2\bar{\eta}}, \quad I_c = 2 \int_0^{\kappa} \frac{(1 - \cos \kappa) d\bar{\eta}}{1 + 4y}, \quad J = 2 \int_0^{\kappa} \frac{I_s d\bar{\eta}}{y f_r}.$$

Подставляя эти решения в правые части (6.5.19) и (6.5.15), получаем уравнение для следующей итерации (6.5.29) и фактор роста ТМ (см. (6.5.32)):

$$\omega_R^{-2} Q'' + Q = (1 + 2\bar{\eta})^{-1} J + I_c, \quad g_M = 6 \int_0^{\kappa} \frac{I_s d\bar{\eta}}{y f_r} - 3.$$

При вычислении этого интеграла можно положить $\kappa = k\eta/\sqrt{3}$, численное значение g_M соответствует оценке (6.5.32).

Численный коэффициент «6» в (6.5.31) и (6.5.32) возникает из комбинации параметров γ/β^2 на радиационно-доминированной стадии. Более точный расчет с учетом изменения функции \mathcal{K} в области рекомбинации подтверждает эти выводы (см. сноску на с. 193).

Эволюционное поведение возмущений компонент материи (6.5.31) в дорекомбинационной Вселенной знаменательно и требует дополнительных комментариев.

6.6. Звуковые колебания барионов

Темная материя гравитационно неустойчива в силу незначительности давления. Логарифмическое возрастание контраста плотности M -компоненты сменяется степенным ростом возмущений нерелятивистской материи ($m = M + b$) на материально-доминированном этапе эволюции:

$$\delta_\varepsilon \simeq \delta_m = -y g_k q_k \quad \text{при } y > 1. \quad (6.6.1)$$

Знак «—» означает ¹⁾, что плотность нарастает в областях пониженного потенциала кривизны; g_k — переходная функция, связывающая пространственные гармоники начального ($y \ll 1$) потенциала q и конечной ($y > 1$) плотности материи $\delta_m = f_M \delta_M + f_b \delta_b$.

Когда температура радиационно-барионной плазмы падает до 3000 К, происходит рекомбинация водорода. В результате число свободных электронов в сопутствующем объеме уменьшается, и длина свободного пробега фотонов возрастает. Излучение уже нельзя считать идеальной средой, и его описание требует кинетического подхода. Весь процесс рекомбинации протекает за короткое время (несколько десятых хаббловского времени), после чего Вселенная оказывается прозрачной для реликтовых квантов — длина их пробега становится больше размера горизонта. Барионы отсоединяются от излучения и начинают *свободное* падение в неоднородном гравитационном поле, усиливая свой контраст плотности и сближая его с профилем распределения ТМ.

Следует напомнить, что к началу рекомбинации плотность нерелятивистской материи втрое превышает плотность излуче-

¹⁾ Если речь идет о спектрах, то знак «—» следует опустить. Пренебрегая в (6.6.1) излучением, не следует забывать о нейтрино, которые присоединяются к m после того, как стали нерелятивистскими ($y > 10$). Однако их доля в материи мала ($f_\nu < 0,07$, см. (10.8.4)), и их необходимо будет учитывать при увеличении точности определения параметров (см. табл. 10.1).

ния ($f_m^{(\text{rec})} \simeq 9/13$, $f_r^{(\text{rec})} \simeq 4/13$), и это соотношение возрастает пропорционально масштабному фактору (см. (6.5.3)). Поэтому динамика развития возмущений в послерекомбинационной Вселенной управляется нерелятивистским веществом, которое можно описывать двумя идеальными компонентами M и b , взаимодействующими друг с другом гравитационным образом.

Среди многих мод решений имеется *единственное* растущее решение ¹⁾, лидирующее при больших y (см. (5.5.3), (6.3.9)):

$$q = q_0 \left(1 + O\left(\frac{1}{y}\right) \right), \quad (6.6.2)$$

$$\hat{\delta} \simeq y \frac{\Delta q_0}{10k_c^2}, \quad v \propto a_1 \eta q_0. \quad (6.6.3)$$

Функция пространственных координат $q_0 = q_0(\mathbf{x})$ содержит всю информацию о линейных возмущениях плотности, из которых образуется структура во Вселенной. Соответствующие контрасты плотности компонент материи ($l = m, b, M$) ведут себя следующим образом при $y > 1$:

$$\delta_l \rightarrow O(q_0), \quad (\delta_l - \hat{\delta}) \rightarrow \text{const}_l = O(q_0).$$

Переходя от k_c к стандартному масштабу k_{eq} (см. (6.5.5)), получаем из (6.6.3) интересующую нас связь между спектральными амплитудами δ_m и q_0 (ср. (6.6.1)):

$$\delta_m = -y \frac{3,2 \cdot k^2}{k_{\text{eq}}^2} q_0(k). \quad (6.6.4)$$

Как видно, обе нерелятивистские компоненты среды падают в общем гравитационном поле и их возмущения плотности нарастают пропорционально масштабному фактору. Однако *разность* парциальных контрастов плотности b и M , оставшаяся из-за различной истории их эволюции в радиационно-доминированную эпоху, сохраняется неизменной во времени. Такую динамику можно сравнить с падением двух материальных точек $l = 1, 2$

¹⁾ Интегрируя уравнение (6.3.11) с учетом (6.6.2), получаем первую итерацию для функции $\hat{\delta}$:

$$\hat{\delta} = \Delta q_0 (\gamma y^2 \bar{H})^{-1} \int \gamma y^2 d\eta.$$

Подставляя сюда из (6.5.2) функции γ и y и используя тождество

$$30(1 + 2\bar{\eta}) \int \gamma y^2 d\bar{\eta} = 18y^3 + \bar{\eta}^3(2 + \bar{\eta}),$$

получаем, переходя к $y > 1$, первую формулу (6.6.3).

с Пизанской башни, отпущенных одновременно в момент времени $\eta = 0$ из разных начальных положений x_{0l} : $x_l = x_{0l} + g\eta^2/2$, где g — гравитационное ускорение. При этом линейный член $V_l\eta$, пропорциональный начальным скоростям компонент, в космологии отсутствует из-за конечности пробега нерелятивистской частицы в материально-доминированной Вселенной: $\int V_l d\eta \propto \int d\eta/a \rightarrow \text{const}$.

Пользуясь линейностью уравнений движения, мы можем связать послерекомбинационную $q_0(k)$ и начальную q_k спектральные амплитуды S-моды возмущений:

$$q_0(k) = T(k) \cdot q_k, \quad (6.6.5)$$

где $T(k)$ — переходная функция, трансформирующая начальный S-спектр в послерекомбинационный спектр неоднородностей кривизны, непосредственно предшествующий эпохе образования галактик. Обращаясь к (6.6.1), получаем соотношение между переходными функциями плотности и кривизны:

$$g_k = \frac{3,2 \cdot k^2}{k_{\text{eq}}^2} T(k). \quad (6.6.6)$$

Асимптотики переходной функции можно найти аналитически. По построению (6.6.5) $T(k \ll k_{\text{eq}}) = 1$. Для малых длин волн из (6.6.1) и (6.5.32) в главном порядке получаем:

$$g_k \sim g_m \equiv \frac{4}{3} f_M g_M \simeq 6,6 \cdot \ln \kappa_{\text{rec}}, \quad (6.6.7)$$

$$T(k \gg k_{\text{eq}}) \sim \frac{2k_{\text{eq}}^2}{k^2} \ln\left(\frac{3k}{2k_{\text{eq}}}\right). \quad (6.6.8)$$

Звуковые колебания барионов с фазой на момент рекомбинации «впечатываются» в переходные функции (6.6.7) и (6.6.8) в виде модуляционной аддитивной добавки к логарифму $\sim 0,5 f_b/f_M \cos \kappa_{\text{rec}} \simeq 0,1 \cos \kappa_{\text{rec}}$ (см. (6.5.31)). Однако такая прямая оценка не вполне корректна не только потому, что при сшивке осциллирующих членов нужно учитывать все ветви решений, что скорректирует фазу и амплитуду модуляционной добавки. Дело в том, что малая ширина рекомбинации нарушает адиабатическое приближение для звуковых колебаний. Более короткие волны диссипируют в процессе рекомбинации из-за конечной длины свободного пробега фотонов (см. раздел 11.6), что приводит к сглаживанию спектральных модуляций.

Таким образом, для вычисления модуляционной компоненты спектра требуется численное моделирование. Напротив, на глад-

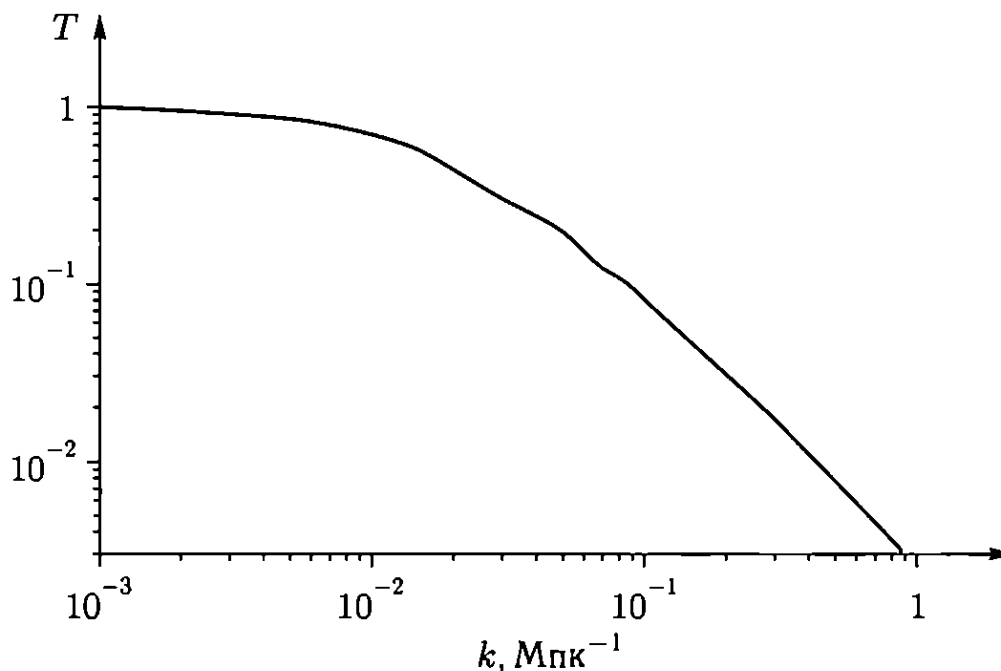


Рис. 6.1. Переходная функция $T(k)$ стандартной модели Вселенной

кую часть спектра (6.6.7) – (6.6.8) процесс рекомбинации влияния не оказывает. Функция $T(k)$ для КСМ представлена на рис. 6.1.

Удивителен тот факт, что весовое доминирование гравитационно стабильного излучения над плотностью ТМ при $\bar{\eta} \ll 0,2$ не препятствует росту неоднородностей ТМ. Напомним, что гравитационное влияние темной энергии полностью заблокирует развитие неустойчивости и гравитационно стабилизирует ТМ в будущем (см. раздел 12.1). Без преувеличения можно сказать, что и существованием, и грандиозным многообразием наблюдаемых галактик мы обязаны именно логарифмическому фактору роста g_m . Если бы не этот дополнительный коэффициент, то при заданных начальных возмущениях и составе материи, сгенерированных в ранней Вселенной, контраст плотности вещества просто не успел бы дорасти до нелинейных значений $\delta_m \sim 1$ и образовать галактики. Если бы ТМ, будучи нерелятивистской, взаимодействовала в той или иной степени с излучением или сама с собой, то логарифмический рост был бы подавлен или вовсе отсутствовал (как, например, для барионов) и в нашем мире не было бы галактик.

Принимая фактор роста плотности в послерекомбинационный период эволюции равным $\sim 10^3$, а на ранней стадии

$$g_m \simeq 6,6 \cdot \ln\left(\frac{0,2}{k_{\text{eq}}}\right) \simeq 20 \quad \text{для } k \sim 0,15 \text{ Мпк}^{-1},$$

получаем суммарный рост возмущений плотности в нерелятивистской материи $\sim 2 \cdot 10^4$ для масштабов скоплений галактик. Отсюда выводим спектральную амплитуду начальных скалярных

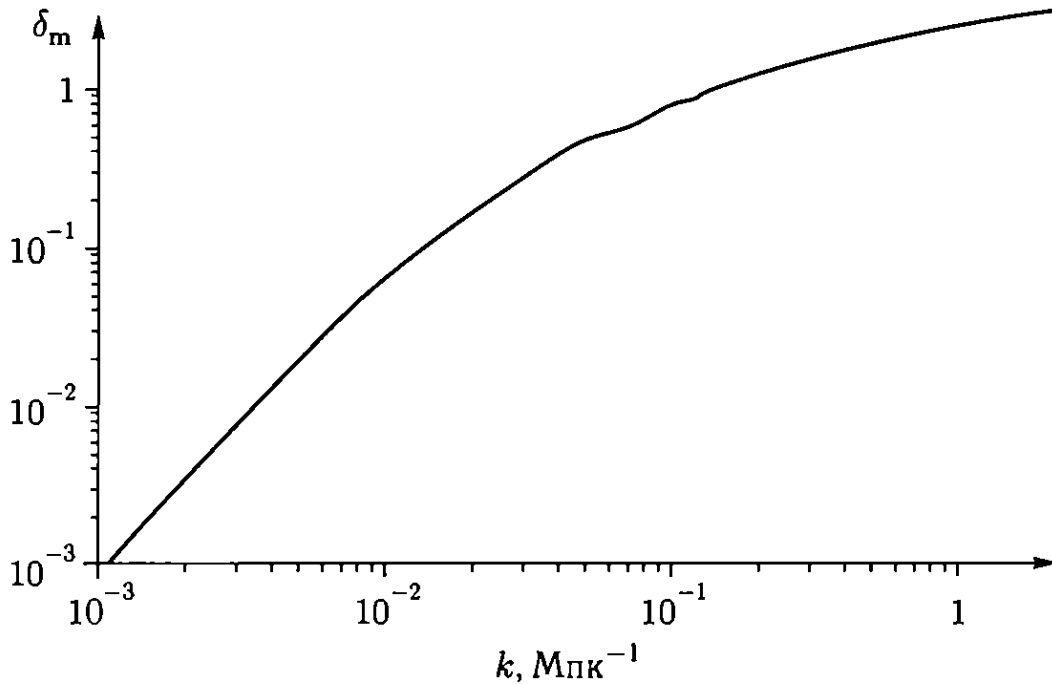


Рис. 6.2. Послерекombинационный спектр возмущений плотности $\delta_m(k)$ неоднородностей, необходимых для образования галактик:

$$q_k \simeq \frac{10^{-3}}{20} = 5 \cdot 10^{-5}. \quad (6.6.9)$$

Эта оценка опирается на астрономические данные о крупномасштабной структуре Вселенной и справедливость ОТО вплоть до минут после Большого взрыва, когда внутрь хаббловского горизонта попадают размеры порядка килопарсека.

Столь ясные и простые выводы основаны на двухкомпонентном приближении эволюции возмущений плотности в ранней Вселенной. Они подтверждаются точными численными расчетами с учетом нейтрино, барионов, а также кинетических эффектов излучения. Рисунок 6.2 демонстрирует спектральную амплитуду линейного возмущения плотности материи $\delta_m = \delta_m(k) = k^3 P(k)$, аппроксимированного к сегодняшнему моменту времени КСМ (ср. (3.4.5), (3.5.3)). Начальная величина S-спектра $q_k = 5 \cdot 10^{-5}$ соответствует спектральной амплитуде возмущений плотности $\delta_m \simeq 1$ при $k_1 \simeq 0,15 \text{ Мпк}^{-1}$, что эквивалентно масштабу неоднородности 10 Мпк ($\sigma_{10} \simeq 1$, см. раздел 3.5). Явно наблюдается логарифмический подъем коротковолновой части спектра, обусловленный фактором g_m , и звуковая модуляция в области $k \sim 0,1 \text{ Мпк}^{-1}$ (см. также рис. 6.3).

Не менее замечателен факт гравитационной стабильности излучения. Можно сказать, что она имеет критический характер. Почему звуковые волны распространяются в радиационно-доминированной плазме с постоянной амплитудой, а не затухают или не растут во времени?

Мы знаем, что свободные колебания пробных полей в расширяющейся Вселенной адиабатически затухают (например, гравитационные волны). Однако звуковые колебания «чувствуют» давление среды, в которой они распространяются. Здесь мы сталкиваемся с эффектом релятивистской неодновременности (см. раздел 1.4).

Действительно, воспользуемся лоренцевыми преобразованиями и перейдем в систему отсчета, сопутствующую звуковой скорости волны. Из-за того что в процессе расширения давление уменьшается ($\dot{p} < 0$), волна в своей собственной системе отсчета находится на падающем профиле давления и «подталкивается» (усиливается) градиентом давления в направлении распространения. Этот эффект можно сравнить с ускорением звуковой волны и превращением ее в ударную при выходе в верхние слои атмосферы.

Для уравнения состояния $p = \varepsilon/3$ увеличение амплитуды волны из-за эффекта релятивистской неодновременности точно компенсируется ее адиабатическим затуханием в расширяющейся Вселенной, в результате чего амплитуда звуковых колебаний остается неизменной. Показательно, что в главном порядке по $k \gg k_{\text{eq}}$ (см. (6.5.31)) плазма не прогибается и не смещается, хотя бы и с постоянной амплитудой, в гравитационные ямы потенциала $q(\mathbf{x})$ (куда устремляется ТМ, подобно тому как дождевая вода стекает в лужи). Это объясняется релятивистской динамикой: скорость звуковых волн близка к скорости света и заведомо превышает «вторую космическую скорость» гравитационных ям. Фононы, подобно свету, пронизывают однородную среду, свободно проходя и минуя области пониженного или повышенного гравитационного потенциала. Единственная реакция со стороны плазмы на возмущения гравитационного поля — это *амплитудная модуляция* звуковых колебаний неоднородностями гравитационного потенциала (см. (6.5.31)).

Осцилляционная компонента в спектре возмущений плотности в области масштабов $k > k_{\text{eq}}$ (сахаровские осцилляции) оставила характерные корреляционные следы в пространственном распределении галактик и эволюции структуры (числа галактик, квазаров, облаков лайман-альфа и др.) с красным смещением. Рисунок 6.3 демонстрирует наблюдаемые отклонения догалактического спектра мощности $P(k)$ от среднего сглаженного $\bar{P}(k)$, согласующиеся с теоретическими расчетами стандартной модели (ср. рис. 3.12). Эти наблюдения прямо свидетельствуют о том, что наша Вселенная действительно проходила стадию *звуковых колебаний барионов* в дорекомбинационную

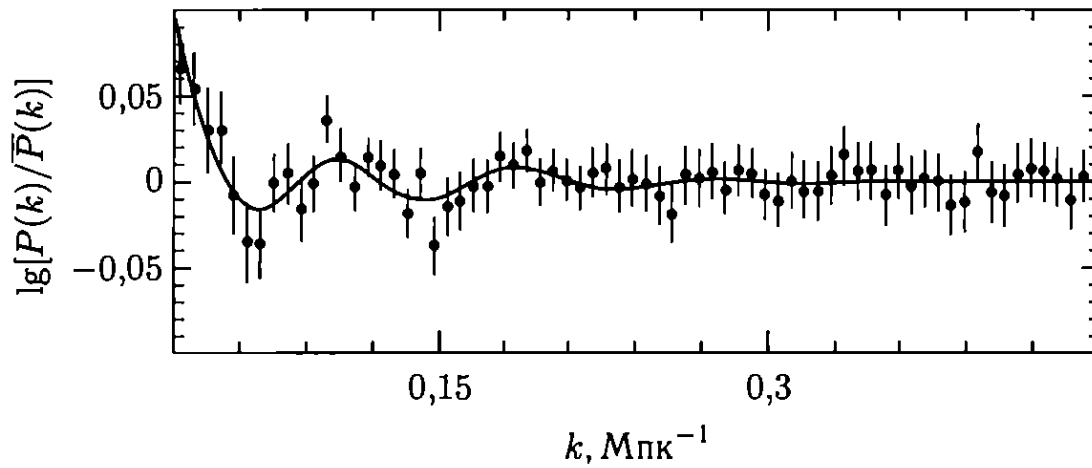


Рис. 6.3. Звуковая модуляция спектра мощности $P(k)$ [Percival et al., 2007]

эпоху и временная фаза этих колебаний соответствовала растущему адиабатическому решению начальных возмущений плотности.

Мы вернемся к образованию крупномасштабной структуры Вселенной в последующих главах. А сейчас обратимся к геометрии ранней Вселенной, построению лагранжевой теории космологических возмущений, к их квантованию и генерации.

КОНФОРМНАЯ НЕИНВАРИАНТНОСТЬ ВОЗМУЩЕНИЙ

Формирование космологического хаббловского потока материи происходит естественным образом в период локально-инфляционного расширения гравитирующей среды. При этом размер начальных неоднородностей материи быстро нарастает, а их амплитуда эффективно подавляется динамически, что, казалось бы, не оставляет никаких шансов для образования новой космологической структуры в эпоху последующего замедленного расширения.

Однако наблюдения свидетельствуют об обратном. На фоне доминирующего хаббловского расширения идет непрерывное строительство крупномасштабной структуры Вселенной. Для запуска этого процесса необходимы начальные возмущения плотности: хаббловское течение должно быть изначально слабо искажено в широком диапазоне длин волн.

Как и почему возникают затравочные возмущения?

Физической причиной их рождения является нестационарность гравитационного поля. Переменный гравитационный потенциал Вселенной параметрически воздействует на динамику мелкомасштабных флуктуаций однородной среды, приводя к их усилению в ходе эволюции.

В этой и последующих главах мы рассмотрим процессы генерации и развития структуры — от создания начальных неоднородностей до образования анизотропии реликтового излучения, галактик и их дальнейшей эволюции.

7.1. Действие возмущений плотности

Лагранжианы S-, V- и T-мод возмущений были получены в главе 5. В этом разделе мы рассмотрим более детально S-моду квазифридмановской геометрии.

Для простоты предположим, что материя однокомпонентна и для нее выполнены оба условия идеальности:

$$S = 0, \quad \delta p_c = \beta^2 \delta \epsilon_c, \quad (7.1.1)$$

где $\beta^2 = \beta^2(t)$ — положительная функция времени.

В этом случае изометрические ветви возмущений отсутствуют ($\widehat{\delta}_p = 0$) и эволюция скаляра кривизны описывается уравнением (6.3.8) в упрощенной форме:

$$\ddot{q} + \left(3H + 2 \frac{\dot{\alpha}}{\alpha}\right) \dot{q} - \beta^2 a^{-2} \Delta q = 0. \quad (7.1.2)$$

Уравнение движения скаляра q может быть получено путем варьирования действия:

$$S[q] = \int L(q) a^3 dt d\mathbf{x} = \int \bar{L} d\eta d\mathbf{x}, \quad (7.1.3)$$

где

$$L = L(q) = \frac{\alpha^2}{2} \left(\dot{q}^2 - \frac{\beta^2}{a^2} \nabla q^2 \right) \quad (7.1.4)$$

— лагранжиан возмущений плотности квази Хаббловского потока гравитирующей материи,

$$\bar{L} \equiv a^4 L = \frac{\bar{\alpha}^2}{2} (q'^2 - \beta^2 \nabla q^2) \quad (7.1.5)$$

— лагранжева плотность поля q в координатах конформно-плоского пространства Минковского, $\bar{\alpha} = \alpha a$.

Так как обратный путь от уравнения к лагранжиану определяет последний с точностью до умножения на константу, то нам следует проверить правильность коэффициента в действии S-моды (7.1.3).

Для этого достаточно проверить нерелятивистский предел. Перепишем действие (7.1.3) в виде:

$$L(q) = \frac{\alpha^2}{2a^2} (q'^2 - \beta^2 \nabla q^2) = \frac{1}{2} (\varepsilon + p) (\beta^2 \delta_\varepsilon^2 - \mathbf{v}^2), \quad (7.1.6)$$

где сопутствующее возмущение плотности и скорость материи соответственно равны

$$\delta_\varepsilon = \frac{q'}{H\beta^2}, \quad \mathbf{v} = -\frac{\nabla q}{H}$$

(здесь \mathbf{v} есть 3-скорость в евклидовой калибровке $A = 0$, см. (4.5.11)).

Оба слагаемых в правой части (7.1.6) описывают плотности потенциальной и кинетической энергий поля q . В этой форме действие напоминает звуковое поле негравитирующей среды:

q выступает в роли потенциала скорости, а \dot{q} — возмущение плотности. Полная плотность энергии возмущений имеет вид:

$$\mathcal{E} = \frac{1}{2} \frac{\beta^2 \delta \epsilon_c^2}{\epsilon + p} + \frac{1}{2} (\epsilon + p) \mathbf{v}^2. \quad (7.1.7)$$

Это соотношение полностью соответствует гидродинамическому пределу звуковой волны с частотой $\omega = c_s k$ и скоростью звука $c_s = \beta$.

Итак, мы построили замкнутую лагранжеву теорию S-моды космологических возмущений для одной материальной среды. Единственным ограничением свойств среды являются условия идеальности (7.1.1), где второе уравнение задает скорость распространения скалярных возмущений в единицах скорости света β (для тензорной моды $\beta = 1$). Данная теория легко обобщается на случай многокомпонентных и других сред, однако каждый раз это приводит к модификации лагранжиана. В этом состоит коренное различие в построении теорий S- и T-мод возмущений: T-мода не требует информации о среде, тогда как для S-моды необходимо знать структуру среды.

В разделе 6.4 мы продемонстрировали, что многокомпонентные идеальные среды квази Хаббловского потока могут быть описаны в рамках теории фундаментальных скалярных полей. Этот результат открывает дорогу для применения идеальных сред в теории очень ранней Вселенной. В силу важности этого вывода рассмотрим более подробно случай одной материальной среды, где эквивалентность теории поля и потенциального течения идеальной среды очевидна.

Скалярное материальное поле φ геометрически тождественно потенциалу скорости материи, поскольку его связь с вектором скорости, по определению, имеет вид:

$$u_\mu = \frac{\varphi_{,\mu}}{w}, \quad (7.1.8)$$

где $w^2 = g^{\mu\nu} \varphi_{,\mu} \varphi_{,\nu}$ — нормировочная функция.

Плотность лагранжиана материи зададим в произвольной форме:

$$\mathcal{L} = \mathcal{L}(\varphi, w). \quad (7.1.9)$$

Остальные характеристики материи получаем из тензора энергии-импульса (6.4.3):

$$\begin{aligned} \epsilon &\equiv n w - \mathcal{L}, & p &\equiv \mathcal{L}, \\ n &\equiv \frac{\partial \mathcal{L}}{\partial w}, & v &= \frac{\delta \varphi}{\dot{\varphi}}, \end{aligned} \quad (7.1.10)$$

откуда следует соотношение между $\delta\varepsilon_c$ и δp_c (см. (4.5.12), (4.5.13)):

$$\frac{\delta\varepsilon_c}{\delta p_c} \equiv \frac{w\delta n - \dot{n}\delta\varphi}{n(\delta w - \dot{w}v)} = \frac{wn_{,w}}{n} \equiv \frac{w}{n} \frac{\partial^2 \mathcal{L}}{\partial w^2} \equiv c_s^{-2} = c_s^{-2}(\varphi, w). \quad (7.1.11)$$

Для линейных возмущений функция $c_s^{-2}(\varphi, w)$ в соотношении (7.1.11) должна быть взята в нулевом порядке, в котором она зависит только от времени: $c_s^{-2} = c_s^{-2}(t)$. Это доказывает, что фоновые функции c_s и β идентичны и, следовательно, оба способа задания характеристик материи (7.1.1) и (7.1.11) для вывода ключевого уравнения (7.1.2) также идентичны.

В следующем разделе мы покажем, что математическая аналогия между потенциальными (невращательными) движениями идеальной среды и теорией скалярного поля идет глубже и может быть распространена на все порядки теории возмущений. А сейчас вернемся к фундаментальному соотношению (7.1.9), которое, бесспорно, является избыточным для построения линейной теории скалярных возмущений.

В полевом подходе лагранжиан космологических возмущений плотности можно получить путем *прямого* разложения материальной плотности (7.1.9) и действия Гильберта вплоть до второго порядка малости по переменным возмущений $\delta g_{\mu\nu} \equiv h_{\mu\nu}$ и $\delta\varphi \equiv wv$ (см. (5.4.13)). Предполагая выполненными уравнения нулевого порядка ($F = m_G = 0$) и ограничения дальнего действия и однокомпонентности первого порядка малости ($\Pi = \hat{\delta} = 0$), а также опуская фридмановскую часть действия и члены с высшими производными в форме полных дивергенций, из (5.4.11) и (5.3.6) получаем:

$$S[q, \mathfrak{S}_i^j] = \int (\bar{L}_S + \bar{L}_T) d\eta dx, \quad (7.1.12)$$

$$\bar{L}_S = \bar{L}(q) = \frac{\bar{\alpha}_S^2}{2} (q'^2 - \beta^2 \nabla q^2), \quad (7.1.13)$$

$$\bar{L}_T = \bar{L}(\mathfrak{S}_i^j) = \frac{\bar{\alpha}_T^2}{2} \eta^{\mu\nu} \mathfrak{S}_{i,\mu}^j \mathfrak{S}_{j,\nu}^i, \quad (7.1.14)$$

$$\bar{\alpha}_S^2 \equiv \frac{a^2 \gamma}{4\pi G \beta^2}, \quad \bar{\alpha}_T^2 \equiv \frac{a^2}{8\pi G}. \quad (7.1.15)$$

Переменные q и \mathfrak{S}_{ij} выделены тем, что входят в собственные лагранжианы в виде первых производных, то есть математически эквивалентны *пробным безмассовым полям* во внешнем фридмановском пространстве-времени. Постоянные факторы в выборе этих переменных фиксированы коэффициентами действия $\bar{\alpha}_S$ и $\bar{\alpha}_T$. Этими условиями канонические переменные q

и \mathfrak{S}_{ij} космологических возмущений квази Хаббловского потока однозначно определены.

7.2. Эффект параметрического усиления

Модель Фридмана связана конформным преобразованием с плоским пространством-временем Минковского. Чтобы в этом убедиться, надо перейти от физического времени t к координатному времени Минковского:

$$\eta = \int \frac{d\eta}{a},$$

оставив пространственные переменные x неизменными. Конформным метрическим параметром является квадрат масштабного фактора $a = a(t)$, описывающий расширение гравитирующей материи (см. (2.7.11)).

Рассматривая эволюцию физических полей в модели Фридмана, мы будем называть поле *конформно-инвариантным*, если существуют такие преобразования полевых переменных, при которых уравнение эволюции данного поля, записанное в координатах Минковского, идентично динамическому уравнению движения соответствующего физического поля в метрике Минковского. Полученные таким образом полевые переменные называются *конформными*, а их связь с физическими полями — *конформными преобразованиями*. Амплитуды гравитирующих конформно-инвариантных полей адиабатически изменяются в соответствии с конформными преобразованиями из-за расширения Вселенной, однако их внутренняя динамика остается той же, что и в отсутствие гравитации.

Примером конформно-инвариантного поля является векторная мода V космологических возмущений. Действительно, переходя к конформным переменным

$$\bar{S} = \frac{S_i}{16\pi G} = \frac{\mathfrak{S}_i}{16\pi G}, \quad \bar{v} = a^3(\varepsilon + p)v_i,$$

получаем уравнения (4.5.5), (4.5.9) в форме

$$\Delta \bar{S} = \bar{v}, \quad (7.2.1)$$

эквивалентной уравнению Эйлера для вихревой моды негравитирующей среды в пространстве Минковского. Никаких новых физических эффектов в гравитирующей среде по сравнению с негравитирующей не возникает. Математическая эквивалентность двух теорий означает, что их физические лагранжианы

в координатах (η, \mathbf{x}) совпадают с точностью до слагаемых с полными производными.

Иначе обстоит дело с модами S и T космологических возмущений.

Для динамического сопоставления гравитирующей и негравитирующей сред перепишем действие (7.1.13) в конформно-сопутствующих координатах Минковского в приведенных конформных переменных:

$$\begin{aligned} \bar{L} &= \frac{1}{2} \bar{\alpha}^2 (q'^2 - \beta^2 \nabla q^2) = \\ &= \frac{1}{2} \left(\bar{q}'^2 - \beta^2 \nabla \bar{q}^2 + U \bar{q}^2 - (\bar{\alpha} \bar{\alpha}' q^2)' \right), \end{aligned} \quad (7.2.2)$$

где

$$\bar{q} = \bar{\alpha} q, \quad U = \frac{\bar{\alpha}''}{\bar{\alpha}}.$$

Последнее слагаемое в правой части (7.2.2) имеет форму полной временной производной, не влияющей на поле q .

Таким образом, с теоретической точки зрения задача о космологических возмущениях плотности в метрике Фридмана эквивалентна задаче о безмассовом действительном поле q в метрике Минковского в переменном внешнем потенциале ¹⁾:

$$\bar{\alpha}^2 = \bar{\alpha}_S^2(\eta) = \frac{a^2 \gamma}{4\pi G \beta^2}.$$

Динамическое влияние этого потенциала на поле q описывается членом с трением ($\propto \dot{q}$) в уравнении (7.1.2). В терминах конформного скаляра \bar{q} действие имеет стандартную форму свободного поля с переменной эффективной массой $-U$. Однако физически правильнее считать, что массовый член описывает *взаимодействие* поля с внешней геометрией, поскольку «масса» в данном случае характеризует не само поле, а его связь с внешней нестационарной метрикой: U зависит от масштабного фактора $a(t)$ и его первых четырех производных.

Эффективный потенциал $U \sim \bar{H}^2$ не имеет нерелятивистского аналога и ответственен за приливное параметрическое воздействие хаббловского потока на поле возмущений плотности. Оно описывается ненулевой правой частью волнового уравнения для конформного поля \bar{q} :

$$\square_\beta \bar{q} = U \bar{q}. \quad (7.2.3)$$

где $\square_\beta \equiv \partial^2 / \partial \eta^2 - \beta^2 \Delta$ — звуковой даламбертиан.

¹⁾ Аналогичная ситуация и с тензорной модой, где $\bar{\alpha}_T = (8\pi G)^{-1/2} a$ (подробнее см. раздел 8.5).

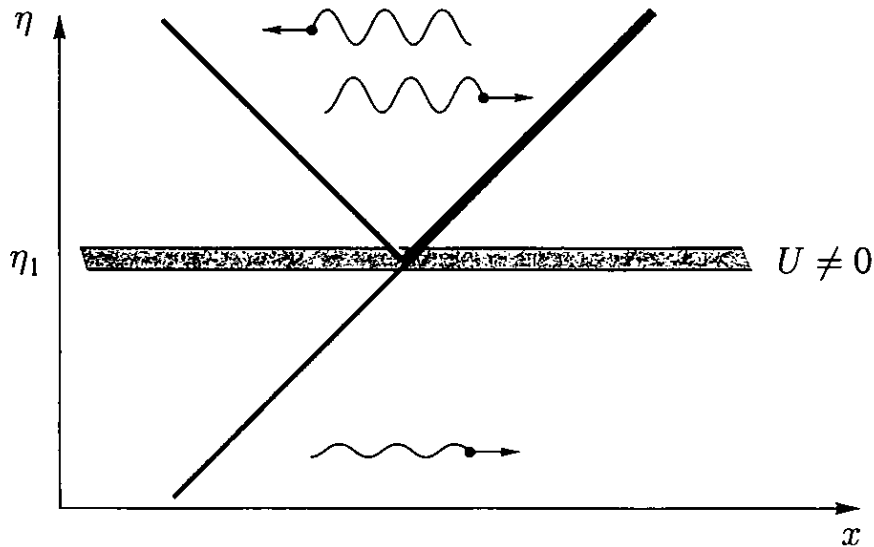


Рис. 7.1. Иллюстрация решения уравнения (7.2.3)

Решение уравнения (7.2.3) хорошо известно в квантовой механике: это распространение волн в пространстве с барьером U . На барьер падает волна и отражается от него, образуя перед барьером стоячую волну, но незначительная доля энергии просачивается через барьер и уходит в виде бегущей волны.

В космологии подобный барьер возникает не в пространстве, а во времени (рис. 7.1). До появления барьера (внешнего воздействия) имеется затравочная флуктуация поля \bar{q} в виде бегущей волны. После параметрического воздействия волна усиливается: рождается стоячая волна (пара частиц) с определенными фазой и амплитудой. Количественно результирующая амплитуда зависит от временной фазы, в которой произошло воздействие на волну. При усреднении по начальной фазе бегущей волны всегда имеется выигрыш в амплитуде (закон роста энтропии или рождение частиц). В последующих главах мы получим фазовые и спектральные характеристики генерированных таким образом возмущений плотности, а сейчас вернемся к параметрическому эффекту усиления.

При $U = 0$ мода S фридмановской метрики удовлетворяет тому же уравнению, что и негравитирующие возмущения плотности в метрике Минковского. В этом случае поле q конформно-инвариантно возмущениям негравитирующей среды. Физические следствия идентичны: отсутствует эффект усиления, поле q эволюционирует в адиабатическом режиме, сохраняется адиабатический инвариант — число квантов поля не меняется со временем. Вселенная с $U = 0$ гравитационно устойчива. Важным случаем адиабатического режима расширения является радиационно-доминированная Вселенная: $a \propto \eta$, $\gamma = 2$, $\beta = 1/\sqrt{3}$, $U = 0$.

В общем случае $U \neq 0$ и поле q конформно инвариантно. Это значит, что космологическая S-мода взаимодействует с фоновой нестационарной метрикой, что принципиально отличает ее от звуковых волн негравитирующей материи. Потенциал U изменяет эффективную частоту звуковых колебаний и приводит к несохранению числа частиц поля q со временем. Происходит *параметрическое усиление* неоднородностей плотности в ходе расширения Вселенной.

В лабораторном пределе частота колебаний велика и потенциал U несущественен ($\omega \gg \bar{H}$). Однако прежде чем попасть в коротковолновую область и образовать галактики, космологические возмущения провели в длинноволновой зоне $\omega < \bar{H}$ большую часть своих хаббловских времен

$$\int H dt = -\ln a \gg 1. \quad (7.2.4)$$

В этом пределе главные члены разложения решения уравнения (6.3.8) по малому параметру ω/\bar{H} имеют вид:

$$q = c_1 + c_2 \int \frac{\beta^2 da}{\gamma a^4 H} + \int \hat{\delta}_p \frac{da}{a} + \dots, \quad (7.2.5)$$

$$p \equiv \frac{\partial \bar{L}}{\partial q'} = \bar{\alpha}^2 q' = \frac{1}{4\pi G} \left(c_2 + \Delta c_1 \int \frac{\gamma da}{H} \right) + \bar{\alpha}^2 \bar{H} \hat{\delta}_p + \dots, \quad (7.2.6)$$

где $c_1 = c_1(\mathbf{x})$ и $c_2 = c_2(\mathbf{x})$ — функции пространственных координат, задающие два фундаментальных решения.

Скаляр q «замораживается» во времени при условии сходимости интегральной амплитуды c_2 в правой части (7.2.5), что выполняется в большинстве сценариев ранней Вселенной. В этом случае предел интегрирования выбран так, чтобы решение $\propto c_2$ затухало с ростом масштабного фактора. Исторически решение $\propto c_1$ называют *растущим*, а $\propto c_2$ — *падающим*. Эти решения ассоциируются соответственно с полем q и с его канонически сопряженным импульсом p .

Отношение между падающей и растущей ветвями решений затухает со временем $\propto \gamma a^3 H$ ($\propto a$ на радиационно-доминированной стадии). Если принять, что в эпоху Большого взрыва при $\bar{H} \sim k$ обе амплитуды были одного порядка, то за все время существования Вселенной начальная падающая ветвь поля уменьшилась по сравнению с растущей более чем на 100 порядков. По этой причине крупномасштабная структура Вселенной обязана своим происхождением исключительно растущей

первичной ветви решения c_1 , а начальную падающую ветвь c_2 современная теория образования структуры игнорирует ¹⁾).

7.3. Потенциальные течения и теория поля

Для описания движений S- и V-мод космологических возмущений необходимы элементарные знания о свойствах материи. Мы видели, что потенциальные течения идеальной среды математически эквивалентны полевому приближению в рамках линейной теории S-моды. Докажем, что эта аналогия простирается гораздо глубже: точное описание потенциальных движений идеальной среды математически тождественно теории скалярного поля во *всех* порядках теории возмущений.

В рамках концепции идеальной среды ее течения можно описывать с помощью линий тока — мировых линий элементов среды на многообразии M . Рассмотрим *конгруэнцию* (пучок) линий тока такой среды. Пусть λ — аффинный параметр вдоль линий тока, а x_a — метка (номер) мировой линии, неизменная на линии тока, $a = 1, 2, 3$. Параметр λ можно выбрать многими способами. Он определен с точностью до преобразований, изменяющих параметризацию линий $x_a = \text{const}$:

$$\lambda \rightarrow \lambda^*(\lambda, x_a).$$

Поскольку мы имеем дело с конгруэнцией, то требуется выполнение лишь одного условия: каждой точке x^μ в M приписано единственное значение λ . При выполнении этого условия выбор функции $\lambda = \lambda(x^\mu)$ задает конкретное *расслоение* на M (рис. 7.2) и справедливы следующие утверждения.

- Функция $\lambda(x^\mu)$ является скаляром, не зависящим от координат, покрывающих многообразие M . На этом этапе не требуется метрика: достаточно иметь многообразие и координатные сетки, покрывающие его.
- Значение $\lambda(x^\mu) = \text{const}$ ставит метку (номер) гиперповерхности, пересекающей линии тока. Это означает, что функция $\lambda = \lambda(x^\mu)$ задает систему гиперповерхностей, осуществляющих расслоение многообразия.

¹⁾ Это не относится к новообразованным решениям падающего типа вследствие взаимодействия адиабатических и изометрических ветвей возмущений в ходе эволюции Вселенной.

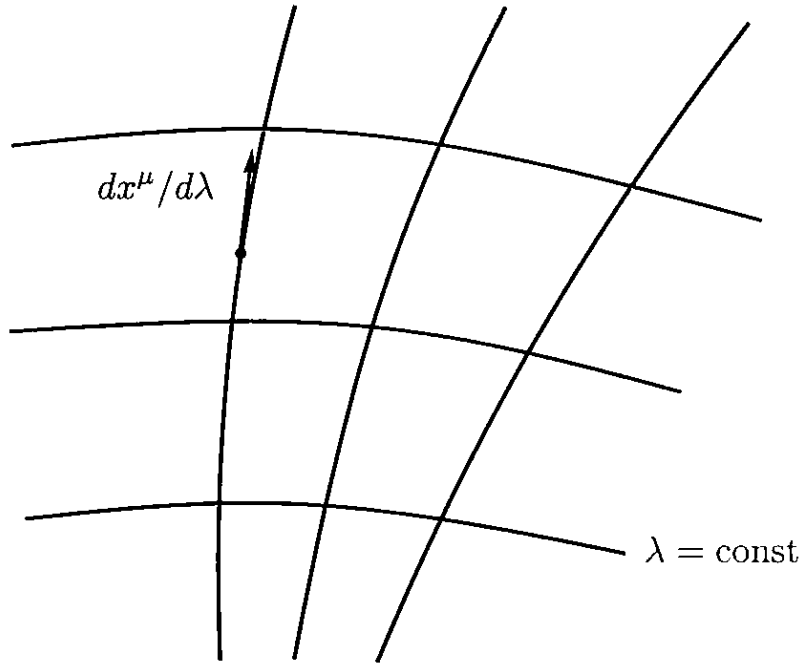


Рис. 7.2. Пучок линий тока идеальной среды; λ — аффинный параметр, задающий расслоение на \mathcal{M} , $dx^\mu/d\lambda$ — касательный вектор к линиям тока, не зависящий от выбора λ

- $\lambda_{,\mu}$ — ковариантный вектор, ортогональный к гиперповерхности $\lambda = \text{const}$:

$$\lambda_{,\mu} \delta x^\mu|_{\lambda=\text{const}} = 0. \quad (7.3.1)$$

Разные $\lambda \rightarrow \lambda^*$ определяют разные расслоения.

- $dx^\mu/d\lambda$ — контравариантный вектор, касательный к конгруэнции линий тока (в каждой точке есть только один вектор). Векторное поле $dx^\mu/d\lambda$ определяет саму конгруэнцию и не зависит от выбора функции λ .

Итак, у нас есть два векторных поля, однако мы не можем сопоставить их между собой, поскольку пока не имеем геометрии. Для определения норм, углов, расстояний и синхронизаций нужна метрика, которую мы обозначим $g_{\mu\nu}$.

В этом случае можно ввести два единичных векторных поля ($u_\mu u^\mu = v_\mu v^\mu = 1$): одно из них — касательное к конгруэнции (не зависит от выбора λ):

$$u^\mu = w \frac{dx^\mu}{d\lambda}, \quad w^{-2} = g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}, \quad (7.3.2)$$

другое — ортогональное к расслоению λ :

$$v_\mu = \frac{\lambda_{,\mu}}{\tilde{w}}, \quad \tilde{w}^2 = g^{\mu\nu} \lambda_{,\mu} \lambda_{,\nu}, \quad (7.3.3)$$

где w и \tilde{w} — нормировочные скаляры. Если бы материя двигалась по геодезическим, то нормировочная функция w была бы интегралом движения и ее можно было совместить с единицей.

Однако в общем случае это не так, зависимость w и \tilde{w} от координат говорит о негеодезическом потоке.

По построению наши поля различны.

Векторное поле v^μ всегда потенциально. Выбором потенциала λ можно задать любое расслоение пространства-времени на времениподобные линии, касательные к v^μ , и ортогональные к ним пространственноподобные сечения $\lambda = \text{const}$.

Поле скорости u^μ определяет конгруэнцию линий тока материи, а ее ковариантная производная определяет деформацию этой конгруэнции:

$$u_{\mu;\nu} = a_\mu u_\nu + \sigma_{\mu\nu} + \omega_{\mu\nu}. \quad (7.3.4)$$

Как видно, полная деформация состоит из трех компонент:

- ускорения

$$a_\mu = u^\nu u_{\mu;\nu}, \quad (7.3.5)$$

- собственно деформации (объемной и сдвиговой)

$$\sigma_{\mu\nu} = \sigma_{(\mu\nu)} - a_{(\mu} u_{\nu)} \quad (7.3.6)$$

- и вращения

$$\omega_{\mu\nu} = \omega_{[\mu\nu]} - a_{[\mu} u_{\nu]}, \quad \Omega^\mu = e^{\mu\nu\alpha\beta} u_\nu \omega_{\alpha\beta}, \quad (7.3.7)$$

где индексы, заключенные в круглые или квадратные скобки, предполагают соответственно симметризацию или антисимметризацию, $e_{\alpha\beta\gamma\delta}$ — абсолютный антисимметричный тензор. Задание скорости вращения материи Ω^μ эквивалентно заданию тензора вращения $\omega_{\mu\nu} = e_{\mu\nu\alpha\beta} u^\alpha \Omega^\beta$.

Все три компоненты ортогональны скорости потока:

$$a_\mu u^\nu = \sigma_{\mu\nu} u^\mu = \Omega_\mu u^\mu = 0. \quad (7.3.8)$$

Следовательно, они описывают деформацию пространственноподобной ортогональной к потоку гиперповерхности и ее эволюцию при переносе вдоль векторного поля u^μ .

В случае движения без вращения ($\Omega^\mu = 0$) вектор u_μ потенциален и описывает единственное расслоение, ортогональное потоку. Поскольку λ перебирает все возможные расслоения в M , то всегда существует такое единственное $\lambda = \varphi$, при котором поля u^μ и v^μ тождественно совпадают:

$$\Omega^\mu = 0: \quad u^\mu = w \frac{dx^\mu}{d\varphi}, \quad u_\mu = \frac{\varphi_{,\mu}}{w}, \quad (7.3.9)$$

где $w^2 = g^{\mu\nu} \varphi_{,\mu} \varphi_{,\nu}$. Уравнения (7.3.9) определяют потенциальные (невращательные) течения идеальной среды во всех поряд-

ках теории возмущений и фиксируют скаляр φ с точностью до преобразований:

$$\varphi \rightarrow \varphi^*(\varphi). \quad (7.3.10)$$

Любая система отсчета $x^\mu = (\varphi, \mathbf{x})$ удовлетворяет условию сопутствия в форме $u_\mu = w^{-1} \delta_\mu^0$, однако сопутствующая система с $\mathbf{x} = x_a$ ортогональна по построению ($g_{0\mu} = w^{-2} \delta_\mu^0$, поскольку $u^\mu = w \delta_0^\mu$).

Итак, для потенциального потока идеальной среды справедлива следующая теорема.

Скаляр сопутствующего времени идеальной материи есть функция потенциала скорости ее движения φ .

При этом собственное время потока имеет вид:

$$\delta\tau = u_\mu dx^\mu = \frac{d\varphi}{w}, \quad (7.3.11)$$

а тензор энергии-импульса среды $T_\mu^\nu = (\varepsilon + p) u_\mu u^\nu - p \delta_\mu^\nu$ (где p и ε — сопутствующие давление и плотность энергии).

Из этой теоремы следует, что потенциальные течения идеальной материи описываются теорией скалярного поля φ , роль которого выполняет *потенциал скорости материи*.

Поясним этот вывод. Если вращение равно нулю, то можно единственным образом выбрать такое расслоение, при котором гиперповерхности $\lambda = \text{const}$ ортогональны конгруэнции. Если вращение не равно нулю, то такого глобального расслоения не существует. Возьмем для примера Землю. Она вращается. Метрика, которую создает Земля, — это метрика с вращением: гравитационное поле непотенциально (гироскопы поворачиваются относительно далеких звезд). Можно синхронизовать часы в какой-то конечной области пространства-времени (например, в Московской области) и создать локально-ортогональное расслоение, но все пространство оно охватить не может.

Существует классический результат, что в любой геометрии, в том числе и с вращением, можно построить синхронную систему отсчета с ортогональным (но не сопутствующим) расщеплением. Методика несложная: берется пространственноподобная гиперповерхность, и на ней в каждой точке необходимо так направить (толкнуть) частицы системы отсчета, чтобы они двигались ортогонально к этой гиперповерхности. Тогда эта ортогональность будет сохраняться во времени. Но такая синхронная система отсчета не будет покрывать все пространство-время: она коллапсирует и сжимается в каустику.

Подведем итог. Идеальная среда подразумевает вполне конкретный вид тензора энергии-импульса. А полевой подход выстраивается от лагранжиана. При определенных условиях эти языки тождественны. В важном для нас случае *потенциальных* течений оба подхода оказываются математически эквивалентными.

Ранее мы сформулировали достаточные условия для определения динамики S-моды космологических возмущений одной среды (см. (7.1.1)):

$$S = 0, \quad \delta\epsilon_c = \beta^{-2}\delta p_c. \quad (7.3.12)$$

Они описывают паскалево (изотропное) давление и вводят новую функцию¹⁾ $\beta(t)$, которая может иметь как гидродинамическую трактовку (скорость звука), так и полевую (вторая производная лагранжиана). Для построения линейной теории возмущений нам необходимо знать функцию β^{-2} (или матрицу β_{ml}^{-2} , если число компонент среды больше единицы), информация же о физической природе этой функции для вывода уравнений движения не требуется и является избыточной.

7.4. Массивные поля в космологии

Важная роль в истории Вселенной отводится слабовзаимодействующим массивным скалярным полям, которые естественным образом приводят к промежуточным стадиям инфляции в ходе своей динамической релаксации к равновесному состоянию. Вопросы о том, как такие поля возникли, каким образом они приобрели массу и почему оказались в неравновесном состоянии, отсылают нас к квантовой эпохе нарушения симметрии в ранней Вселенной.

Наблюдения свидетельствуют как минимум о двух космологических стадиях ускоренного расширения. Одна из них связана с эпохой Большого взрыва и длилась не менее 50 хаббловских времен. В результате сформировалась та область пространства-времени, которую мы наблюдаем сегодня в пределах горизонта видимости.

Вторая стадия инфляции связана с доминированием темной энергии. Эта стадия началась около трех с половиной миллиар-

¹⁾ Заметим, что эта характеристика среды не требуется для описания динамики хаббловского потока материи в нулевом порядке теории возмущений. Необходимость в функции β возникает при построении первого порядка возмущений (моды S).

дов лет тому назад и не прошла еще даже одного хаббловского времени, которое составляет для нее 17 миллиардов лет:

$$H_E \simeq 60 \text{ км} \cdot \text{с}^{-1} \cdot \text{Мпк}^{-1}. \quad (7.4.1)$$

Возможно, во Вселенной были и другие инфляционные стадии, обусловленные существованием массивных слабых полей, оставшихся после Большого взрыва.

Мы полагаем, что в условиях неравновесной динамики, обусловленной гравитационным взаимодействием и иными вновь образованными взаимодействиями в ранней Вселенной, могли возникнуть квазиклассические массивные поля, в том числе и скалярные, которые эволюционировали от своих начальных состояний, где они оказались в эпоху образования, к положениям ближайших локальных минимумов их общей потенциальной энергии само- и взаимодействий. Многообразие возможных конфигураций и «ландшафтов» полей и взаимодействий, в том числе многокомпонентных (см., напр., (6.4.1)), диктует многообразие форм динамической релаксации и ассоциированных с ними функций $a(t)$ расширяющейся материальной среды. При определенных условиях (например, в состояниях «ожидания», см. (7.4.4)) от систем полей могли «отсоединиться» обособленные группы полей, слабым образом взаимодействовавшие с остальной материей, или отдельные сверхслабые поля, подразумевающие существование в общем лагранжиане аддитивных членов, зависящих каждый от своего поля.

Для того чтобы понять, как и почему начинаются, эволюционируют и заканчиваются стадии космологической инфляции, рассмотрим поведение классических скалярных полей в геометрии Фридмана.

Возьмем одно из таких материальных полей φ с лагранжевой плотностью

$$\mathcal{L}(\varphi, w) = \frac{w^2}{2} - V(\varphi), \quad (7.4.2)$$

у которой кинетическая энергия имеет «стандартный» вид ($w^2 = = \varphi_{,\mu} \varphi^{,\mu}$), а массовый член описывается функцией $V = V(\varphi)$, называемой *потенциальной энергией* поля. Форма функции $V(\varphi)$ определяет динамику поля (о классификации динамических траекторий см. раздел 10.5). В первую очередь нас будут интересовать свойства массивных гравитирующих полей ($V = m^2 \varphi^2 / 2$), поэтому по умолчанию будем считать потенциальную энергию положительной величиной, имеющей минимум при $V = 0$, который мы совместим с точкой $\varphi = 0$.

В силу пространственной симметрии поле $\varphi = \varphi(t)$ и кинетический скаляр $w = \dot{\varphi}$ являются функциями времени ¹⁾, и уравнение движения поля имеет вид:

$$\ddot{\varphi} + 3H\dot{\varphi} + V_{,\varphi} = 0, \quad (7.4.3)$$

$$H^2 = \frac{4\pi G}{3} (\dot{\varphi}^2 + 2V + 2\varepsilon), \quad \gamma = \frac{4\pi G}{H^2} (\dot{\varphi}^2 + \varepsilon + p),$$

где ε и p — соответственно плотность и давление остальной материи.

При $H > \sqrt{V}/M_P$ член с трением в (7.4.3) определяет эволюцию, в результате чего поле быстро релаксирует к промежуточному состоянию ожидания:

$$\varphi \simeq \varphi_0 = \text{const}, \quad (7.4.4)$$

в котором оно пребывает до тех пор, пока убывающая хаббловская функция $H(t)$ не сравняется с потенциальной энергией $\sqrt{V_0}/M_P$, где $V_0 = V(\varphi_0)$. Застывание происходит при значении φ_0 , близком к начальной величине классической части поля в эпоху ее появления в очень ранней Вселенной (рис. 7.3). Для определенности будем считать $\varphi_0 > 0$.

Темп космологического расширения на этой стадии определяется остальной материей (излучением или другими частицами или полями). Так, при квазистепенном расширении $Ht \simeq \gamma^{-1} \simeq \text{const}$ решение (7.4.3) в состоянии ожидания имеет вид:

$$\varphi = \varphi_0 \left(1 - \frac{c_0}{2} t^2 \right), \quad c_0 = \frac{\gamma V_{0,\varphi}}{(3 + \gamma) \varphi_0} \simeq \text{const}, \quad (7.4.5)$$

где $|c_0|t^2 \ll 1$ при $\varphi_0 > M_P$.

В процессе эволюции функция H убывает и сравнивается с $\sqrt{V_0}/M_P$. В этот период решение «садится» на затягивающую инфляционную сепаратрису:

$$H^2 \simeq \frac{8\pi V}{3M_P^2}, \quad \dot{\varphi} = -\frac{V_{,\varphi}}{3H}, \quad \frac{3\dot{\varphi}^2}{2V} \leq \gamma < 1, \quad (7.4.6)$$

¹⁾ Член с лапласианом в уравнении (7.4.3) следует учитывать при рассмотрении пространственных флуктуаций поля. В квазифридмановском потоке приближенно имеем ($m^2 = 2 d^2 V/d\varphi^2$):

$$\ddot{\varphi} + 3H\dot{\varphi} + (m^2 - a^{-2}\Delta)\varphi = 0.$$

Член с лапласианом быстро затухает с ростом $a(t)$, и им можно пренебречь при рассмотрении эволюции классической части поля $\varphi = \langle \varphi \rangle$.

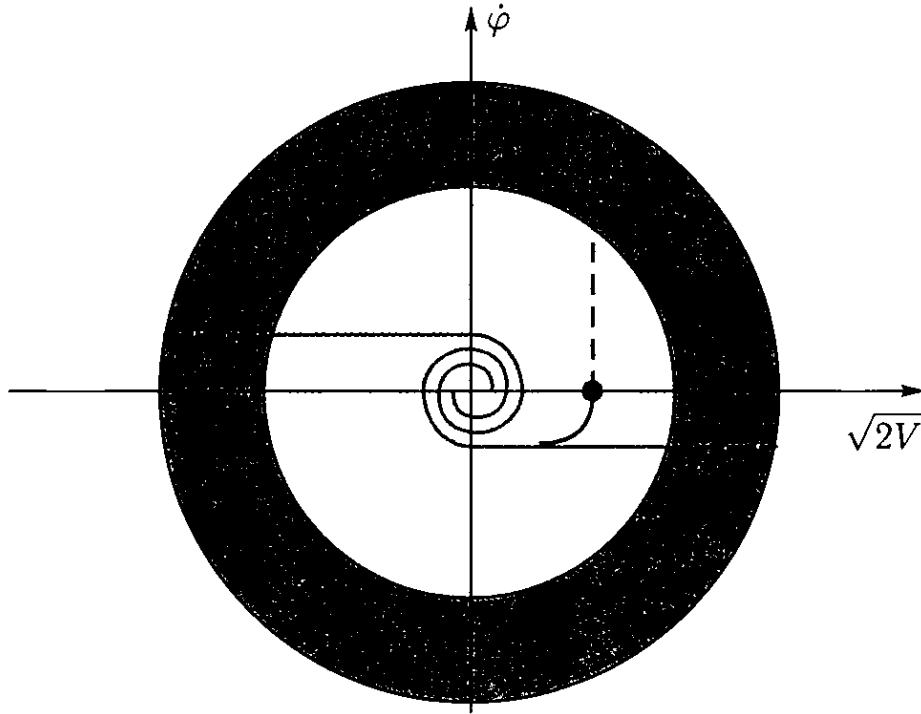


Рис. 7.3. Фазовая диаграмма уравнения (7.4.3) при $V > \epsilon$ (светлый круг). Темное кольцо — область доминирования квантово-гравитационных флуктуаций. Красная точка — состояние ожидания $\varphi \simeq \varphi_0$ (см. решение (7.4.5)). Горизонтальные линии — инфляционная сепаратриса (см. решение (7.4.6))

которая существует, пока член со старшей производной в уравнении (7.4.3) остается малым:

$$\left| \frac{\ddot{\varphi}}{\dot{\varphi}H} \right| < 1. \quad (7.4.7)$$

Поскольку с учетом (7.4.6) имеем

$$\frac{\ddot{\varphi}}{\dot{\varphi}H} = \gamma - \epsilon, \quad \epsilon \equiv \frac{1}{3H^2} \frac{d^2V}{d\varphi^2} \simeq \frac{M_P^2 V_{,\varphi\varphi}}{8\pi V}, \quad (7.4.8)$$

неравенство (7.4.7) можно переписать в виде, называемом *приближением медленного скатывания*:

$$|\epsilon| < 1. \quad (7.4.9)$$

Поле φ , ответственное за существование и динамику инфляционной стадии Вселенной, называется *инфлатоном*.

Инфляция ($\gamma < 1$) может иметь разнообразные формы, однако инфляционная сепаратриса существует тогда и только тогда, когда выполняются оба неравенства:

$$\gamma < 1 \quad \text{и} \quad |\epsilon| < 1. \quad (7.4.10)$$

При этих условиях функции $X = (\varphi, \dot{\varphi}, V, H, N)$ медленно (адиабатически) дрейфуют со временем ($|\dot{X}/X| < H$), тогда как

масштабный фактор ускоренно возрастает (см. (7.2.4)):

$$a = a_m e^{-\mathcal{N}}, \quad t_m - t = \int_0^{\mathcal{N}} \frac{d\mathcal{N}}{H}, \quad (7.4.11)$$

где t_m (или t_m , где m — масса инфлатона) — момент окончания инфляции ($\gamma = 1$), $a_m = a(t_m)$,

$$\mathcal{N} = \mathcal{N}(\varphi) = 8\pi G \int_{\varphi_m}^{\varphi} \frac{V d\varphi}{V_{,\varphi}} \quad (7.4.12)$$

— число инфляционных хаббловских эпох между t и t_m .

Заметим, что условия (7.4.10) не означают еще, что функция γ адиабатична. По отношению к поведению $\gamma(t)$ инфляционная сепаратриса разбивается на два периода.

На первом этапе функция γ эволюционирует под действием обычной материи и может изменяться как степень масштабного фактора, то есть *неадиабатически* (см. (7.4.3)):

$$\gamma \simeq \frac{3(\varepsilon + p)}{2V}. \quad (7.4.13)$$

Только позже, когда член $\varepsilon + p$ затухнет из-за красного смещения и окажется меньше, чем $\dot{\varphi}^2$, эволюция $\gamma(t)$ определяется инфлатоном и становится адиабатической:

$$\gamma \simeq \frac{3\dot{\varphi}^2}{2V} = \frac{1}{\pi} \left(\frac{M_P V_{,\varphi}}{4V} \right)^2. \quad (7.4.14)$$

Мы будем называть этот второй динамический период инфляционной сепаратрисы *адиабатической инфляцией*.

Условия (7.4.10) накладывают некоторые ограничения на гладкость функции $V(\varphi)$:

$$\left| \frac{dV}{d\varphi} \right| < \frac{7V}{M_P}, \quad \left| \frac{d^2V}{d\varphi^2} \right| < \frac{25V}{M_P^2}, \quad (7.4.15)$$

которые выполняются во многих моделях инфляции.

В ходе инфляции потенциальная энергия уменьшается и поле «скатывается» к минимуму $V(\varphi)$. Инфляция заканчивается в момент, когда поле уменьшается до значения φ_m , определяемого условием $\gamma = 1$ ($2V/|dV/d\varphi| \simeq 0,3 M_P$), после этого инфлатон переходит в режим осцилляций. Начальные колебания коррелированы по фазе, но в дальнейшем фаза «расстраивается» под воздействием флуктуаций в причинно-несвязанных областях пространства, и эволюция продолжается в режиме холодных массивных частиц поля φ . Параллельно, быстропеременное поле φ

параметрически воздействует на другие поля и «раскачивает» их степени свободы. Коэффициенты связи (сечения взаимодействий) могут быть малыми, но энергия накачки в условиях ранней Вселенной чрезвычайно велика ($\sim H^2/M_P^2$). При определенной подстройке параметров поле φ распадается на более легкие частицы, продукты распада термализуются, и возникает радиационно-доминированная тепловая «баня» частиц.

Подобный процесс релаксации, по всей видимости, произошел в ранней Вселенной: об этом свидетельствуют оставшиеся от тепловой «бани» частиц реликтовые фотоны. Однако сам расчет стадий осцилляций и разогрева зависит от многих неизвестных нам параметров и по этой причине не может пока быть признан завершенным.

Намного проще обстоит дело с инфляцией. В рассмотренном выше примере, обладающем достаточной степенью общности, динамика определяется одной функцией $V(\varphi)$ и одним параметром — значением замороженного инфлатона φ_0 или полным числом хаббловских эпох \mathcal{N}_0 на данной инфляционной стадии. Предположения о свойствах доинфляционной материи не столь существенны. Более того, поскольку поле эволюционирует очень медленно, то вся наблюдательная космология укладывается в короткий динамический интервал изменения φ , не превышающий единицу¹⁾. Принимая условия гладкости (7.4.15), приходим к выводу, что от всей функции $V(\varphi)$ нам надо знать только несколько ее первых производных в интересующей нас области поля. То есть речь идет о трех-четырех параметрах инфлатона, полностью задающих структуру ранней Вселенной.

¹⁾ Динамический диапазон от пк до Гпк занимает девять порядков и соответствует $\Delta\mathcal{N} = 9 \ln 10 \simeq 20$. Учитывая, что для космологических масштабов $\mathcal{N}_i > 50$, получаем интервал изменения инфлатона, соответствующий диапазону масштабов наблюдательной космологии:

$$\frac{\Delta\varphi}{\varphi_i} \sim \frac{\Delta\mathcal{N}}{2\mathcal{N}_i} < 0,2.$$

Для оценки \mathcal{N}_i положим, что после окончания инфляции при $t \sim t_m$ вся плотность перешла в излучение:

$$H_m \sim \ell_P k_{\text{rad}}^2 e^{2N} \sim m \sim 10^{-6} M_P, \quad e^N \sim \frac{\sqrt{m} M_P}{k_{\text{rad}}} \sim 10^{30},$$

где $k_{\text{rad}} \sim \text{см}^{-1}$ — частота реликтового кванта (см. (8.3.8)), $N = N_m = -\ln a_m$ — число хаббловских времен, прошедших после момента t_m . Полагая, что масштаб k_i равен современному горизонту, получаем оценку величины \mathcal{N}_i :

$$k_i \sim H_i e^{-\mathcal{N}_i - N} \sim H_0, \quad k_m \sim m e^{-N}, \quad e^{\mathcal{N}_i} \sim \frac{k_{\text{rad}}}{H_0} \frac{\varphi_i}{M_P} \sqrt{\frac{m}{M_P}} \sim 10^{25} \sim e^{60}.$$

РОЖДЕНИЕ СТРУКТУРЫ

Геометрия ранней Вселенной содержит малый параметр — амплитуду начальных космологических неоднородностей, из которых развилась структура, — что позволяет исследовать ее с помощью теории возмущений. Откуда взялись эти начальные возмущения? Ответ дают квантовая физика и теория гравитации.

Строго однородное распределение материи невозможно из-за принципа неопределенности Гейзенберга: всегда имеются вакуумные флуктуации метрики и плотности вещества. Эти флуктуации существуют в релятивистском гравитационном поле расширяющейся Вселенной. Гравитационное воздействие переменного во времени однородного потенциала на малые отклонения от однородности и изотропии приводит к параметрическому усилению гармоник с частотами порядка хаббловских ($\omega \sim \bar{H}$). Так происходит рождение частиц *любых* маломассивных физических полей ($m < H$), взаимодействующих с фоновой метрикой минимальным образом в рамках ОТО. Неоднородные распределения метрики и плотности являются примерами таких безмассовых полей.

Эффект или процесс спонтанной генерации физических степеней свободы с конечной длиной волны под действием нестационарного однородного гравитационного поля мы называем *параметрическим* или *квантово-гравитационным*, подчеркивая квантовую природу затравочных флуктуаций и их гравитационную связь с внешним переменным полем.

8.1. Квантование возмущений плотности

В уравнениях S-моды космологических возмущений геометрический скаляр кривизны q предстает в виде действительного *пробного* поля во внешней фридмановской метрике. Поэтому квантование q осуществляется стандартными методами квантовой теории вложенного поля на заданном классическом фоне¹⁾.

¹⁾ Напомним, что в q входят метрические переменные, например гравитационный потенциал Φ . Это означает, что мы квантуем не только материальные степени свободы, но и физические возмущения метрики и частично координаты. Классическими остаются лишь квазиоднородные длинноволновые степени свободы. Используемая процедура квантования является квазиклассическим пределом квантовой теории гравитации (см. раздел 5.4).

В ковариантном обобщении действие и классическое уравнение движения для q имеют вид (см. главы 5, 7):

$$S[q] = \int L_S \sqrt{-g} d^4x, \quad L_S = \frac{1}{2} D^{\mu\nu} q_{,\mu} q_{,\nu}, \quad (8.1.1)$$

$$(D^{\mu\nu} q_{,\mu})_{;\nu} = 0, \quad (8.1.2)$$

$$D_{\mu\nu} = \alpha_S^2 (u_\mu u_\nu - \beta_S^2 p_{\mu\nu}), \quad \alpha_S^2 = \frac{\varepsilon + p}{H^2 \beta_S^2} = \frac{\gamma}{4\pi G \beta_S^2},$$

где u^μ , β_S , $g_{\mu\nu}$ и $p_{\mu\nu} = u_\mu u_\nu - g_{\mu\nu}$ — соответственно скорость хаббловского потока материи, скорость распространения скалярных (звуковых) возмущений в среде, фридмановский метрический тензор и проекционный тензор.

В уравнение движения входит только β^{-2} . Для определенности будем считать $\beta^2 > 0$. Тогда положительный квадратный корень из этой величины есть $\beta > 0$. В случае, когда скорость звука равна скорости света: $\beta = 1$, $D_{\mu\nu} = \alpha^2 g_{\mu\nu}$ и $L = \alpha^2 q_{,\mu} q^{,\mu} / 2$. Здесь и далее индекс «S» скалярной моды и фоновый индекс «F» по возможности опускаем.

Введем канонически сопряженный к q скаляр p :

$$p = p(x^\mu) \equiv \frac{\partial L(q)}{\partial \dot{q}} = \alpha^2 \dot{q} = \frac{\delta \varepsilon_c}{H} = \frac{\varepsilon + p}{H} \delta \varepsilon, \quad (8.1.3)$$

где $\dot{q} \equiv u^\mu q_{,\mu}$, $\delta \varepsilon_c$ — сопутствующее возмущение плотности (4.5.12). Квантование вводится с помощью стандартного одновременного коммутационного соотношения между операторами q и p :

$$[q(t, \mathbf{x}) p(t, \mathbf{x}')] \equiv qp - pq = i \frac{\delta^{(3)}(\mathbf{x} - \mathbf{x}')}{\sqrt{-g}}, \quad (8.1.4)$$

которое является математическим выражением принципа неопределенности Гейзенберга. Мы видим, что q выступает в роли координаты, а p — импульса. Процедура квантования не зависит от выбора гиперповерхности одновременности Σ_μ , однако для простоты предполагаем фридмановскую гиперповерхность $t = \text{const}$ с мерой на ней $d\Sigma_\mu = t_{,\mu} a^3 dx$, где $\sqrt{-g} = a^3$, $\delta^{(3)}(\mathbf{x})$ — трехмерная дельта-функция Дирака.

Восстанавливая размерные константы, перепишем соотношение (8.1.4) с учетом (8.1.3):

$$\left[\frac{q}{H} \delta \varepsilon_c \right] = i \hbar \frac{\delta^{(3)}(\mathbf{x} - \mathbf{x}')}{a^3}. \quad (8.1.5)$$

Коммутатор (8.1.5) напоминает квантование фононов в гидродинамике. В пределе коротких длин волн квантовая теория S-моды

переходит в обычную теорию фононов негравитирующей однородной среды в пространстве Минковского:

$$[v \delta \rho] = i\hbar \delta^{(3)}(\mathbf{x} - \mathbf{x}'). \quad (8.1.6)$$

Используя эту аналогию, мы будем называть кванты поля q гравитирующими фононами или просто *фононами*.

Полный гамильтониан системы (8.1.1)–(8.1.3) имеет стандартный вид интеграла по 3-мерному объему:

$$\mathcal{H} = a^3 \int \mathcal{E} d\mathbf{x}, \quad (8.1.7)$$

где \mathcal{E} — локальная энергия поля q (см. (7.1.3)–(7.1.7)):

$$\mathcal{E} = \frac{\alpha^2}{2} \left[\dot{q}^2 + \beta^2 \left(\frac{\nabla q}{a} \right)^2 \right] = \frac{\varepsilon + p}{2} (\beta^2 \delta_\varepsilon^2 + \mathbf{v}^2). \quad (8.1.8)$$

Данная процедура квантования, основанная на одновременном коммутационном соотношении, вполне корректна в квазиклассическом пределе, поскольку *однородная* часть пространства-времени остается классической и понятие одновременных событий в нем строго определено. Квантуется только *неоднородная* часть пространства-времени, предполагаемая малой, что и является условием квазиклассического предела квантовой гравитации.

8.2. Элементарные осцилляторы

Для представления вторичного квантования поля плотности рассмотрим гильбертово пространство всех классических решений уравнения (8.1.2) ¹⁾. Определим меру в этом пространстве скалярным произведением (числом), которое ставится в соответствие любой паре решений q_1 и q_2 с помощью интегрирования тока — билинейной формы J_{12}^μ , составленной из этих решений, — на пространственноподобной гиперповерхности:

$$(q_1, q_2) = \int_{\Sigma} J_{12}^\mu d\Sigma_\mu, \quad (8.2.1)$$

где

$$J_{12}^\mu = iD^{\mu\nu} (q_1^* q_{2,\nu} - q_{1,\nu}^* q_2), \quad J_{12;\mu}^\mu = 0, \quad (8.2.2)$$

¹⁾ Это уравнение действительное, но его решения комплексные.

звездочкой обозначено комплексное сопряжение. Интеграл не зависит от выбора гиперповерхности Σ_μ в силу закона сохранения тока.

Выберем ортонормированный базис в гильбертовом пространстве в виде *плоских волн*:

$$(\nu_{k_1} e^{i\mathbf{k}_1 \mathbf{x}}, \nu_{k_2} e^{i\mathbf{k}_2 \mathbf{x}}) = (2\pi)^3 \delta^{(3)}(\mathbf{k}_1 - \mathbf{k}_2), \quad (8.2.3)$$

и разложим квантовое поле в интеграл по плоским волнам:

$$q = \int \left(a_{\mathbf{k}} \nu_{\mathbf{k}} e^{i\mathbf{k}\mathbf{x}} + a_{\mathbf{k}}^\dagger \nu_{\mathbf{k}}^* e^{-i\mathbf{k}\mathbf{x}} \right) \frac{d\mathbf{k}}{(2\pi)^{3/2}}, \quad (8.2.4)$$

где интегрирование ведется по всем волновым векторам \mathbf{k} ,

$$\nu_{\mathbf{k}} = \nu_{\mathbf{k}}(\eta) \equiv \frac{\bar{\nu}_{\mathbf{k}}}{\bar{\alpha}}, \quad \bar{\alpha} \equiv \alpha a, \quad (8.2.5)$$

крест означает эрмитово сопряжение.

Представление (8.2.4) является полным в классе равномерно-ограниченных в пространстве функций $|q| < \text{const}$, поскольку оно сводится с помощью простого переобозначения к стандартному преобразованию Фурье между функциями q и $q_{\mathbf{k}}$, заданными соответственно в координатном \mathbf{x} и импульсном \mathbf{k} евклидовых пространствах:

$$q = \int q_{\mathbf{k}} e^{i\mathbf{k}\mathbf{x}} \frac{d\mathbf{k}}{(2\pi)^{3/2}}, \quad q_{\mathbf{k}} = a_{\mathbf{k}} \nu_{\mathbf{k}} + a_{-\mathbf{k}}^\dagger \nu_{\mathbf{k}}^*. \quad (8.2.6)$$

Фурье-гармоники обладают естественной симметрией в силу действительности поля q и называются *элементарными осцилляторами*, поскольку зависят только от времени:

$$q_{\mathbf{k}} = q_{\mathbf{k}}(\eta) = q_{-\mathbf{k}}^\dagger.$$

Коммутационное соотношение (8.1.4) между операторами q и p в позиционном пространстве переносится на коммутаторы операторов уничтожения $a_{\mathbf{k}}$ и рождения $a_{\mathbf{k}}^\dagger$ частиц (фононов) в импульсном пространстве:

$$\left[a_{\mathbf{k}_1} a_{\mathbf{k}_2}^\dagger \right] = \delta^{(3)}(\mathbf{k}_1 - \mathbf{k}_2), \quad [a_{\mathbf{k}_1} a_{\mathbf{k}_2}] = 0. \quad (8.2.7)$$

Очевидно, частицы тождественны своим античастицам. *Постоянные* операторы $a_{\mathbf{k}}$ и $a_{\mathbf{k}}^\dagger$ воздействуют на состояния поля, которые также не зависят от координат (представление Гейзенберга).

Эволюция поля q описывается классическими функциями (8.2.5), удовлетворяющими условиям нормировки (8.2.3):

$$\bar{v}_k'' + (\omega^2 - U) \bar{v}_k = 0, \quad (8.2.8)$$

$$\bar{v}_k \bar{v}_k^{*'} - \bar{v}_k^{*'} \bar{v}_k' = i, \quad (8.2.9)$$

где

$$\omega = \beta k \quad \text{и} \quad U = \frac{\bar{\alpha}''}{\bar{\alpha}}$$

— частота и эффективный потенциал \mathbf{k} -осциллятора соответственно. Гармоники с разными волновыми векторами эволюционируют *независимо* друг от друга и ощущают только динамическое влияние фонового гравитационного поля, в котором они распространяются, — масштабного фактора и ряда его производных.

Вернемся к общему разложению поля q по плоским волнам и докажем, что полное действие $S[q]$ разбивается на сумму действий элементарных осцилляторов:

$$S[q] = \int S_{\mathbf{k}} d\mathbf{k}, \quad S_{\mathbf{k}} = S[q_{\mathbf{k}}]. \quad (8.2.10)$$

Действительно, подставляя (8.2.6) в (8.1.1) и интегрируя по x , получаем лагранжиан \mathbf{k} -осциллятора:

$$S_{\mathbf{k}} = \int L_{\mathbf{k}} d\eta, \quad L_{\mathbf{k}} = \frac{\bar{\alpha}^2}{2} (|q_{\mathbf{k}}'|^2 - \omega^2 |q_{\mathbf{k}}|^2), \quad (8.2.11)$$

где обозначение $|X|^2 \equiv X X^\dagger = X^\dagger X$ относится к любому самокоммутирующему оператору $[X X^\dagger] = 0$. Это выражение выглядит как лагранжиан обычного математического маятника в пространстве Минковского, находящегося под воздействием внешнего переменного поля $\bar{\alpha} = \bar{\alpha}(\eta)$. Функция $\bar{\alpha}(\eta)$ несет информацию о фоновой геометрии Фридмана. Очевидно, речь идет о двух эквивалентных формулировках одного и того же явления, что позволяет нам воспользоваться известными ранее результатами для параметрического маятника в плоском пространстве-времени и применить их к Вселенной.

По своему построению поля $q_{\mathbf{k}}$ являются *комплексными* операторами, обладающими определенной симметрией. Более естественно было бы разбить поле плотности на независимые *действительные* осцилляторы.

Это можно сделать, рассматривая вместо каждого $q_{\mathbf{k}}$ пару эрмитовых полей $q_{\mathbf{k}+}$ и $q_{\mathbf{k}-}$:

$$q_n = q_n^\dagger = \frac{q_{\mathbf{k}+} + sq_{\mathbf{k}-}}{\sqrt{2s}}, \quad (8.2.12)$$

где $s = +, -$ есть знаковый индекс, а $n = \mathbf{k}s$ обозначает совокупность чисел \mathbf{k} и s . Обратное преобразование имеет вид:

$$q_{\mathbf{k}} = \frac{q_{\mathbf{k}+} + iq_{\mathbf{k}-}}{\sqrt{2}}.$$

Подставляя его в (8.2.11), получаем искомое разбиение и лагранжианы n -осцилляторов:

$$L_{\mathbf{k}} = \frac{1}{2} \sum_s L_n, \quad L_n = \frac{\bar{\alpha}^2}{2} (q_n'^2 - \omega^2 q_n^2). \quad (8.2.13)$$

Независимые друг от друга n -осцилляторы перемешивают гармоники с противоположными волновыми векторами \mathbf{k} и $-\mathbf{k}$. Поля $q_{\mathbf{k}+}$ и $q_{\mathbf{k}-}$ есть симметричная и антисимметричная комбинации $q_{\mathbf{k}}$ и $q_{-\mathbf{k}}$. Поэтому $L_{\mathbf{k}+} + L_{\mathbf{k}-} = L_{\mathbf{k}} + L_{-\mathbf{k}}$.

Таким образом, квантовое поле q в геометрии Фридмана представляет собой ансамбль действительных элементарных осцилляторов с частотами ω и лагранжианами L_n :

$$q_n = a_n \nu_{\mathbf{k}} + a_n^\dagger \nu_{\mathbf{k}}^*, \quad a_n = \frac{a_{\mathbf{k}} + sa_{-\mathbf{k}}}{\sqrt{2s}}. \quad (8.2.14)$$

Ассоциированный с (8.2.13) гамильтониан имеет вид:

$$\mathcal{H}_n = \mathcal{H}(q_n, p_n) = \frac{1}{2} \left(\frac{p_n^2}{\bar{\alpha}^2} + \omega^2 \bar{\alpha}^2 q_n^2 \right), \quad (8.2.15)$$

где $p_n = \frac{\partial L_n}{\partial q_n'} = \bar{\alpha}^2 q_n'$ — импульс n -осциллятора. Нетрудно видеть, что полный гамильтониан (8.1.7) также разбивается на сумму парциальных гамильтонианов (8.2.15).

Для упрощения записи часто используют переменные, нормированные на пространственный объем $V \equiv \int d\mathbf{x}$:

$$\frac{(2\pi)^{3/2}}{\sqrt{V}} q_n \rightarrow q_n, \quad \frac{(2\pi)^3}{V} L_n \rightarrow L_n, \quad \frac{V}{(2\pi)^3} \int d\mathbf{k} \rightarrow \sum_n,$$

что позволяет заменить интегральное действие (8.1.1) на сумму действий n осцилляторов по элементарным ячейкам фазового

пространства (ср. (8.2.10), (8.2.11)):

$$S[q] = \sum_n S_n, \quad S_n = S[q_n] = \int L_n d\eta. \quad (8.2.16)$$

Подчеркнем, что k - и n -представления эквивалентны, в зависимости от задачи можно пользоваться любым из них.

В канонических переменных

$$\hat{q}_n = \bar{\alpha} \sqrt{\omega} q_n, \quad \hat{p}_n = \frac{\partial L_n}{\partial \hat{q}'_n} = \frac{p_n}{\bar{\alpha} \sqrt{\omega}} \quad (8.2.17)$$

лагранжиан (8.2.13) принимает стандартный вид:

$$L_n = \frac{1}{2} \omega (\hat{p}_n^2 - \hat{q}_n^2), \quad [\hat{q}_n \hat{p}_n] = i\hbar. \quad (8.2.18)$$

Однако гамильтониан осциллятора оказывается деформированным:

$$\mathcal{H}_n = \frac{1}{2} \omega \left(\hat{p}_n^2 + \frac{\hbar}{\omega} \hat{q}_n \hat{p}_n + \frac{\hbar}{\omega} \hat{p}_n \hat{q}_n + \hat{q}_n^2 \right), \quad (8.2.19)$$

где

$$\hbar = \frac{(\bar{\alpha} \sqrt{\beta})'}{\bar{\alpha} \sqrt{\beta}}.$$

Фактор \hbar/ω пропорционален отношению длины волны к звуковому горизонту. Члены, пропорциональные \hbar , описывают приливное гравитационное воздействие, испытываемое n -осциллятором в расширяющейся Вселенной. В явном виде это прослеживается из связи канонического импульса с координатной производной (8.2.17):

$$\omega \hat{p}_n = \hat{q}'_n - \hbar \hat{q}_n.$$

Лагранжиан (8.2.18) сохраняет свою каноническую форму при преобразованиях Лоренца:

$$\tilde{q}_n = \frac{\hat{q}_n + \nu \hat{p}_n}{\sqrt{1 - \nu^2}}, \quad \tilde{p}_n = \frac{\hat{p}_n + \nu \hat{q}_n}{\sqrt{1 - \nu^2}}, \quad (8.2.20)$$

с помощью которых можно привести гамильтониан (8.2.19) к диагональному виду. Выбирая функцию времени

$$\nu = \frac{\hbar}{\omega + \sqrt{\omega^2 - \hbar^2}},$$

получаем:

$$\mathcal{H}_n = \frac{1}{2} \sqrt{\omega^2 - \hbar^2} (\tilde{p}_n^2 + \tilde{q}_n^2). \quad (8.2.21)$$

Операторы \tilde{q}_n и \tilde{p}_n эрмитовы тогда и только тогда, если

$$|\hbar| < \omega, \quad (8.2.22)$$

что предполагает причинную (микроскопическую) зону эволюции. Параметр ν в этом случае характеризует хаббловскую скорость в масштабе волны. При $|\hbar| < \omega$ гамильтониан элементарного осциллятора положительно определен и у него существует *единственное* основное энергетическое состояние.

8.3. Начальные условия

Эволюция элементарного осциллятора (8.2.8) зависит от его эффективной частоты, определяемой соотношением между скалярами ω^2 и U .

При $|U| < \omega^2$ осциллятор находится в *волновой* зоне и совершает свободные адиабатические колебания, затухающие вследствие расширения Вселенной:

$$\nu_k = \frac{\exp(-i\kappa)}{\bar{\alpha}\sqrt{2\omega}}. \quad (8.3.1)$$

Фазовая функция

$$\kappa = \kappa(\eta) \equiv \int \omega d\eta = - \int \frac{\omega}{H} dN,$$

монотонно растущая со временем, определена с точностью до произвольной аддитивной постоянной κ_0 (начальной фазы). Выбор функции ν_k в виде *бегущей* волны соответствует представлению частиц на стадии (8.3.1). Здесь и далее предполагается адиабатичность β при переходе от стадии (8.3.1) к стадии (8.3.2)¹⁾.

При $\omega^2 < |U|$ осциллятор оказывается в *параметрической* зоне и решение в главном порядке имеет вид:

$$\nu_k = \frac{c - i}{C\sqrt{2k}}, \quad (8.3.2)$$

где

$$c = c(\eta) \equiv -kC^2 \int \frac{d\eta}{\bar{\alpha}^2} = kC^2 \left(\frac{1}{\bar{\alpha}\bar{\alpha}'} + \int \frac{U}{\bar{\alpha}'^2} d\eta \right). \quad (8.3.3)$$

¹⁾ Это относится только к S-моду, поскольку для T-мод $\beta = 1$. Для последующей эволюции адиабатичность не предполагается. В отсутствие частиц бегущую волну (8.3.1) называют нулевым или вакуумным колебанием поля.

С учетом ограничения (8.2.9) комплексная функция ν_k задается тремя действительными постоянными. Одна из них определяет общий фазовый фактор $e^{-i\kappa_0}$ решения (8.3.2), который здесь для простоты опущен. Две других — это константа интегрирования функции c и амплитудный коэффициент $C = \text{const}$. Для нахождения всех трех констант параметрической зоны необходимо решить уравнение (8.2.8) с начальным условием (8.3.1).

В рассматриваемой области параметров ($\gamma < 3$) интегралы (8.3.3) сходятся на верхнем пределе, и осциллятор застывает с течением времени:

$$\begin{aligned} c &\rightarrow c_k = \text{const}, \\ \sqrt{2k} \nu_k &\rightarrow \frac{c_k - i}{C} \equiv \frac{2\pi}{k} q_k e^{-i\psi_k}. \end{aligned} \quad (8.3.4)$$

«Замороженное» поле соответствует растущей ветви общего решения, а действительные постоянные q_k и ψ_k — это его амплитуда и фаза¹⁾. Ансамбль этих постоянных для всех элементарных осцилляторов описывает суммарное поле $q(\mathbf{x})$ космологических возмущений.

Амплитуда

$$q_k = \frac{k \sqrt{1 + c_k^2}}{2\pi C}$$

содержит всю спектральную информацию о структуре ранней Вселенной. Константа c_k является интегральной характеристикой истории осциллятора в ранней Вселенной, включая сам переход из (8.3.1) в (8.3.2) и начальный период параметрической стадии, где насыщается интеграл (8.3.3). Постоянная $C = C(k)$ задана с момента возникновения параметрической зоны и сохраняет свое значение в ходе последующей эволюции при $\bar{H} > \omega$ (интеграл движения).

С наступлением радиационно-доминированной стадии космологического расширения k -осциллятор попадает в волновую зону ($U = 0$), на которой

$$\gamma = 2, \quad \beta_S = \frac{1}{\sqrt{3}} \quad \text{и} \quad \bar{\alpha}' = \text{const} \quad (8.3.5)$$

— это инварианты эволюции. Константа $\bar{\alpha}'$ связана с полной энтропией Вселенной. Определяя параметр

$$k_{\text{rad}} = \epsilon_{\text{rad}}^{1/4} a = \text{const} \quad (8.3.6)$$

¹⁾ Начальная фаза κ_0 остается в качестве аддитивного слагаемого ψ_k .

по полной плотности энергии излучения ϵ_{rad} в начале радиационно-доминированной стадии, получаем с учетом уравнения Фридмана:

$$\bar{\alpha}'_S = 2k_{\text{rad}}^2, \quad \bar{\alpha}'_T = \frac{k_{\text{rad}}^2}{\sqrt{3}}. \quad (8.3.7)$$

Энтропийная константа слабо зависит от числа безмассовых степеней свободы в ранней Вселенной и близка к частоте реликтового фотона:

$$k_{\text{rad}} \sim \text{см}^{-1}. \quad (8.3.8)$$

На радиационно-доминированной стадии осциллятор описывает преимущественно *стоячую* волну (ср. (8.3.1)), так как в ходе эволюции выделяется одна из временных фаз общего решения (8.2.8), и функция (8.3.2) приобретает вид:

$$\nu_k = \frac{c_k - i}{C\sqrt{2k}} + \frac{1}{R\bar{\alpha}\sqrt{2\omega}}, \quad (8.3.9)$$

где последнее слагаемое описывает падающую ветвь решения, а ее амплитуда характеризуется постоянной

$$R = R(k) \equiv \frac{\bar{\alpha}'}{kC\sqrt{\beta}} = \text{const}$$

(при $R > 1$ падающая ветвь подавлена, см. раздел 7.2).

Напомним, что в начале радиационно-доминированной стадии масштаб космологических возмущений превосходил хаббловский размер:

$$\bar{H} \gg \omega. \quad (8.3.10)$$

Такие сверхдлинные поля всегда стационарны, и их можно рассматривать в классическом пределе. В результате эволюции мы получаем классические поля с реальным *неоднородным* распределением плотности энергии, тогда как затравочные возмущения могли быть квантовыми, не нарушающими фоновое *однородное* распределение плотности и гравитационного потенциала.

Итак, для реализации параметрического эффекта усиления необходимо выполнение двух условий: наличие затравочного возмущения и переход из адиабатической в параметрическую зоны эволюции. Если изначально осциллятор не возбужден, то и общее решение равно нулю.

Затравочные возмущения могли быть любыми в зависимости от постановки задачи и начальных условий. Однако при любом подходе существует минимально возможный уровень возбужде-

ния осцилляторов, в котором изначально *отсутствуют* частицы поля q :

$$a_{\mathbf{k}}|0\rangle = 0, \quad (8.3.11)$$

или, эквивалентно,

$$a_n|0\rangle = 0,$$

где операторы уничтожения соответствуют представлению частиц (8.3.1). Состояние $|0\rangle$ не изменяется с течением времени, но оно соответствует физическому вакууму поля q только в зоне (8.3.1). В ходе дальнейшей эволюции это состояние описывает уже неминимальные возбуждения с конечным числом частиц.

Ниже мы предполагаем существование волновой зоны и основного состояния $|0\rangle$ для всех элементарных осцилляторов. Предположение о *начальной адиабатической зоне* составляет главную гипотезу ранней Вселенной. Реализация этой гипотезы во фридмановской геометрии делает выбор состояния (8.3.11) *единственным*, поскольку при $U = 0$ поле q конформно-инвариантно свободному осциллятору в пространстве Минковского, а такой осциллятор, как известно из квантовой механики, имеет единственное состояние, в котором отсутствуют частицы.

В вакуумном состоянии элементарные осцилляторы не коррелированы: их начальные фазы случайны и не зависят друг от друга. В параметрической зоне осцилляторы описывают реальные (надвакуумные) космологические возмущения. Эволюционный переход каждого из осцилляторов в параметрическую зону регулируется внешним фактором — темпом расширения Вселенной — и никак не зависит от фазы его собственных колебаний. Колеблющийся осциллятор застывает в том положении, в котором его застал вдруг появившийся потенциал U , поэтому родившееся суммарное поле элементарных осцилляторов $q(\mathbf{x})$ распределено в *пространстве* случайным образом¹⁾. Более того, оно носит гауссовый характер в силу того, что вакуумные затравочные возмущения по своей природе изначально гауссовы и это качество передается линейным образом результирующему распределению поля $q(\mathbf{x})$.

Среднеквадратичное ожидание поля в состоянии, описываемом (8.3.11), имеет вид:

$$\langle q^2 \rangle = \int_0^\infty q_k^2 \frac{dk}{k}, \quad q_k = \frac{k\sqrt{1+c_k^2}}{2\pi C}, \quad (8.3.12)$$

¹⁾ Иначе дело обстоит с *временной* фазой (подробнее см. раздел 8.6).

где q_k — дисперсия возмущений с длиной волны $2\pi/k$ (8.3.4). Важно подчеркнуть, что характер космологического расширения после перехода в параметрическую зону — поведение масштабного фактора, а также свойства и процессы преобразования материи (рождение частиц, разогрев вещества, трансформация к радиационно-доминированной стадии и пр.) — не оказывают никакого влияния на крупномасштабный спектр q_k . Он зависит лишь от космологических параметров в момент перехода k -осциллятора из начальной волновой в параметрическую зону эволюции.

Для определения спектральных и фазовых характеристик космологических возмущений следует решить соответствующее дифференциальное уравнение (8.2.8) с начальным условием (8.3.1) и найти константы C и c_k как функции волнового числа k . Однако если область перехода достаточно мала, можно поступить проще и прямо связать хаббловские частоты в момент перехода с волновым числом осциллятора, разрешая алгебраическое уравнение:

$$|U| = \omega^2. \quad (8.3.13)$$

Анализ этого уравнения показывает, какие именно космологические параметры индуцируют параметрическое усиление возмущений и какое соотношение между частотами необходимо для реализации самого эффекта.

Выпишем в явном виде эффективный потенциал элементарных осцилляторов (см. (8.2.8)):

$$\frac{U}{\bar{H}^2} = \begin{cases} 2 - \gamma = 1 - \bar{\gamma}, \\ (2 - \gamma + \Gamma)(1 + \Gamma) + \frac{\Gamma'}{\bar{H}} = (1 - \bar{\Gamma})(1 + \Gamma), \end{cases}$$

$$\Gamma \equiv \frac{\alpha'}{\alpha\bar{H}}, \quad \bar{\Gamma} \equiv -\frac{(\alpha(1 + \Gamma)\bar{H})'}{\alpha(1 + \Gamma)\bar{H}^2}, \quad (8.3.14)$$

верхняя строка скобки относится к моде T, а нижняя — к S.

Наиболее просто устроены T-осцилляторы: для них потенциал зависит лишь от двух параметров \bar{H} и γ . В рассматриваемой области значений $-1 \lesssim \gamma < 3$ получаем следующее ограничение на частоты в переходной области:

$$|U| = \omega^2: \quad \bar{H} \gtrsim \omega. \quad (8.3.15)$$

Для S-осцилляторов число параметров потенциала в два раза больше (\bar{H} , γ , γ' и γ'') и функция $1 + \Gamma$ может менять знак.

Однако в большинстве рассматриваемых моделей модуль правой части в (8.3.14) не превосходит нескольких единиц. Предполагая ограниченность $|U|/\bar{H}^2$, приходим опять к условию (8.3.15).

В зависимости от того, расширялась ли Вселенная на стадии (8.3.1) с ускорением или замедлением, начальное состояние (8.3.11) должно быть соответствующим образом определено либо в причинной ($\bar{H} < \omega$), либо в не причинной ($\bar{H} > \omega$) области пространства-времени. В микроскопической области определение вакуума как состояния поля с минимальной энергией физически осмысленно и однозначно (см. (8.2.22)). Напротив, постановка начальных условий в причинно-несвязанной области вызывает вопросы, поскольку гамильтониан (8.2.21) положительно не определен.

При совместном выполнении условий (8.3.1) и (8.2.22):

$$|U|^{1/2} < \omega, \quad |\hbar| < \omega, \quad (8.3.16)$$

состояние (8.3.11) отвечает минимуму начальной энергии возбуждения элементарных осцилляторов.

8.4. Фононы

Рассмотрим более подробно волновую зону $U \rightarrow 0$.

Важной характеристикой поля на этой стадии является *адиабатический инвариант*, определяющий число частиц в волновой зоне. Для его описания введем понятие энергии частиц на гиперповерхности $\eta = \text{const}$ в следующей форме (мы отбросили нелокальный член в виде полной дивергенции в правой части (7.2.2)):

$$\begin{aligned} \bar{\mathcal{H}} &= \int \bar{\mathcal{E}} d\mathbf{x} = \int \bar{\mathcal{E}}_{\mathbf{k}} d\mathbf{k}, \\ \bar{\mathcal{E}} &= \frac{1}{2} (\bar{q}'^2 + \beta^2 \nabla \bar{q}^2 - U \bar{q}^2), \end{aligned} \quad (8.4.1)$$

где оператор

$$\bar{\mathcal{E}}_{\mathbf{k}} = \frac{1}{2} (|\bar{q}'_{\mathbf{k}}|^2 + (\omega^2 - U) |\bar{q}_{\mathbf{k}}|^2) \quad (8.4.2)$$

представляет энергию частиц (и античастиц) поля $q_{\mathbf{k}}$.

На стадии $U = 0$ из (8.3.1) в случае $\beta = \text{const}$ имеем $\bar{\mathcal{E}}_{\mathbf{k}} = \text{const}$ и $\bar{\mathcal{H}} = \text{const}$. При адиабатическом изменении скорости β сохраняются адиабатические инварианты $\bar{\mathcal{E}}_{\mathbf{k}}/\omega$ и общий инвариант $\bar{\mathcal{H}}/\beta$, регуляризованные части которых соответствуют

операторам числа частиц полей $q_{\mathbf{k}}$:

$$\bar{\mathcal{E}}_{\mathbf{k}} = \frac{1}{2} \left(a_{\mathbf{k}}^{\dagger} a_{\mathbf{k}} + a_{-\mathbf{k}} a_{-\mathbf{k}}^{\dagger} \right), \quad (8.4.3)$$

$$N_{\mathbf{k}} = \frac{1}{2} \left(a_{\mathbf{k}}^{\dagger} a_{\mathbf{k}} + a_{-\mathbf{k}}^{\dagger} a_{-\mathbf{k}} \right), \quad (8.4.4)$$

$$\bar{\mathcal{H}}_{reg} = \int \omega N_{\mathbf{k}} d\mathbf{k} = \int \omega a_{\mathbf{k}}^{\dagger} a_{\mathbf{k}} d\mathbf{k}. \quad (8.4.5)$$

Представим адиабатический инвариант с помощью операторов стоячих волн, фиксирующих определенную временную фазу колебаний поля $q_{\mathbf{k}}$ в волновой зоне. По определению *растущие* и *падающие* операторы частиц соответственно равны

$$C_1 = \frac{a_{\mathbf{k}} - a_{-\mathbf{k}}^{\dagger}}{i\sqrt{2}} \quad \text{и} \quad C_2 = \frac{a_{\mathbf{k}} + a_{-\mathbf{k}}^{\dagger}}{\sqrt{2}}, \quad (8.4.6)$$

где $C_m = C_m(\mathbf{k}) = C_m^{\dagger}(-\mathbf{k})$, $m = 1, 2$. Подставляя эти выражения в (8.2.6) и (8.4.1), с учетом (8.3.1) получаем:

$$\sqrt{\omega} \bar{q}_{\mathbf{k}} = C_1 \sin \kappa + C_2 \cos \kappa, \quad (8.4.7)$$

$$\bar{\mathcal{E}}_{\mathbf{k}} = \frac{1}{2} \omega (|C_1|^2 + |C_2|^2). \quad (8.4.8)$$

Мы видим, что операторы временных фаз осциллятора дают пропорциональный вклад в энергию его колебаний. Важно подчеркнуть, что разделение полей по временной фазе имеет смысл только в адиабатических зонах, где осцилляторы свободны от внешнего воздействия. В общем виде такое разделение невозможно из-за взаимодействия осцилляторов с фоновой метрикой и перепутывания временных фаз.

Усредняя квадраты операторов (8.4.6) по состоянию поля q , получаем численные значения амплитуд растущих ($m = 1$) и падающих ($m = 2$) ветвей колебаний свободных осцилляторов (см. (8.2.16)):

$$c_m = \left(\frac{(2\pi)^3}{V} \langle |C_m|^2 \rangle \right)^{1/2}, \quad (8.4.9)$$

где индекс \mathbf{k} для простоты опущен (ср. (8.4.6)).

Мы будем считать, что состояние \mathbf{k} -осциллятора *сжато*, если для него $c_1 \neq c_2$ (выделенная временная фаза). При $c_1 = c_2$ \mathbf{k} -осциллятор находится в состоянии с произвольной (случайной или неопределенной) временной фазой. В частности, в вакуумном состоянии (8.3.11):

$$c_1^2 = c_2^2 = \frac{\hbar}{2} \quad (8.4.10)$$

(здесь восстановлена размерность).

Из (8.4.8) и (8.4.9) получаем *среднюю* плотность энергии фононов в заданном состоянии поля q :

$$\varepsilon_q \equiv \frac{\langle \mathcal{H} \rangle}{a^4 V} = \frac{1}{(2\pi a)^3} \int E_k \left(n_{\mathbf{k}} + \frac{1}{2} \right) d\mathbf{k}, \quad (8.4.11)$$

$$\begin{aligned} \langle a_{\mathbf{k}}^\dagger a_{\mathbf{k}'} \rangle &= n_{\mathbf{k}} \delta(\mathbf{k} - \mathbf{k}'), \\ n_{\mathbf{k}} &= \frac{1}{2} (c_1^2 + c_2^2 - 1), \end{aligned} \quad (8.4.12)$$

где $E_k = \omega/a$ — энергия одного фонона, $n_{\mathbf{k}}$ — числа заполнения фононов с импульсом \mathbf{k} .

В вакуумном состоянии (8.4.10) частицы поля отсутствуют: $n_{\mathbf{k}} = 0$. Член $1/2$ под интегралом в (8.4.11) описывает энергию нулевых (вакуумных) колебаний¹⁾.

Рассмотрим фононные представления с точки зрения полного гамильтониана S -поля (8.2.15).

Канонические операторы на стадии (8.3.1) задаются выражениями (8.2.18):

$$\begin{aligned} \hat{q}_n &= \frac{\hat{a}_n e^{-i\kappa} + \hat{a}_n^\dagger e^{i\kappa}}{\sqrt{2}}, \\ \hat{p}_n &= \frac{\hat{a}_n e^{-i\kappa} - \hat{a}_n^\dagger e^{i\kappa}}{i\sqrt{2}} - \frac{\hbar}{\omega} \hat{q}_n, \end{aligned} \quad (8.4.13)$$

а их среднеквадратичные ожидания в вакуумном состоянии соответственно равны:

$$\langle \hat{q}_n^2 \rangle = \frac{\hbar}{2}, \quad \langle \hat{p}_n^2 \rangle = \frac{\hbar}{2} \left(1 + \frac{\hbar^2}{\omega^2} \right). \quad (8.4.14)$$

В микроскопической области (8.2.22) период колебаний элементарного осциллятора меньше хаббловского времени и гамильтониан системы (8.2.19) положительно определен. При выполнении обоих условий:

$$|U|^{1/2} \ll \omega \quad \text{и} \quad |\hbar| \ll \omega, \quad (8.4.15)$$

квантово-механические операторы \hat{q}_n и \hat{p}_n описывают *свободный* гармонический осциллятор, не подверженный внешнему гравитационному влиянию со стороны расширяющейся Вселенной. В адиабатической зоне, но при конечных \hbar , элементарный осциллятор испытывает приливное воздействие фонового гравитационного потенциала (см. (8.4.14)): фононный вакуум поляризован, однако число частиц в нем сохраняется равным нулю,

¹⁾ При регуляризации оператора (8.4.5) мы этот член отбросили.

так как они не рождаются в адиабатической зоне (конформная инвариантность).

8.5. Гравитоны

Аналогично скалярной моде строится квантовая теория тензорных возмущений кривизны в геометрии Фридмана. Напомним вкратце ее основные этапы и поясним физическое сходство S- и T-мод космологических возмущений.

Как и в моде S, квантуются только физические степени свободы. Действие и классическое уравнение движения поперечно-бесследового поля метрического тензора \mathfrak{S}_{ij} имеют вид (см. (5.3.6), (7.1.14)) ¹⁾:

$$S [\mathfrak{S}_i^j] = \int L_T \sqrt{-g} d^4x, \quad (8.5.1)$$

$$L_T = \frac{\alpha_T^2}{2} \mathfrak{S}_{i,\mu}^j \mathfrak{S}_j^{i,\mu}, \quad \alpha_T^2 = \frac{1}{8\pi G} = \text{const},$$

$$\ddot{\mathfrak{S}}_{ij} + 3H\dot{\mathfrak{S}}_{ij} - a^{-2}\Delta\mathfrak{S}_{ij} = 0. \quad (8.5.2)$$

В представлении вторичного квантования имеем:

$$\mathfrak{S}_{ij} = \sum_{\lambda} \int \left(\mathcal{P}_{ij}^{\lambda} a_{\mathbf{k}\lambda} \nu_{\mathbf{k}\lambda} e^{i\mathbf{k}\mathbf{x}} + \mathcal{P}_{ij}^{\lambda*} a_{\mathbf{k}\lambda}^{\dagger} \nu_{\mathbf{k}\lambda}^* e^{-i\mathbf{k}\mathbf{x}} \right) \frac{d\mathbf{k}}{(2\pi)^{3/2}}.$$

Постоянные поляризаационные тензоры \mathbf{k} -волн

$$\mathcal{P}_{ij}^{\lambda} = \mathcal{P}_{ij}^{\lambda}(\mathbf{k}) = \mathcal{P}_{ij}^{\lambda*}(-\mathbf{k}) \quad (8.5.3)$$

имеют две степени свободы $\lambda = \oplus, \otimes$ и удовлетворяют условиям нормировки:

$$\mathcal{P}_i^{\lambda i}(\mathbf{k}) = k^i \mathcal{P}_{ij}^{\lambda}(\mathbf{k}) = 0, \quad \mathcal{P}_{ij}^{\lambda_1*}(\mathbf{k}) \mathcal{P}^{\lambda_2 ij}(\mathbf{k}) = \delta_{\lambda_1 \lambda_2}.$$

При ориентации оси x^3 вдоль вектора \mathbf{k} и произвольном позиционном угле в ортогональной плоскости (но при сохранении общей право- или левосторонней ориентации всех осей) ненулевые компоненты $\mathcal{P}_{ij}^{\lambda}$ могут быть выбраны равными $\mathcal{P}_{11}^{\oplus} = -\mathcal{P}_{22}^{\oplus} = \mathcal{P}_{12}^{\otimes} = \mathcal{P}_{21}^{\otimes} = 1/\sqrt{2}$.

¹⁾ Напомним, что манипуляции с латинскими индексами производятся с помощью единичного тензора δ_{ij} , а с греческими — с помощью метрического тензора Фридмана.

Для ненулевых коммутаторов операторов рождения и уничтожения гравитонов с импульсом \mathbf{k} и поляризацией λ получаем:

$$\left[a_{\mathbf{k}_1 \lambda_1}, a_{\mathbf{k}_2 \lambda_2}^\dagger \right] = \delta_{\lambda_1 \lambda_2} \delta^{(3)}(\mathbf{k}_1 - \mathbf{k}_2). \quad (8.5.4)$$

Классические функции времени

$$\nu_{\mathbf{k}\lambda} = \nu_{\mathbf{k}\lambda}(\eta) \equiv \frac{\bar{\nu}_{\mathbf{k}\lambda}}{\bar{\alpha}_T}, \quad \bar{\alpha}_T \equiv \alpha_T a \quad (8.5.5)$$

описывают эволюцию \mathbf{k} -осцилляторов в каждой из поляризаций λ и подчинены тем же уравнениям (8.2.8) и условиям (8.2.9), что и для S-моды, с той лишь разницей, что в них следует заменить β на единицу, а U_S на U_T . В силу постоянства α_T : $U_T = a''/a$. Таким образом, можно сказать, что тензорное поле более фундаментально, чем скалярное, поскольку его связь с фоном зависит от меньшего числа производных масштабного фактора.

Покажем, что аналогия в описании мод T и S глубже, чем кажется на первый взгляд.

Из (8.5.1)–(8.5.5) следует, что фурье-гармоники

$$q_{\mathbf{k}\lambda} = q_{\mathbf{k}\lambda}(\eta) = a_{\mathbf{k}\lambda} \nu_{\mathbf{k}\lambda} + a_{-\mathbf{k}\lambda}^\dagger \nu_{\mathbf{k}\lambda}^* \quad (8.5.6)$$

имеют форму скаляров (ср. (8.2.6)):

$$\mathfrak{S}_{ij} = \sum_{\lambda} \int \mathcal{P}_{ij}^{\lambda} q_{\mathbf{k}\lambda} e^{i\mathbf{k}\mathbf{x}} \frac{d\mathbf{k}}{(2\pi)^{3/2}}, \quad (8.5.7)$$

представляющих полное действие (8.5.1) в виде суммы действий элементарных осцилляторов (8.5.6), которых, заметим, вдвое больше, чем у S-моды:

$$S[q_{\mathbf{k}\lambda}] = \sum_{\lambda} \int L_{\mathbf{k}\lambda} d\eta, \quad (8.5.8)$$

$$L_{\mathbf{k}\lambda} = \frac{\bar{\alpha}_T^2}{2} (|q'_{\mathbf{k}\lambda}|^2 - k^2 |q_{\mathbf{k}\lambda}|^2).$$

Точно такое же лагранжево представление справедливо для *двух* пробных безмассовых действительных *скалярных* полей $q_{\lambda} = q_{\lambda}(t, \mathbf{x})$, минимальным образом связанных с фоновым пространством-временем:

$$q_{\lambda} = \int \left(a_{\mathbf{k}\lambda} \nu_{\mathbf{k}\lambda} e^{i\mathbf{k}\mathbf{x}} + a_{\mathbf{k}\lambda}^\dagger \nu_{\mathbf{k}\lambda}^* e^{-i\mathbf{k}\mathbf{x}} \right) \frac{d\mathbf{k}}{(2\pi)^{3/2}}, \quad (8.5.9)$$

$$S[q_{\lambda}] = \int L_{\lambda} \sqrt{-g} d^4x, \quad L_{\lambda} = \frac{\alpha_T^2}{2} q_{\lambda, \mu} q_{\lambda}^{\prime \mu}, \quad (8.5.10)$$

что проверяется прямой подстановкой (8.5.9) в (8.5.10) и интегрированием по \mathbf{x} .

Таким образом, квантование гравитационных волн в модели Фридмана эквивалентно квантованию двух действительных скалярных полей q_λ с действиями (8.5.10):

$$S [\mathfrak{G}_i^j] = S [q_\oplus] + S [q_\otimes]. \quad (8.5.11)$$

Аналогично, вариация Т-поля в состоянии (8.3.11) равна сумме вариаций скаляров q_λ , равных между собой:

$$\langle \mathfrak{G}_{ij} \mathfrak{G}^{ij} \rangle = \langle q_\oplus^2 \rangle + \langle q_\otimes^2 \rangle = 2 \langle q_\lambda^2 \rangle. \quad (8.5.12)$$

По этой причине при исследовании параметрического эффекта мы акцентируем внимание на скалярных космологических возмущениях. Все полученные результаты переносятся на каждую поляризацию Т-моды при замене β на единицу и α_S на α_T .

8.6. Фазовый портрет структуры

Эффект параметрического рождения возмущений кривизны удобно исследовать в постановке задачи рассеяния, поскольку имеются хорошо определенные начальная (*in*) и конечная (*out*) волновые зоны, в которых можно пересчитать и идентифицировать частицы соответствующих полей.

В ранней Вселенной все элементарные осцилляторы независимо от их частоты попадают в *out*-зону при $a > a_2$ (рис. 8.1). Это радиационно-доминированная стадия, наступившая в результате разогрева и термализации вещества. Она характеризуется высокой хаббловской частотой по сравнению с про-

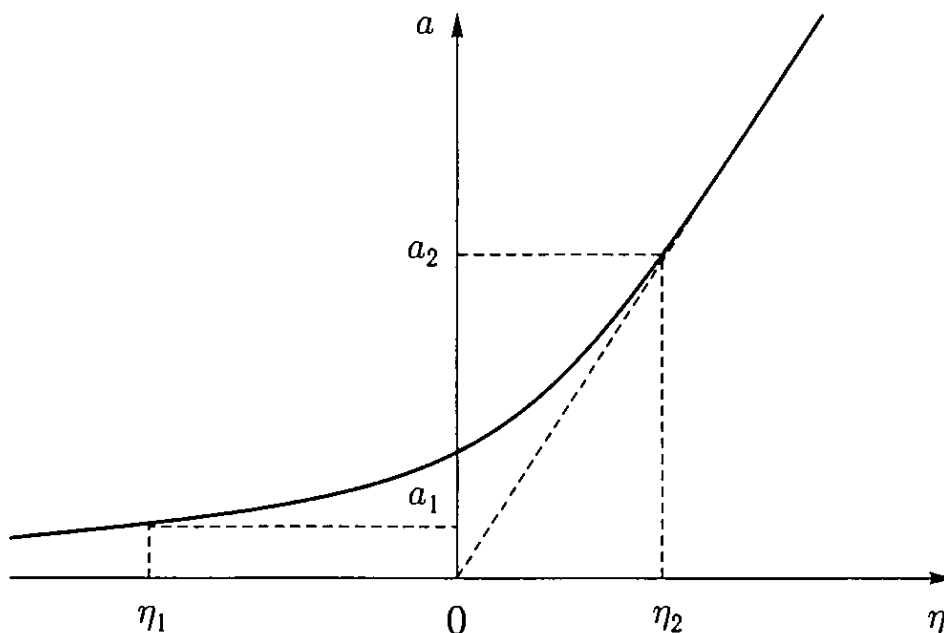


Рис. 8.1. Эволюция масштабного фактора в ранней Вселенной

странственной частотой элементарных осцилляторов: $\bar{H}_2 \gg k$ (см. (8.3.5)–(8.3.10)).

В более раннюю эпоху k -осцилляторы находились в основном состоянии в своих in -зонах (см. (8.3.1), (8.3.11)). Важным моментом в жизни каждого из них был переход из начальной волновой в параметрическую зону при $a_1 < a_2$ и $\bar{H}_1 \gtrsim \omega$ (см. (8.3.15)). Его наступление зависит, вообще говоря, от волнового числа k (в отличие от момента a_2): $a_1 = a_1(\mathbf{k})$.

О рождении полей в промежуточной области $a_1 < a < a_2$ мы можем судить, сравнивая между собой характеристики частиц в начальной и конечной волновых зонах.

Выберем аддитивную постоянную конформного времени таким образом, чтобы в начале out -стадии масштабный фактор был пропорционален η (рис. 8.1):

$$\bar{\alpha}(> \eta_2) = \bar{\alpha}'_2 \eta, \quad \mathfrak{h}(> \eta_2) = \bar{H} = \frac{1}{\eta}, \quad (8.6.1)$$

где $\bar{\alpha}'_2(> \eta_2) = \bar{\alpha}'_2 = \text{const}$ (см. (8.3.7)). Обозначим операторы уничтожения и рождения частиц в out -представлении через $b_{\mathbf{k}}$ и $b_{\mathbf{k}}^\dagger$ соответственно.

Поле осциллятора в out -представлении имеет вид:

$$q_{\mathbf{k}} = b_{\mathbf{k}} \nu_{\mathbf{k}}^{(out)} + b_{-\mathbf{k}}^\dagger \nu_{\mathbf{k}}^{(out)*},$$

$$\nu_{\mathbf{k}}^{(out)}(\eta > \eta_2) = \frac{e^{-i\omega_2 \eta}}{\bar{\alpha}'_2 \eta \sqrt{2\omega_2}}, \quad (8.6.2)$$

где $\omega_2 = k$ или $k/\sqrt{3}$ для полей T или S соответственно.

Поскольку оба представления $a_{\mathbf{k}}$ и $b_{\mathbf{k}}$ описывают одну и ту же переменную $q_{\mathbf{k}}$, они должны быть связаны между собой преобразованием Боголюбова:

$$b_{\mathbf{k}} = \alpha_{\mathbf{k}} a_{\mathbf{k}} + \beta_{\mathbf{k}}^* a_{-\mathbf{k}}^\dagger,$$

$$|\alpha_{\mathbf{k}}|^2 - |\beta_{\mathbf{k}}|^2 = 1, \quad (8.6.3)$$

где $\alpha_{\mathbf{k}}$ и $\beta_{\mathbf{k}}$ — коэффициенты Боголюбова. В соответствии с законом сохранения импульса оператор уничтожения b -частицы с импульсом \mathbf{k} является суперпозицией операторов уничтожения a -частицы с импульсом \mathbf{k} и рождения a - (анти)частицы с импульсом $-\mathbf{k}$.

Усредняя операторы $b_{\mathbf{k}}$ по состоянию (8.3.11), получаем числа заполнения частиц в out -зоне (ср. (8.4.11)):

$$\langle b_{\mathbf{k}}^\dagger b_{\mathbf{k}'} \rangle = n_{\mathbf{k}} \delta^{(3)}(\mathbf{k} - \mathbf{k}'), \quad n_{\mathbf{k}} = |\beta_{\mathbf{k}}|^2. \quad (8.6.4)$$

Числа n_k характеризует вариацию поля возмущений при $\eta > \eta_2$, но они ничего не говорят о фазовом портрете распределения возмущений. Для этого необходимо найти соотношение между фазовыми амплитудами поля на радиационно-доминированной стадии расширения.

Перепишем формулу для коэффициентов Боголюбова в несколько ином виде:

$$(\alpha_k - \beta_k)(\alpha_k + \beta_k)^* = 1 + i c_k, \quad (8.6.5)$$

где c_k — действительные числа. Не умоляя общности, мы можем считать суммы $\alpha_k + \beta_k$ также действительными числами. Обозначим обратные к ним через $R = R(k)$. Тогда из (8.6.5) получаем:

$$\alpha_k + \beta_k = \frac{1}{R}, \quad \alpha_k - \beta_k = R(1 + i c_k). \quad (8.6.6)$$

Вводя фазовые операторы стоячих волн (ср. (8.4.6))

$$C_1^{(out)} = \frac{b_k - b_{-k}^\dagger}{i\sqrt{2}}, \quad C_2^{(out)} = \frac{b_k + b_{-k}^\dagger}{\sqrt{2}}, \quad (8.6.7)$$

$$q(> \eta_2) = \frac{1}{\bar{\alpha}_2} \int \left(C_1^{(out)} \frac{\sin \omega \eta}{\omega \eta} + C_2^{(out)} \frac{\cos \omega \eta}{\omega \eta} \right) e^{ikx} \frac{\sqrt{\omega} dk}{(2\pi)^{3/2}},$$

получаем адиабатический инвариант в *out*-зоне:

$$\bar{\mathcal{E}}_k^{(out)} = \frac{1}{2} \omega \left(|C_1^{(out)}|^2 + |C_2^{(out)}|^2 \right), \quad (8.6.8)$$

и соотношения между фазовыми *out*- и *in*-операторами:

$$\begin{aligned} C_1^{(out)} &= R \left(C_1^{(in)} + c_k C_2^{(in)} \right), \\ C_2^{(out)} &= R^{-1} C_2^{(in)}. \end{aligned} \quad (8.6.9)$$

Напомним, что *in*-амплитуды в состоянии $|0\rangle$ равны между собой (см. (8.4.6), (8.4.10)). Однако *out*-амплитуды, как мы видим, различаются, то есть поле *сжато*.

Параметры c_k и R тождественно совпадают с аналогичными параметрами в формулах (8.3.9) и (8.3.12). Коэффициент R описывает степень сжатия k -осциллятора в *out*-зоне:

$$\begin{aligned} \nu_k^{(in)}(\eta > \eta_2) &= \alpha_k \nu_k^{(out)} + \beta_k \nu_k^{(out)*} = \\ &= \frac{(c_k - i) R^2 \sin(\omega_2 \eta) + \cos(\omega_2 \eta)}{\bar{\alpha} R \sqrt{2\omega_2}}. \end{aligned} \quad (8.6.10)$$

При условии $k\eta_2 \ll 1$ это решение аналитически продолжается в область (8.3.2), откуда и следуют известные нам соотношения для спектральной амплитуды q_k и коэффициента сжатия R (см. (8.3.4), (8.3.9)).

Очевидно, оба фактора пропорциональны друг другу:

$$q_k \propto C^{-1} \propto R.$$

Это означает, что достаточно большая величина S-моды, требуемая для образования нелинейной структуры через миллиард лет после Большого взрыва, может быть получена только при значительном сжатии начальных возмущений в *out*-зоне:

$$R = \frac{n_0 k_{\text{rad}}^2 q_k}{k^2 \sqrt{1 + c_k^2}} \sim 10^{45} \left(\frac{M_{\text{пк}}^{-1}}{k} \right)^2 \left(\frac{q_k}{10^{-5} \sqrt{1 + c_k^2}} \right), \quad (8.6.11)$$

где $n_0 = 4\pi 3^{1/4}$ (см. (8.3.7)). Аналогичное соотношение выполняется и для T-мод с той лишь разницей, что для них $n_0 = 2\pi 3^{-1/2}$.

Таким образом, при ограниченных значениях коэффициента c_k степень сжатия космологических возмущений составляет колоссальную величину, свидетельствующую о безусловном приоритете растущей фазовой ветви начальных неоднородностей в образовании структуры Вселенной (ср. (8.3.12)):

$$c_1^{(out)} \gg c_2^{(out)}. \quad (8.6.12)$$

8.7. Амплитудная информация

Обратимся к спектральным характеристикам рожденных неоднородностей.

Ожидания линейных полей возмущений по определению равны нулю, но имеются вариации полей:

$$\langle q^2 \rangle = \int_0^{\infty} S \frac{dk}{k}, \quad S \equiv q_k^2, \quad (8.7.1)$$

$$\langle \mathcal{G}_{ij} \mathcal{G}^{ij} \rangle = \int_0^{\infty} T \frac{dk}{k}, \quad T \equiv 2q_{k\lambda}^2, \quad (8.7.2)$$

где $q_{k\lambda} = q_{k\oplus} = q_{k\otimes}$ — спектр любой из двух поляризаций гравитационных волн (8.5.12), скобки $\langle \dots \rangle$ означают усреднение по состоянию (8.3.11).

В силу случайности пространственных фаз родившихся полей возмущений это усреднение эквивалентно усреднению по всем реализациям или по большому пространственному объему (эргодическая теорема). Более того, как мы уже отмечали ранее, пространственные распределения полей гауссовы; это означает, что корреляторы второго порядка q_k и $q_{k\lambda}$ содержат всю статистическую информацию об их распределении (через них, в частности, выражаются корреляторы высших порядков). Поэтому скобки в (8.7.1) и (8.7.2) можно считать усреднением по гауссовому ансамблю пространственных реализаций скалярного поля.

С точки зрения наблюдательной космологии нам интересен диапазон пространственных частот, который в ранней Вселенной находился в причинно-несвязанной области пространства. На этом начальном этапе развития космологической структуры поля S и T постоянны во времени и их спектры имеют вид (8.3.12):

$$q_{k(\lambda)} = \frac{k \sqrt{1 + c_{k(\lambda)}^2}}{2\pi C}, \quad (8.7.3)$$

где значения космологических параметров берутся в момент генерации η_1 , зависящий от волнового числа k .

Амплитуды q_k , $q_{k\oplus}$ и $q_{k\otimes}$ задают дисперсии соответствующих гармоник полей q , q_{\oplus} и q_{\otimes} в диапазоне частот $\delta k = k$. Интегрирование по логарифмическому частотному интервалу в (8.7.1) и (8.7.2) реализует теорему Пифагора: средний квадрат поля равен сумме средних квадратов его гармоник. Квадраты соответствующих амплитуд называются *спектрами мощности* скалярных $S = S(k)$ и тензорных $T = T(k)$ возмущений.

Индексы наклонов и отношение спектров мощности определены следующим образом:

$$n_S \equiv \frac{d \ln S}{d \ln k}, \quad n_T \equiv \frac{d \ln T}{d \ln k}, \quad (8.7.4)$$

$$r \equiv \frac{T}{S}, \quad r_S = \frac{d n_S}{d \ln k}. \quad (8.7.5)$$

По историческим причинам часто вместо n_S используют индекс $n = n_S + 1$, сдвинутый на единицу относительно n_S . Мы также ввели фактор r_S изменения наклона спектра мощности моды S .

Все спектральные параметры являются функциями волнового числа k . Если (в каком-то интервале длин волн) спектры не зависят от k , их называют *плоскими* или спектрами Харрисона–Зельдовича; для них $n_S = n_T = 0$, а $n = 1$. Спектры, растущие с ростом k , называются *голубыми*, а падающие — *красными*.

Начальные спектры возмущений плотности, пекулярной скорости и гравитационного потенциала зависят от времени и связаны с q_k следующим образом:

$$\langle \delta_\varepsilon^2 \rangle = \int_0^\infty \delta_k^2 \frac{dk}{k}, \quad \langle \mathbf{v}_{\text{pec}}^2 \rangle = \int_0^\infty v_k^2 \frac{dk}{k}, \quad \langle \Phi^2 \rangle = \int_0^\infty \Phi_k^2 \frac{dk}{k},$$

$$\delta_k = g k^2 q_k, \quad v_k = \nu k q_k, \quad \Phi_k = \phi q_k, \quad (8.7.6)$$

$$\nu = \frac{1}{a^2} \int_0^a \frac{da}{H}, \quad \phi = \gamma \bar{H}^2 g = 1 - \frac{H}{a} \int_0^a \frac{da}{H}. \quad (8.7.7)$$

Как и q_k (см. (8.7.3)), эти спектры определены в причинно-несвязанной области пространства. Факторы роста $g = g(a)$, $\nu = \nu(a)$ и $\phi = \phi(a)$ являются функциями масштабного фактора a и не зависят от k (см. (5.5.4)).

8.8. Решение проблемы начальных неоднородностей

Как найти явную зависимость спектров от частоты k ?

Для ответа на этот вопрос необходимо решить уравнение (8.2.8) в переходной области $a \sim a_1$ и определить параметры возмущений.

Полная эволюция функции $\sqrt{2k} \nu_k(\eta)$ выглядит следующим образом (см. (8.3.1), (8.3.2), (8.3.9), (8.6.10)):

$$\frac{e^{-i\kappa}}{\bar{\alpha} \sqrt{\beta}} \xrightarrow{1} \frac{c-i}{c} \xrightarrow{2} \frac{c_k-i}{c} + \frac{1}{\bar{R}\eta}, \quad (8.8.1)$$

где $\bar{R} \equiv \bar{\alpha}'_2 \sqrt{\beta_2} R = \text{const}$, нижние индексы 1 и 2 относятся к соответствующим меткам времени (см. раздел 8.6). Напомним, что скорость звука в области a_1 эволюционирует адиабатически, а вблизи a_2 $\beta = 3^{-1/2}$. Для тензорных мод $\beta = 1$, индекс поляризации λ по возможности опускаем.

Предполагая, что переход из адиабатической зоны в параметрическую произошел за короткое время порядка космологического, мы можем просто сшить модули¹⁾ амплитуд (8.3.1) и (8.3.2)

¹⁾ Чтобы не учитывать начальную фазу.

и их производных до и после момента времени η_1 :

$$|\nu_k|_1 \simeq \frac{1}{\bar{\alpha}_1 \sqrt{2\omega_1}}, \quad \left| \frac{(\bar{\alpha}\nu_k)'}{\bar{\alpha}\omega} + i\nu_k \right|_1 \simeq 0. \quad (8.8.2)$$

В результате фиксируем значения констант (см. (8.3.2)):

$$c_1 \simeq \frac{\mathfrak{h}_1}{\omega_1}, \quad (8.8.3)$$

$$C \simeq \bar{\alpha}_1 \sqrt{\beta_1(1+c_1^2)} \simeq \frac{|\bar{\alpha}'_1| \sqrt{1+c_1^{-2}}}{k \sqrt{\beta_1}}, \quad (8.8.4)$$

и функцию c в области $\eta > \eta_1$:

$$c \simeq k C^2 \left(\frac{1}{\bar{\alpha}\bar{\alpha}'} + \int_{\eta_1}^{\eta} \frac{U d\eta}{\bar{\alpha}'^2} \right) - c_1^{-1}, \quad (8.8.5)$$

где $c_1 = c(\eta_1)$. Устремляя время к моменту η_2 , получаем кумулятивную постоянную (напомним, что $U(\eta > \eta_2) = 0$):

$$c_k \simeq (1+c_1^{-2}) \frac{\bar{\alpha}'_1{}^2}{\omega_1} \int_{\eta_1}^{\infty} \frac{U d\eta}{\bar{\alpha}'^2} - c_1^{-1}. \quad (8.8.6)$$

В общем случае

$$\mathfrak{h}_1 \simeq \frac{\bar{\alpha}'_1}{\bar{\alpha}_1} = \bar{H}_1(1+\Gamma_1).$$

Для динамических режимов расширения с $|1+\Gamma_1| \simeq 1$ из (8.3.14), (8.3.15) имеем:

$$|c_1| \simeq \frac{\bar{H}_1}{\omega_1} \gtrsim 1, \quad C \simeq \frac{|\bar{\alpha}'_1|}{k\sqrt{\beta_1}},$$

$$|c_k| \lesssim \frac{|U_{max}|_1}{\omega_1 \bar{H}_1} \lesssim |c_1|, \quad (8.8.7)$$

где $|U_{max}|_1$ — максимальное значение потенциала $|U|$ вблизи перехода $\eta \sim \eta_1$. Отсюда получаем искомые выражения для коэффициентов сжатия:

$$R \simeq \frac{\bar{\alpha}'_2}{|\bar{\alpha}'_1|} \left(\beta_1 \sqrt{3} \right)^{1/2}, \quad R_\lambda \simeq \frac{a'_2}{a'_1}, \quad (8.8.8)$$

и спектральных амплитуд (8.7.3):

$$q_{k(\lambda)} \simeq \sqrt{\frac{2}{\pi}} \frac{\omega_1 \ell_P \sqrt{1 + c_k^2}}{a_1 |c_1| f} \lesssim \frac{\omega_1 \ell_P}{a_1 f} \simeq \frac{H_1}{M_P |c_1| f} \lesssim \frac{H_1}{M_P f}, \quad (8.8.9)$$

где $f = \sqrt{2\gamma_1 \beta_1}$, $f_\lambda = 1$.

Этот достаточно общий результат получен в предположении о короткой переходной области вблизи a_1 , которое, как правило, выполняется и позволяет сшить асимптотики (8.3.1) и (8.3.2), не прибегая к решению уравнений (8.2.8).

Более того, в пределе $|c_1| \gg 1$ полученные формулы оказываются точными, так как в области $|U| \sim \omega^2$ определена малая функция $\kappa = \omega/\hbar$, что проверяется прямым дифференцированием (см. (8.8.3)):

$$\left(\frac{\kappa}{\omega}\right)' = 1 - \frac{U}{\hbar^2} \simeq 1.$$

«Сквозная» фазовая функция $\kappa(\eta)$ дает возможность аналитически продолжить оба решения (8.3.1) и (8.3.2) в смежную область $\sim \eta_1$:

$$e^{-i\kappa} \simeq 1 - i\kappa = \frac{(c - i)\bar{a}_1 \sqrt{\beta_1}}{c}, \quad (8.8.10)$$

где в зоне перекрытия решения тождественны, что позволяет найти связь между параметрами (ср. (8.8.3), (8.8.4)):

$$c_1 = \kappa_1^{-1}, \quad C = c_1 \bar{a}_1 \sqrt{\beta_1}. \quad (8.8.11)$$

Как следует из (8.8.6)–(8.8.9), модуль фактора c_k ограничен модулем c_1 , а максимальные амплитуды $q_{k(\lambda)}$, генерируемые при заданных H и γ , реализуются в предельном случае малых $|c_1| \sim 1$. В этих условиях всегда $R \gg 1$ (см. (8.6.11)), что требует выполнения неравенства:

$$\bar{a}'_2 \gg |\bar{a}'_1|. \quad (8.8.12)$$

Таким образом, мы можем сделать важный вывод: условия, обеспечившие образование космологической структуры, были прямым образом связаны с производством энтропии Вселенной и с высокой степенью сжатия поля начальных возмущений (преимущество растущей ветви решения). Существенный рост функции $\bar{\alpha}$ в период генерации космологических возмущений¹⁾ говорит о том, что эффективный потенциал находился в положи-

¹⁾ Это не означает, что на отдельных этапах функция $a\sqrt{\gamma}$ не могла убывать со временем (см. далее раздел 9.4).

тельной области значений $U > 0$ и, следовательно, эффективное давление материи было меньше трети ее плотности энергии ($p < \varepsilon/3$). Тогда условие перехода элементарных осцилляторов в параметрическую область в ранней Вселенной выглядело следующим образом:

$$\eta = \eta_1: \quad \omega^2 = U > 0. \quad (8.8.13)$$

Соотношения (8.7.1)–(8.8.13) решают задачу о генерации начальных космологических возмущений в ранней Вселенной. Явный вид спектра q_k зависит от динамики космологического расширения и требует обращения уравнения (8.8.13) либо точного решения уравнений (8.2.8).

Подведем краткие итоги анализа свойств возмущений.

При жестких уравнениях состояния в период генерации возмущений ($p > \varepsilon/3$) выделяется падающая ветвь решения ($R < 1$), однако ее амплитуда чрезвычайно мала ($c_1 \gg 1$) и недостаточна для реализации структуры.

При мягком эффективном давлении $p < \varepsilon/3$ растущая ветвь усиливается ($R > 1$) и амплитуда возрастает. Наиболее высокие амплитуды возмущений получаются в предельно мягких (инфляционных) сценариях ранней Вселенной, где параметр c_1 понижается до величины порядка единицы. В этом случае быть и оставаться большой позволено только растущей ветви решения (см. (8.6.12)).

Такой результат неудивителен и объясняется общим принципом Ле Шателье.

Для генерации достаточно больших амплитуд ($\sim 10^{-5}$) требуются большие H , то есть ранние эпохи эволюции. Растущая ветвь возмущений, ассоциированная с константой c_1 на радиационно-доминированной стадии (см. (8.6.12)), не расходится при $\omega\eta_2 \ll 1$ и имеет конечную амплитуду, совместимую с фридмановской геометрией при сколь угодно малых временах (см. (8.6.7)). Напротив, падающая ветвь ($\propto c_2$) должна быть подавлена в ранней Вселенной, поскольку ее амплитуда расходится как η^{-1} и может нарушить фоновую симметрию. Однородный хаббловский поток с изначально малыми возмущениями ($\sim 10^{-5}$) не может саморазрушиться за короткий отрезок времени. Рост неоднородностей требует длительного времени ($\Delta N \sim 10$), тогда они могут вырасти и образовать нелинейную структуру лишь в причинно-связанных областях пространства, если дополнительно выполняются условия гравитационной неустойчивости (в реальной Вселенной все это заняло миллиар-

ды лет). Однако производство самих затравочных возмущений происходит за времена порядка хаббловских и не связано с гравитационной неустойчивостью, мы говорим о другом эффекте — гравитационно-параметрическом, не имеющем аналога в нерелятивистской физике.

Этот эффект можно проиллюстрировать на примере краткосрочной параметрической зоны в ходе радиационно-доминированного расширения. Пусть при малых $\eta \ll \omega^{-1}$ начальная фаза осциллятора случайна и он минимально возбужден: $c_1 \sim c_2 \sim 1$ (см. (8.4.10)). При появлении краткодействующего потенциала $U > 0$ в момент η_1 происходит «удар» по осциллятору и сбиваются его временные фазы. Сама амплитуда за это время не успевает измениться ($q \sim \text{const}$), но она перераспределяется по новым фазам. В результате после удара осциллятор возбужден в растущей фазе колебаний: $c_1 \sim (\omega\eta_1)^{-1} \gg c_2 \sim 1$.

Формулы (8.7.3) описывают спектры первичных космологических возмущений в зависимости от значений динамических параметров Вселенной в эпоху генерации возмущений. Наблюдательная космология ставит перед собой обратную задачу: по известным спектрам, полученным из данных о структуре Вселенной, определить параметры Большого взрыва.

Корректность решения обратной задачи требует отдельного исследования, однако, как это часто случается на практике, существует более прозаическая и очевидная проблема. Дело в том, что один из ключевых параметров спектра c_k не носит явной локальной зависимости от a_1 , а является интегральной характеристикой истории расширения Вселенной в конечный период времени, примыкающий к a_1 . Поэтому в реальности мы можем тестировать отдельные, хотя и достаточно общие, модели ранней Вселенной, предсказывающие динамику масштабного фактора в зависимости от внутренних параметров модели, и определять эти параметры путем сравнения теории с наблюдениями.

Вся схема исследования выглядит следующим образом.

Сначала строится физически мотивированная модель изотропного космологического расширения, зависящая от некоторого числа свободных параметров. Потом по формулам (8.7.3) (или с помощью точных решений (8.2.8)) для нее вычисляются спектры космологических возмущений. Эти спектры сопоставляются с наблюдениями и определяются параметры модели (либо модель отвергается на текущем уровне наблюдательных данных). Очевидно, этот путь может быть успешным только при достаточно большом объеме и точности имеющихся данных о структуре Вселенной.

Современная наблюдательная космология отвечает этим требованиям.

К счастью, амплитуды и формы спектров рожденных возмущений чрезвычайно чувствительны к динамическим параметрам ранней Вселенной (см. (8.7.3)). На протяжении последних лет данный метод доказал свою эффективность. Он привел к важным открытиям в теории ранней Вселенной и сохраняет все шансы для восстановления динамической истории и построения модели Большого взрыва в эпоху генерации возмущений — в первую очередь, функций $H(a)$ и $\gamma(a)$.

СПЕКТРЫ ВОЗМУЩЕНИЙ ФРИДМАНОВСКИХ МИРОВ

В этой главе мы продолжим изучение спектров космологических неоднородностей на примере динамических моделей ранней Вселенной и рассмотрим условия, необходимые для создания как начальных возмущений кривизны, из которых сформировалась структура Вселенной, так и самой крупномасштабной «подложки», деформированной этими возмущениями, — квазифридмановского мира.

9.1. Адиабатическая релаксация

Начнем с простого примера космологического расширения.

В случае постоянных β и γ (см. (4.1.9)) масштабный фактор пространственно-плоской Вселенной имеет степенной вид:

$$\gamma = \text{const}: \quad a \propto t^{1/\gamma} \propto |\eta|^{1/\bar{\gamma}}, \quad \bar{H} = \frac{1}{\bar{\gamma}\eta}, \quad (9.1.1)$$

а поскольку $\Gamma = 0$, эффективные потенциалы полей S и T тождественно совпадают (см. (8.3.14)).

Этой же аппроксимацией можно пользоваться, если функции β и γ *адиабатически* изменяются в ходе эволюции по сравнению с хаббловским масштабом:

$$|\Gamma| < 1, \quad |\dot{\Gamma}| < H: \quad U_S \simeq U_T = \bar{H}^2(2 - \gamma). \quad (9.1.2)$$

Такие *квазистепенные* режимы расширения характерны для ряда космологических эпох, например для радиационно-доминированных или адиабатически-инфляционных стадий (см. раздел 7.4).

Помимо этого во Вселенной были и периоды неадиабатических изменений функций β или γ , например ранние этапы материально-доминированных или инфляционных стадий. При смене уравнения состояния, когда очередная компонента материи начинает доминировать по плотности и влиять на динамику космологического расширения, релаксация β и/или γ протекает, как правило, в неадиабатическом режиме. И только позже,

по прошествии некоторого времени, она принимает спокойный адиабатический характер.

Так, в современную эпоху мы находимся вначале инфляционной неадиабатической стадии с

$$\begin{aligned}\gamma_0 &= \frac{3}{2} \Omega_m \simeq 0,4, \\ \Gamma_0 &= \frac{\gamma'}{2\gamma\bar{H}} = -\frac{3}{2} \Omega_E \simeq -1.\end{aligned}\tag{9.1.3}$$

Этот и другие примеры неадиабатического расширения мы рассмотрим в последующих разделах, а сейчас обратимся к квазистепенной эволюции и спектрам возмущений.

Из (9.1.2) получаем следующую оценку интеграла для функции c_k (см. (8.8.6)):

$$\frac{\bar{\alpha}'^2}{\omega} \int_{\eta_1} U \frac{d\eta}{\bar{\alpha}'^2} \simeq \frac{2-\gamma}{\omega} \int_{\eta_1} \frac{d\eta}{(a/a_1)^2} \simeq \frac{2-\gamma}{3-\gamma} c_1, \quad c_1 \simeq \frac{\bar{H}_1}{\omega}.\tag{9.1.4}$$

Здесь предполагается, что режим с квазипостоянным γ примыкает к зоне подбарьерного перехода η_1 , где функция $U(\eta)$, выведенная из нуля в положительную область, пересекает уровень ω^2 ; индекс «1» у γ и ω опущен.

В зависимости от того, больше или меньше единицы значение γ , решение принадлежит одному из двух классов с существенно различным поведением функции $\bar{H}(\eta)$ (см. (9.1.1)).

Если $\gamma \in (1, 2)$, функция $\bar{H}(\eta)$ затухает со временем (рис. 9.1). Здесь удобно считать, что в области возникновения барьера $\gamma(\eta)$ плавно изменяется от 2 до своего конечного значения γ за характерное хаббловское время. Поскольку это изменение не превышает единицы, выполняются условия (9.1.2) и зависимостью параметров a_1 и H_1 от k в интересующей нас области частот $\omega \ll \bar{H}_1$ можно пренебречь. Факторы c_1 и c_k оказываются больше единицы (см. (9.1.4)). Отсюда получаем следующие выражения для спектров мод T и S космологических возмущений (см. (8.8.9)):

$$q_{k(\lambda)} \simeq \frac{(2-\gamma)\omega \ell_P}{\sqrt{\pi} a_1 f} \ll \frac{H_1}{M_P},\tag{9.1.5}$$

$$\Gamma = \frac{4\gamma}{\beta} > 1,\tag{9.1.6}$$

которые оказываются «голубыми» с индексами наклонов $n_T = n_S = 2$ (соответственно $n = 3$).

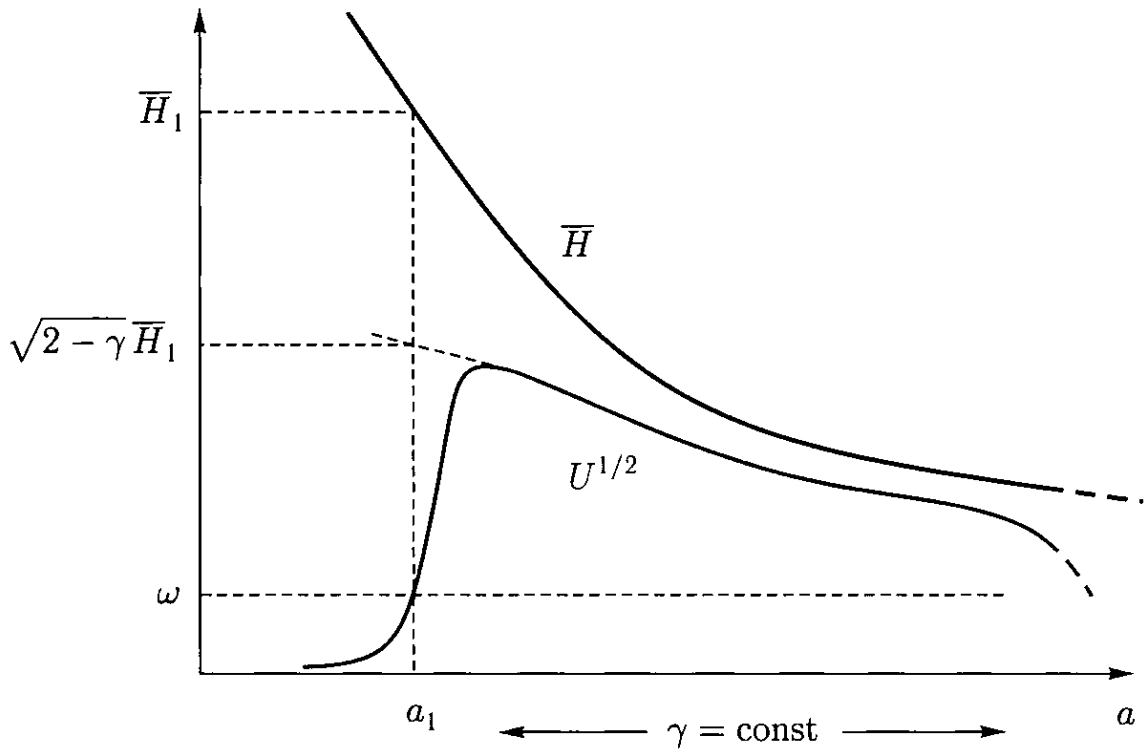


Рис. 9.1. Функции \bar{H} и $U^{1/2}$ на квазистепенной стадии эволюции с $\gamma \in (1, 2)$

Если $\gamma \in (0, 1)$, то функция $\bar{H}(\eta)$ монотонно возрастает со временем (рис. 9.2). Здесь можно считать, что период с квазипостоянным γ включает в себя область самого подбарьерного перехода. Момент η_1 наступает естественным образом в ходе роста функции $\bar{H}(\eta)$ и зависит от длины волны: $\bar{H}_1 = \omega$ (см. (9.1.2)).

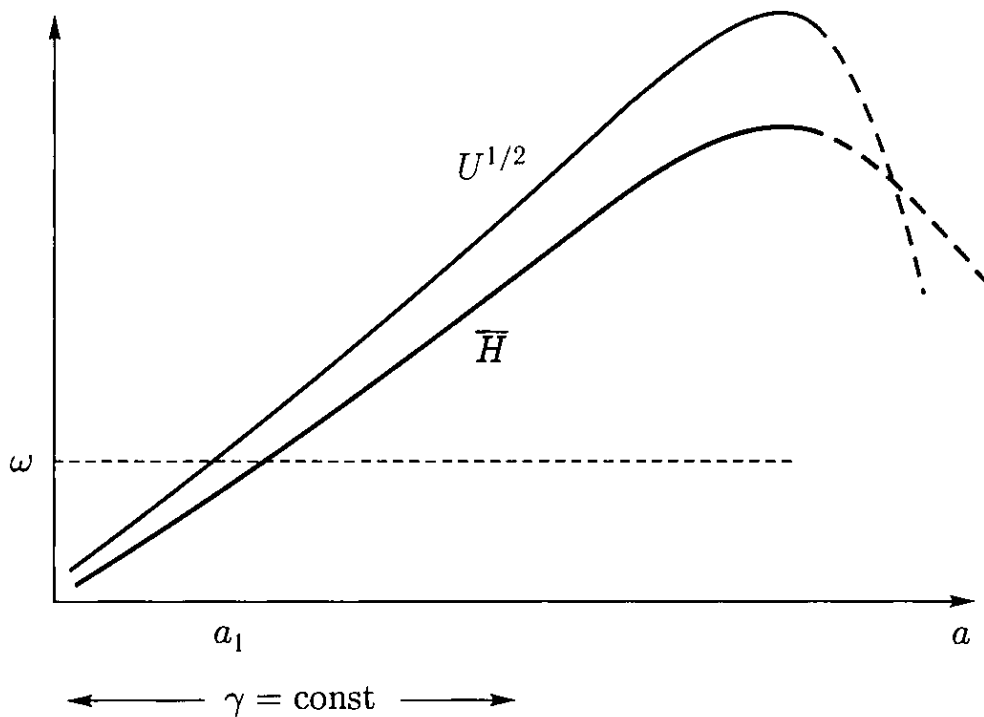


Рис. 9.2. Функции \bar{H} и $U^{1/2}$ на квазистепенной стадии эволюции с $\gamma \in (0, 1)$

В этих условиях $0 \leq \epsilon_k < \epsilon_1 = 1$ (см. (9.1.4), (8.8.7)) и спектры космологических возмущений принимают вид:

$$q_k(\lambda) \simeq \sqrt{\frac{2}{\pi}} \frac{\omega \ell_P}{a_1 f} \simeq \frac{H_1}{M_P f} \propto \omega^{-\gamma/(1-\gamma)}, \quad (9.1.7)$$

$$r = 4\gamma\beta^{(1+\gamma)/(1-\gamma)}. \quad (9.1.8)$$

Здесь мы получили «красные» спектры¹⁾ с индексами наклонов $n_T = n_S = -2\gamma/(1-\gamma)$ (и $n = (1-3\gamma)/(1-\gamma)$).

Таким образом, при уменьшении γ фактор $r = T/S$ становится меньше единицы:

$$\gamma \ll 1: \quad r = 4\gamma\beta \ll 1, \quad (9.1.9)$$

что кардинально отличается от (9.1.6).

В важном частном случае, когда скорость звука равна скорости света, из (9.1.5) и (9.1.7) следует простая формула для отношения спектров мощности при *любых* значениях γ в интервале $\gamma \in (0, 2)$:

$$\beta = 1: \quad r = 4\gamma. \quad (9.1.10)$$

Здесь следует остановиться и задать вопрос. Почему при одинаковых эффективных потенциалах (9.1.2) и скоростях распространения возмущений результирующие решения оказываются различными и $r \neq 1$?

Действительно, при $\gamma = \text{const}$ и $\beta = 1$ уравнения (8.2.8) тождественны,

$$\ddot{q}(\lambda) + 3H\dot{q}(\lambda) - a^{-2}\Delta q(\lambda) = 0, \quad (9.1.11)$$

независимо от того, имеем ли мы дело со скалярным q или тензорными q_λ ($\lambda = \oplus, \otimes$) полями возмущений. Поэтому все \mathbf{k} -осцилляторы ведут себя одинаково. Неудивительно, что наклоны результирующих спектров мощности оказываются идентичными, однако почему же различны их амплитуды?

Все дело в начальных условиях: осцилляторы имеют разные начальные амплитуды. Амплитуда нулевых вакуумных флуктуаций продольной части кривизны в $(2\gamma)^{-1/2} > 1$ раз выше, чем у гравитационно-волновых нулевых флуктуаций вакуума.

Действительно, для лагранжевых плотностей полей q и q_λ в этом случае имеем (см. (8.1.1), (8.5.10)):

$$L = \frac{\gamma}{8\pi G} q_{,\mu} q'^{\mu}, \quad L_\lambda = \frac{1}{16\pi G} q_{\lambda,\mu} q_\lambda'^{\mu}, \quad (9.1.12)$$

¹⁾ Формулы (9.1.7) и (9.1.8) нельзя аппроксимировать к значению $\gamma = 1$, так как они выведены в предположении роста функции $\overline{H}(\eta)$.

откуда и следует отношение затравочных амплитуд элементарных осцилляторов в начальной волновой зоне (равные энергии k -осцилляторов, см. (8.3.1), (8.3.11)):

$$\frac{\langle |q_{k\lambda}|^2 \rangle}{\langle |q_k|^2 \rangle} = 2\gamma. \quad (9.1.13)$$

Это соотношение справедливо для каждой из двух поляризаций гравитационных волн. Умножая его на два по числу поляризаций, получаем фундаментальную формулу (9.1.10)¹⁾.

Мы видим, что теория никак не дискриминирует величину моды T по отношению к S . Все дело в значении фактора γ в эпоху генерации возмущений. При $\gamma \sim 1$ амплитуда T -моды порядка или больше S -амплитуды, при $\gamma \ll 1$ мода T подавлена с фактором 4γ .

Наблюдательные данные указывают на то, что в реальной Вселенной величина T -моды как минимум в десять раз меньше моды S :

$$r < 0,1. \quad (9.1.14)$$

Этот факт непосредственно указывает на то, что генерация космологических возмущений происходила до завершения инфляционной стадии Большого взрыва, поскольку γ меньше единицы (подробнее см. раздел 10.5).

Рассмотренные примеры демонстрируют основные спектральные свойства возмущений, возникающие и при более общих предположениях об эпохе Большого взрыва. В принципе, можно получить любые спектры, моделируя скорость и ускорение масштабного фактора в ранней Вселенной. При неадиабатической динамике спектры T - и S -мод, как мы увидим ниже, будут различаться не только по величине, но и по форме, однако общее утверждение о малости T -моды на стадии инфляции сохраняется.

Подбарьерный переход элементарного осциллятора, необходимый для его параметрической раскачки, происходит в момент времени η_1 , когда эффективная частота обращается в нуль ($U = \omega^2$). Потенциал U состоит из двух множителей: \bar{H}^2 и членов после фигурной скобки (8.3.14). Его можно обратить в нуль

¹⁾ Энергия возбуждений безмассовых полей (9.1.12) равномерно распределена по модам q и q_λ . При $\gamma < 1$ наблюдатели в де-ситтеровском конусе измеряют тепловые флуктуации полей с характерной энергией $E_q = k/a \sim H$ и плотностью энергии $\epsilon_q \sim H^4 \sim \langle L \rangle = \langle L_\lambda \rangle$ (см. раздел 4.2, (8.4.11)), отсюда получаем дисперсии полей $q_\lambda = \sqrt{2\gamma} q \sim H/M_P$ (ср. (9.1.7))

($U = 0$) в двух случаях: либо при $\gamma = 2$ (см. сноску на с. 226), либо при $\bar{H} \rightarrow 0$.

Таким образом, в общем случае, как и для квазистепенной эволюции, вывести функцию $U(\eta)$ из нуля в область положительных значений можно двумя способами: либо уменьшая от 2 $\gamma(\eta)$, либо увеличивая $\bar{H}(\eta)$. В первом случае $\dot{\gamma} < 0$, и мы имеем переход от первичной радиационно-доминированной эволюции к более мягкому уравнению состояния с $\gamma < 2$. Во втором случае $\bar{H} > 0$, то есть $\bar{\gamma} < 0$, а это инфляция.

Ниже мы рассмотрим физические примеры обеих возможностей.

9.2. Тяжелые частицы в ранней Вселенной

Пусть в начальной волновой зоне $\gamma = 2$ и $U = 0$. Для возрастания потенциала $U(\eta)$ необходимо, чтобы функция γ уменьшилась и стала меньше 2. Примером подобной эволюции может быть закалка и отсоединение нерелятивистских частиц от тепловой «бани» излучения с последующим их доминированием. В дальнейшем тяжелые частицы распадаются на более легкие, происходит термализация и восстанавливается релятивистское расширение с $\gamma = 2$.

Рассмотрим этот вариант эволюции подробнее.

Назовем тяжелые частицы для краткости m -частицами и обозначим безразмерный масштабный фактор в интересующий нас период времени через

$$y = \frac{a}{a_m}, \quad \text{где} \quad a_{\text{eq}} = \frac{a_m}{4} = \text{const}$$

— значение масштабного фактора в момент начала доминирования m -частиц в ранней Вселенной (равные плотности энергий частиц и излучения).

Вся эволюция разбивается на четыре этапа:

- 1) $y < y_1 \ll 1$,
- 2) $y_1 < y < y_d$,
- 3) $y_d < y < y_2$,
- 4) $y > y_2 \gg 1$.

На первом ($y < y_1 \ll 1$) и четвертом ($y > y_2 \gg 1$) этапах $\gamma = 2$ ($U = 0$), где $y_{1,2}$ соответствуют моментам заделки количества m -частиц в сопутствующем объеме и началу тепловой *out*-зоны.

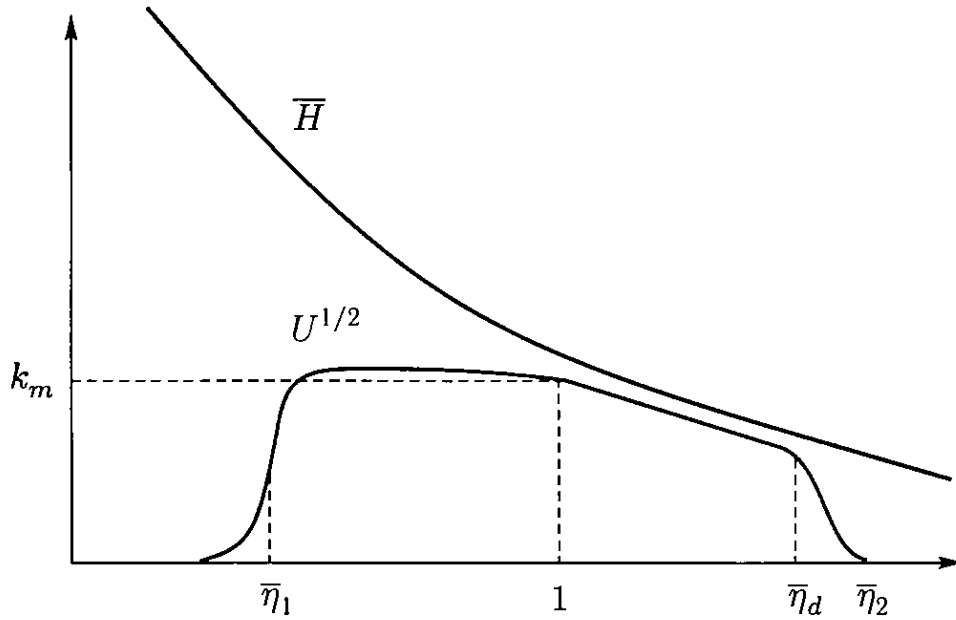


Рис. 9.3. Параметрический потенциал $U(\bar{\eta})$ в модели с тяжелыми частицами. Красная и синяя линии соответствуют функциям \sqrt{U} и \bar{H}

В области $y_1 < y < y_d \sim y_2$ мы имеем стабильные частицы:

$$y = \bar{\eta}(1 + \bar{\eta}), \quad \gamma = \frac{2(1 + 3y)}{1 + 4y}, \quad (9.2.1)$$

где $\bar{\eta} = k_m(\eta + \eta_0)$ — безразмерное время, а k_m и η_0 — постоянные. Эффективная скорость звука на этом этапе уменьшается: $\sqrt{3}\beta = (1 + 3y)^{-1/2}$.

На коротком интервале времени $y \in (y_d, y_2)$ происходит распад m -частиц и разогрев вещества, и скорость звука восстанавливается до прежней величины $3^{-1/2}$.

Заметим, что условие $y_{eq} = 1/4$ следует из уравнения Фридмана $(dy/d\bar{\eta})^2 = 1 + 4y$, в котором факторы 1 и $4y$ соответствуют энергиям излучения и m -частиц на стадии (9.2.1). Константы a_m и k_m связаны условием:

$$a_m k_m = \sqrt{\frac{8\pi}{3}} \frac{k_{\text{rad}}^2 \ell_P}{R}, \quad R = \frac{a'_2}{a'_1} \simeq 1 + 2\bar{\eta}_2 \simeq 2e^{\mathcal{N}_2/2} > 1,$$

где коэффициент сжатия R универсален для всех элементарных осцилляторов, а $\mathcal{N}_2 = \ln(a_2/a_{\text{eq}}) \simeq 2 \ln \bar{\eta}_2$ — число хаббловских эпох на стадии доминирования m -частиц. Нас интересуют длинноволновые осцилляторы с $k \ll k_m$. В этом пределе обе среды движутся совместно в поле адиабатических возмущений и справедлива формула (8.3.2)¹⁾.

¹⁾ Для более коротких волн $k \sim k_m$ следует учитывать двухкомпонентность среды для S-осцилляторов.

В период закалки m -частиц параметрический потенциал $U(\bar{\eta})$ резко возрастает и далее медленно затухает по степенному закону (рис. 9.3):

$$U_T(> x_1) = \frac{2k_m^2}{y}, \quad U_S(> x_1) \simeq \frac{4k_m^2}{y}.$$

Поэтому осцилляторы попадают в параметрическую зону практически одновременно в момент времени $\bar{\eta} \simeq \bar{\eta}_1 \ll 1$, и в области перехода $s = \kappa^{-1}$, где малые фазовые функции задаются выражениями:

$$\kappa_T = \frac{k\bar{\eta}}{k_m}, \quad \kappa_S = \frac{k}{\sqrt{3}k_m} \int_0^{\bar{\eta}} \frac{d\eta}{\sqrt{1+3y}}.$$

Имея эти начальные условия, вычисляем константы C и интегральные функции $c(\bar{\eta})$ для T- и S-мод возмущений (см. (8.3.2), (8.8.11)):

$$C = \frac{ck_{\text{rad}}^2}{kR}, \quad c_T = 3^{-1/2}, \quad c_S = 2 \cdot 3^{1/4}, \quad (9.2.2)$$

$$\frac{k c_T}{k_m} = - \int \frac{d\bar{\eta}}{y^2} = \frac{1+2\bar{\eta}}{y} + 2 \ln \left(\frac{\bar{\eta}^2}{y\bar{\eta}_1} \right) - 1, \quad (9.2.3)$$

$$\frac{k c_S}{k_m \sqrt{3}} = - \int \frac{(1+4y)d\bar{\eta}}{y^2(1+3y)^2} = \frac{(1+2\bar{\eta})(1+6y)}{y(1+3y)} + 4 \ln \left(\frac{\bar{\eta}^2}{y\bar{\eta}_1} \right) - \frac{13}{4},$$

где мы воспользовались общей формулой:

$$\int \left(\frac{3\beta^2}{\gamma} - 1 \right) \frac{d\eta}{a^2} = \frac{1}{a^2 \gamma H} + \text{const}, \quad \beta^2 = \frac{p'}{\epsilon'}.$$

Переходя к пределу больших $\bar{\eta}$, получаем спектральные постоянные:

$$c_k = \bar{c} \frac{k_m}{k} \gg 1, \quad (9.2.4)$$

где

$$\bar{c}_T = -2 \ln \bar{\eta}_1 - 1 \simeq 2\mathcal{N}_1,$$

$$\bar{c}_S = -4\sqrt{3} \left(\ln \bar{\eta}_1 + \frac{13}{16} \right) \simeq 4\sqrt{3}\mathcal{N}_1,$$

$\mathcal{N}_1 = |\ln \bar{\eta}_1|$ — число хаббловских эпох между моментами отсоединения m -частиц и смены темпа расширения.

Подставляя вычисленные постоянные C и c_k в (8.3.12), получаем результирующие амплитуды скалярных и тензорных полей возмущений:

$$q_{k(\lambda)} \sim \frac{kk_m}{k_{\text{rad}}^2} \mathcal{N}_1 R \ll 1, \quad (9.2.5)$$

$$r \simeq 2\sqrt{3} > 1. \quad (9.2.6)$$

Как и ожидалось, сгенерированные спектры оказались голубыми с малыми амплитудами и доминированием тензорной моды. Отметим, что голубые спектры возникают всегда при резком возрастании потенциала U .

9.3. Массивное поле и излучение

Рассмотрим другую модель. Пусть наравне с радиацией в ранней Вселенной существовало скалярное поле φ с массой $m > 0$ и потенциальной энергией $m^2\varphi^2/2$, изначально находившееся в неравновесном состоянии и взаимодействовавшее с излучением гравитационным образом.

В контексте данной модели отличие этой постановки задачи от предыдущей состоит в том, что исходное значение слабодействующего поля большое и оно дрейфует в сторону квазиравновесного состояния осцилляций (частиц) в ходе гравитационной релаксации. Нас интересуют характеристики космологических возмущений, рожденных в результате гравитационного параметрического воздействия на однородную фридмановскую геометрию такого неравновесного состояния поля.

На стадии доминирования излучения член с трением в уравнении (7.4.3) велик ($H \gg m$), в результате чего поле находится в застывшем (закаленном) состоянии, где его кинетическая энергия мала по сравнению с потенциальной:

$$|\dot{\varphi}| \sim \frac{m^2\varphi_0}{H} \ll m\varphi_0 = \text{const} \quad (9.3.1)$$

(для определенности считаем $\varphi_0 > 0$). Закалка происходит при «случайном» (случившемся) значении $\varphi_0 > M_P/3$, связанном с физикой возникновения и условиями отсоединения поля от остальной материи на более ранних этапах эволюции. В таком состоянии ожидания поле пребывает вплоть до начальной фазы инфляции, когда плотность энергии излучения, затухающая из-за расширения сопутствующего объема, становится меньше застывшей плотности энергии поля (см. раздел 7.4).

Введем нормированный масштабный фактор $y = a/a_{i0}$, равный $\sqrt{2}$ в момент начала инфляции. Тогда в области перехода от радиационно-доминированной к инфляционной стадии имеем следующее решение уравнения Фридмана:

$$y = \sqrt{2 \operatorname{sh} 2x}, \quad H = H_{i0} \operatorname{cth} 2x, \quad \gamma = \frac{2}{\operatorname{ch}^2 2x}, \quad (9.3.2)$$

где $x \equiv H_{i0}(t + t_i)$ — нормированное время, а a_{i0} , t_i и

$$H_{i0} = \sqrt{\frac{4\pi G}{3}} m\varphi_0 = m\sqrt{\frac{2}{3}\mathcal{N}_0} \quad (9.3.3)$$

— постоянные величины.

На стадии инфляции поле начинает чувствовать собственную гравитацию, «размораживается» и эволюционирует к состоянию своего устойчивого равновесия $\varphi = \dot{\varphi} = 0$.

Процесс релаксации поля можно разбить на ряд этапов. Сначала, при $1 < x < x_m = 2(\mathcal{N}_0 - \sqrt{\mathcal{N}_0/3})$, поле дрейфует вдоль инфляционной сепаратрисы в режиме медленного скатывания от φ_0 к $\varphi_m \simeq M_P/\sqrt{6\pi}$:

$$\varphi = \varphi_0 \left(1 - \frac{x}{2\mathcal{N}_0}\right) = M_P \sqrt{\frac{\mathcal{N}}{2\pi}}, \quad (9.3.4)$$

$$y = \exp \left[x \left(1 - \frac{x}{4\mathcal{N}_0}\right) \right] = \exp [\mathcal{N}_0 - \mathcal{N}], \quad (9.3.5)$$

$$H = m\sqrt{\frac{2}{3}\mathcal{N}}. \quad (9.3.6)$$

В качестве переменной времени мы использовали число инфляционных хаббловских эпох, отсчитываемых от момента окончания инфляции ¹⁾:

$$\mathcal{N} = 2\pi G \varphi^2 \simeq \frac{1}{3} + \ln \frac{a_m}{a}, \quad (9.3.7)$$

$$\mathcal{N}_0 = 2\pi G \varphi_0^2 = 2\pi \frac{\varphi_0^2}{M_P^2} > 1. \quad (9.3.8)$$

Затем, при $x \gtrsim x_m$, эволюция переходит в стадию осцилляций с эффективным уравнением состояния $p = 0$.

В ходе осцилляций происходит потеря общей фазы колебаний из-за пространственно-некоррелированных флуктуаций, ко-

¹⁾ Для удобства представления (9.3.4)–(9.3.6) в (7.4.12) введена постоянная интегрирования $1/3$ (см. (9.3.7)). Решение (9.3.4)–(9.3.8) естественно продолжается за пределы своего определения с помощью фридмановского уравнения:

$$H^2 = \frac{4\pi G}{3} (m^2\varphi^2 + \dot{\varphi}^2) + \frac{4H_{i0}^2}{y^4}.$$

При $x \ll x_m$ оно стыкуется с (9.3.2), а при $x \gtrsim x_m$ переходит в осцилляции:

$$\varphi = \frac{M_P}{\sqrt{3\pi}} \frac{\cos z}{z}, \quad y = y_m \left(\frac{z}{\sqrt{2}} \right)^{2/3}, \quad H = \frac{2m}{3z},$$

где $z = m(t - t_m) + \sqrt{2}$, $y_m = \exp(\mathcal{N}_0 - 1/3)$.

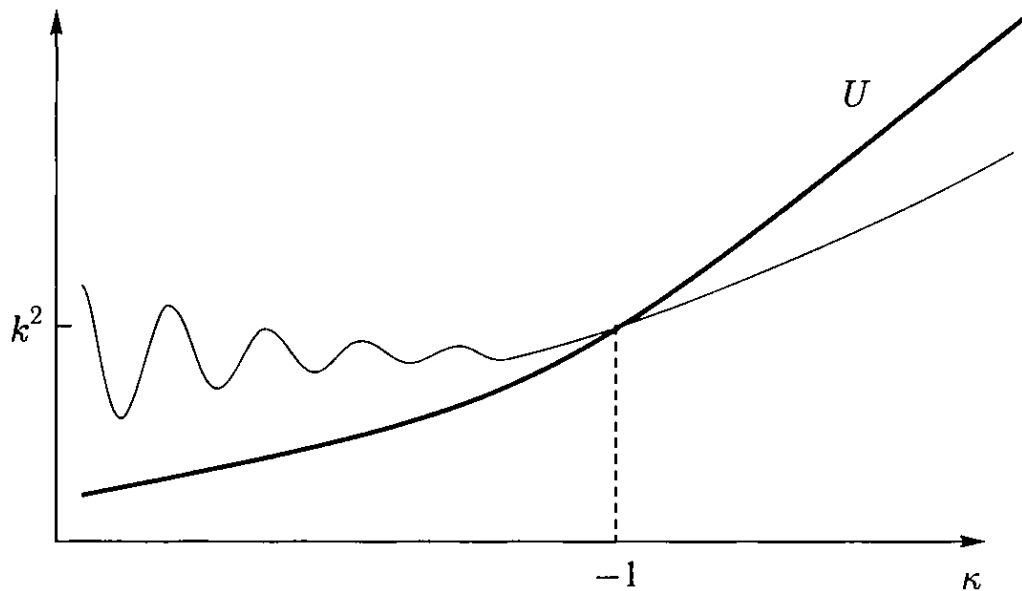


Рис. 9.4. Параметрический потенциал гравитационных волн $U(\kappa)$ на стадии $\gamma \ll 1$ в модели с массивным полем и излучением. Красной линией схематично показано поведение элементарного осциллятора $\bar{\nu}_{k\lambda}(\kappa)$ (см. (9.3.12))

герентность состояния поля нарушается и Вселенная эволюционирует к стадии доминирования тяжелых частиц поля φ . Далее нерелятивистские частицы распадаются на легкие, продукты распада термализуются и восстанавливается радиационная динамика расширения.

Таким образом, инфляция существует ограниченный период времени, пока $H > m$, и заканчивается при $H = m$, имея конечное число хаббловских эпох \mathcal{N}_0 .

У нас имеются две характеристики инфляционной стадии: продолжительность \mathcal{N}_0 и энергия E в момент завершения инфляции:

$$E = \sqrt{mM_P} \quad (9.3.9)$$

(по определению, полная плотность энергии равна E^4). Энергетический масштаб инфляции (9.3.9) связан с массой инфлатона m , а параметр (9.3.8) — с его закаленной величиной φ_0 (или, эквивалентно, с H_{i0}/m): при условии $\varphi_0 > M_P$ (или $H_{i0} > m$) $\mathcal{N}_0 > 1$. Что касается негравитационных взаимодействий поля φ с другими частицами и полями, то в данной модели они должны быть достаточно слабыми, чтобы распад инфлатона произошел не ранее окончания инфляционной стадии и перехода ее в стадию частиц.

Рассмотрим динамику элементарных осцилляторов.

Начнем с моды Т. На стадии инфляции $\gamma \ll 1$ барьер имеет простой вид (рис. 9.4):

$$U_T = \frac{a''}{a} \simeq \frac{2k^2}{\kappa^2}, \quad (9.3.10)$$

где

$$a \simeq -\frac{k}{\kappa H}, \quad \kappa = k(\eta + \eta_i) \simeq -\frac{k}{H},$$

$\eta_i = \text{const}$. Он растет со временем и пересекает уровни частот k^2 в *разные* моменты времени $\kappa_1 = -1$. Величину k удобно связать с числом инфляционных хаббловских эпох $\mathcal{N} = \mathcal{N}_1$, прошедших с момента попадания k -осциллятора в параметрическую зону:

$$k = k_m \sqrt{\mathcal{N}} e^{-\mathcal{N}} \in (k_0, k_m),$$

где

$$\begin{aligned} k_0 &= \bar{H}_{i0} = k_m \sqrt{\mathcal{N}_0} e^{-\mathcal{N}_0}, \\ k_m &= m a_m \equiv m e^{-N_m}, \end{aligned} \quad (9.3.11)$$

N_m — количество хаббловских эпох, прошедших с момента окончания инфляции до сегодняшнего дня.

В области перехода в параметрическую зону (9.3.10) уравнение (8.2.8) решается точно:

$$\bar{v}_{k\lambda} = \frac{e^{-i\kappa}}{\sqrt{2\omega}} \left(1 + \frac{1}{i\kappa}\right) = -\frac{i + \kappa^3/3}{\kappa\sqrt{2\omega}} (1 + O(\kappa^2)), \quad (9.3.12)$$

где при $|\kappa| < 1$ это решение отождествляется с длинноволновой асимптотикой (8.3.2) с параметрами:

$$C = \frac{k \alpha_T}{H_1}, \quad (9.3.13)$$

$$c(\kappa) = -\frac{\kappa^3}{3} \left(\frac{H}{H_1}\right)^2, \quad c_k = 0. \quad (9.3.14)$$

Подставляя полученные константы в (8.3.12), получаем спектр мощности гравитационных волн для $k \in (k_0, k_m)$:

$$\sqrt{T} = \frac{2H}{\sqrt{\pi} M_P} \simeq \frac{m\sqrt{\mathcal{N}}}{M_P} \simeq \frac{m}{M_P} \ln^{1/2} \left(\frac{k_m}{k} \sqrt{\ln \frac{k_m}{k}} \right). \quad (9.3.15)$$

Это слабокрасный спектр с насыщением до плоского в длинноволновой области $k \sim k_0$.

При частотах, меньших чем k_0 , Т-спектр трансформируется в голубой:

$$k < k_0: \quad \sqrt{T} \sim \frac{H_{i0}}{M_P} \frac{k}{k_0}. \quad (9.3.16)$$

Чтобы убедиться в этом, обратимся к (9.3.2) и вычислим эффективный потенциал гравитационных волн:

$$y < 1: \quad U_T = 2k_0^2 y^2. \quad (9.3.17)$$

В переходной области $y \sim k/k_0 < 1$ определена малая фазовая функция

$$\kappa = \frac{k\sqrt{x}}{k_0} \ll 1,$$

позволяющая прямо отождествить константы параметрической зоны с помощью точного решения (8.8.10):

$$c_1 = \frac{1}{\kappa_1}, \quad C = \frac{M_P k_0^2}{\sqrt{2\pi} H_{i0} k}. \quad (9.3.18)$$

В результате получаем функцию $c(x)$:

$$c = -\frac{4k_0}{k} \int \frac{dx}{y^3} = \frac{2k_0}{yk} + \frac{k_0}{k} \int_1^{\sqrt{\text{th } x}} \frac{dx}{\text{ch } x}$$

и необходимую для определения спектра константу c_k :

$$c_k = \frac{c_0 k_0}{k}, \quad c_0 \simeq \int_0^{\infty} \frac{\sqrt{\text{th } x}}{\text{ch } x} dx \sim 1, \quad (9.3.19)$$

откуда и следует (9.3.16).

9.4. Спектр возмущений плотности

Спектр возмущений плотности, генерируемый в модели массивного поля и излучения, качественно напоминает T-спектр с той лишь разницей, что он несколько выше и его максимум смещен в сторону коротких длин волн.

Параметрические потенциалы S- и T-мод космологических возмущений представлены на рис. 9.5.

Критическим для S-спектра является поведение функции γ при $y < y_m$ (см. (9.3.4)–(9.3.6)):

$$\gamma = \frac{2 \left(1 + \frac{y^4}{16\mathcal{N}_0} \right)}{1 + y^4 \left(\frac{\varphi}{2\varphi_0} \right)^2}. \quad (9.4.1)$$

Наибольший интерес вызывает ее инфляционная часть $1 < y < y_m$, где $\ln y = \mathcal{N}_0 - \mathcal{N} \simeq x$:

$$\frac{1}{2} < \mathcal{N} < \mathcal{N}_0 - \frac{1}{2}: \quad \gamma = \frac{1 + 16\mathcal{N}_0 e^{4(\mathcal{N}-\mathcal{N}_0)}}{2\mathcal{N}} < 1.$$

В начальный период инфляции $1 < y < y_c \equiv 2\mathcal{N}_0^{1/4}$ функция γ убывает до значения меньше единицы и экспоненциально затухает из-за красного смещения радиации, достигая затем насыщения и своих минимальных значений в области $y \in (y_c, y_c^2)$:

$$\begin{aligned} 1 < y < y_c^2: \quad \gamma &\simeq \frac{1 + (y_c/y)^4}{2\mathcal{N}_0}, \\ \gamma_{\min} &\simeq \frac{1}{2\mathcal{N}_0} = \frac{M_P^2}{4\pi\varphi_0^2}. \end{aligned} \quad (9.4.2)$$

Далее под влиянием поля φ она растет степенным образом, достигая прежнего уровня $\gamma = 1$ к концу инфляции:

$$y_c^2 < y < y_m: \quad \gamma \simeq \frac{1}{2\mathcal{N}}. \quad (9.4.3)$$

Этой аппроксимацией можно пользоваться в расширенной области $y \in (y_c, y_m)$, поскольку вблизи минимума γ величина $\mathcal{N} \simeq \mathcal{N}_0 = \text{const}$.

На поздней инфляционной стадии $y \in (y_c, y_m)$ функция γ изменяется адиабатически ($|\Gamma| < 1$), вследствие чего применимы все формулы (9.3.10)–(9.3.13) с заменой α_T на α_S . В результате получаем S-спектр в диапазоне масштабов $k \in (k_c, k_m)$, где $k_c = k_0 y_c$:

$$\sqrt{S} = \frac{H}{\sqrt{\pi\gamma} M_P} \simeq \frac{m\mathcal{N}}{M_P} \simeq \frac{m}{M_P} \ln \left(\frac{k_m}{k} \sqrt{\ln \frac{k_m}{k}} \right). \quad (9.4.4)$$

На начальной инфляционной стадии $y \in (1, y_c)$ функция γ эволюционирует в неадиабатическом режиме, в результате чего параметрический потенциал скалярных возмущений экспоненциально подавлен:

$$U_S = \frac{6k_0^2 y^6}{y^4 + y_c^4} \left(1 - \frac{1}{3\gamma\mathcal{N}_0} \right) \simeq \frac{2k_0^2 y^6 (y^4 + 3y_c^4)}{(y^4 + y_c^4)^2}. \quad (9.4.5)$$

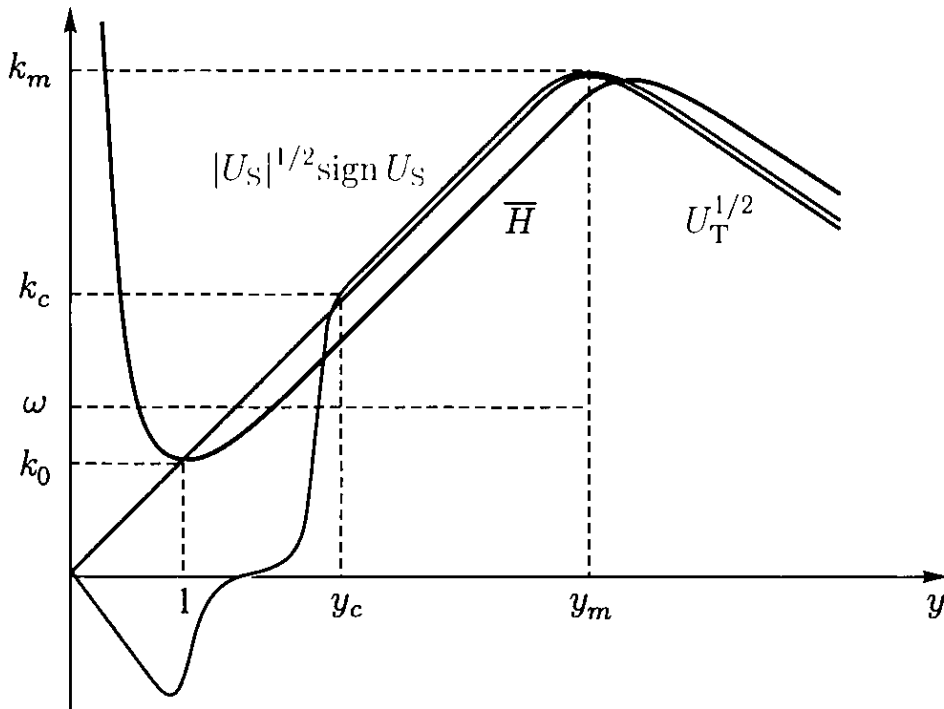


Рис. 9.5. Параметрические потенциалы скалярных возмущений $U_S(y)$ (зеленая линия), гравитационных волн $U_T(y)$ (красная линия) и хаббловская функция $\bar{H}(y)$ (синяя линия) в модели с массивным полем и излучением

Генерируемые гармоники S-моды в области частот $k \in (k_0, k_c)$ попадают в параметрическую зону в период времени $y \in (y_c^{2/3}, y_c)$, на котором

$$\frac{\alpha_S}{\alpha_T} = \sqrt{2\gamma} \simeq \frac{4}{y^2}, \quad U_S \simeq 6k_c^2 \left(\frac{y}{y_c}\right)^6 < \bar{H}^2 \quad (9.4.6)$$

(рис. 9.5). Как мы уже знаем, при столь резком подъеме потенциала формируются голубые спектры мощности.

Действительно, всегда в этом случае мы можем определить малую фазовую функцию

$$-\kappa = \frac{k}{k_0 y} = \frac{k}{\bar{H}} < 1,$$

которая дает возможность «прошить» интересующее нас решение уравнения (8.2.8) из волновой в параметрическую зону и найти спектральные постоянные (ср. (8.8.10)):

$$C = \frac{\bar{\alpha}_S}{|\kappa|} = \frac{4\alpha_T k_0^2}{H_{i0} k}, \quad c_1 = \frac{1}{\kappa_1}. \quad (9.4.7)$$

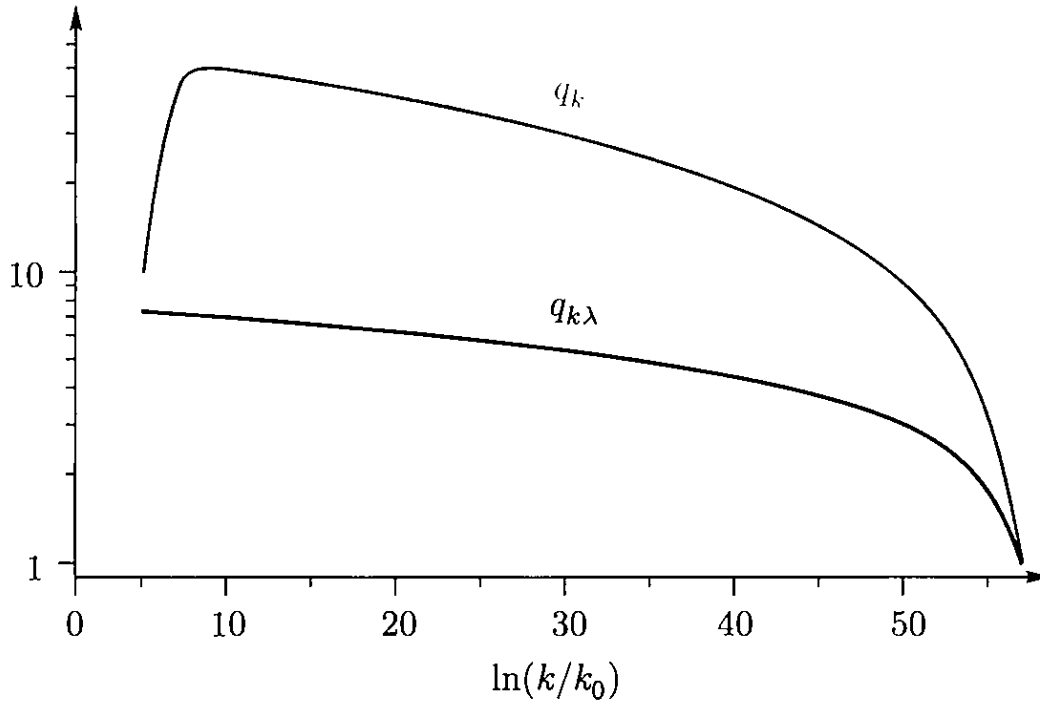


Рис. 9.6. Спектры скалярных (красная линия) и тензорных (черная линия) возмущений в интервале $k \in (k_0, k_m)$ в единицах m/M_P (здесь $\mathcal{N}_0 = 55$)

Отсюда находим функцию

$$c(x) = -\frac{kC^2}{k_0} \int \frac{dx}{y\bar{\alpha}_S^2} = -\frac{k_0}{k} \int_0^{\infty} \frac{dy}{1 + (y/y_c)^4} \quad (9.4.8)$$

и спектральную константу

$$c_k = -\frac{c_0 k_c}{k}, \quad c_0 = \int_0^{\infty} \frac{dx}{1 + x^4} \sim 1. \quad (9.4.9)$$

Итак, у нас имеются все необходимые параметры для определения спектра мощности. Собирая их в формулу (8.3.12), получаем S-спектр в диапазоне частот $k \in (k_0, k_c)$:

$$q_k \sim \frac{H_{i0} \sqrt{\mathcal{N}_0} k}{M_P k_c} \simeq \frac{m \mathcal{N}_0 k}{M_P k_c}. \quad (9.4.10)$$

Полный спектр S-моды представлен на рис. 9.6. При малых частотах $k < k_0$ спектр падает более круто, чем (9.4.10), однако вычисления в этой зоне выходят за рамки рассматриваемой модели, так как начальные условия при $k < k_0$ приходится задавать в причинно-несвязанной области пространства. Подчеркнем в этой связи обоснованность голубого участка спектра (9.4.10), поскольку затравочные возмущения для него задаются в *микроскопической* зоне $y \sim 1$ (рис. 9.5).

Подводя итоги, мы можем выписать аппроксимацию спектров космологических возмущений в интервале $k \in (k_0, k_m)$ следующим образом (рис. 9.6):

$$q_{k\lambda} \simeq \frac{m\sqrt{\mathcal{N}}}{M_P}, \quad n_T \simeq -\frac{1}{\mathcal{N}}, \quad (9.4.11)$$

$$q_k \simeq \frac{m\mathcal{N}}{M_P \sqrt{1 + (k_c/k)^2}}, \quad (9.4.12)$$

$$n_S \simeq 2 \left(\frac{k_c^2 - k^2/\mathcal{N}}{k_c^2 + k^2} \right), \quad (9.4.13)$$

$$r \simeq \frac{2}{\mathcal{N}} \left(1 + \frac{k_c^2}{k^2} \right), \quad (9.4.14)$$

$$\mathcal{N} \simeq 1 + \ln \left(\frac{\sqrt{\mathcal{N}} k_m}{k} \right),$$

где $k_c = 2k_0 \mathcal{N}_0^{1/4}$, $\mathcal{N}_0 - 1 \simeq \ln(\sqrt{\mathcal{N}_0} k_m/k_0)$ — полное число хаббловских эпох на стадии инфляции (см. сноску на с. 256).

Перечислим наиболее важные выводы данной модели.

- Величина спектральной амплитуды S-моды выше, чем амплитуда гравитационных волн, и находится в интервале:

$$\frac{m}{M_P} < q_k < \frac{m\mathcal{N}_0}{M_P}.$$

Полагая $\mathcal{N}_0 \sim 100$ и $q_k \sim 10^{-5}$, получаем массу инфлатона в области, примыкающей к энергии Великого объединения:

$$10^{12} < \frac{m}{\text{ГэВ}} < 10^{14}.$$

- Отношение $r = T/S$ ограничено снизу:

$$r > \frac{2}{\mathcal{N}_0},$$

что повышает шансы обнаружения космологических гравитационных волн. При $\mathcal{N}_0 \sim 100$ их амплитуда всего лишь в семь раз меньше амплитуды скалярных неоднородностей ($\sqrt{r} \sim 1/7$).

- Спектр S-моды растет логарифмически с увеличением масштаба и достигает широкого максимума при $k_{\max} \simeq k_c \sqrt{\mathcal{N}_0}$, а затем при дальнейшем увеличении масштаба ($k < k_c$) резко затухает степенным образом. Этот эффект способен объяснить аномально малую величину наблюдаемой квадрупольной анизотропии реликтового излучения. Например,

при $\mathcal{N}_0 \sim 70$ и $k_m^{-1} \sim 1$ см, масштаб $k_c^{-1} \sim 3$ Гпк находится в пределах современного хаббловского радиуса.

- В красной области спектра выполняются равенства:

$$r = \frac{2}{\mathcal{N}} = -2n_T = -n_S = \sqrt{-2\tau_S},$$

которые можно проверить экспериментально. Так, при $n_S = -0,04$ и $n = 0,96$ получаем:

$$r = 0,04, \quad \tau_S = -8 \cdot 10^{-4}, \quad \mathcal{N} = 50 < \mathcal{N}_0,$$

что близко к современным экспериментальным возможностям обнаружения.

9.5. Самодействие и Λ -инфляция

Исследованные нами простые динамические модели демонстрируют основные свойства генерируемых неоднородностей метрики. Возникающие формы спектров возмущений достаточно разнообразны, но приемлемые амплитуды S-моды реализуются лишь при $s_1 \sim 1$ ($\bar{H}_1 \sim \omega$), что возможно в условиях инфляции.

Простейшей причиной инфляции является скалярное массивное поле, взаимодействующее с остальной материей гравитационным образом. Однако реальное поле, определившее инфляцию в ранней Вселенной, могло быть сложнее, например нелинейным, самодействующим, комбинированным и пр. В данном разделе мы рассмотрим свойства космологических возмущений, генерируемых нелинейным и/или самодействующим инфлатоном.

Лагранжева плотность произвольного скалярного поля может быть представлена в виде функции двух скалярных переменных $\mathcal{L} = \mathcal{L}(w, \varphi)$ — самого поля φ и его кинетической части $w^2 \equiv \varphi_{,\mu} \varphi^{,\mu}$ (см. (7.1.9)). Нас интересуют режимы доминирования плотности энергии, связанные с неравенством $\gamma < 1$. С помощью (7.1.10) перепишем это условие в следующем виде:

$$\gamma = \frac{3w^2}{2w^2 - \mathcal{L}/\mathcal{L}_{,w^2}}, \quad w^2 < -\frac{\mathcal{L}}{\mathcal{L}_{,w^2}}, \quad (9.5.1)$$

где частная производная берется при постоянном φ .

Подавление кинетического скаляра w^2 означает, что динамика контролируется первыми исчезающими членами разложения \mathcal{L} в ряд Тейлора по малому параметру w^2 . Полагая для

простоты $\beta = 1$ (см. (7.1.11)), получаем следующее выражение¹⁾:

$$\mathcal{L} = \mathcal{L}_0(\varphi) + \mathcal{L}_1(\varphi) w^2 + \dots, \quad (9.5.2)$$

где $\mathcal{L}_1 > 0$ (отсутствие тахионной неустойчивости). Ограничиваясь первыми двумя членами разложения и переопределяя полевую переменную ($\int \sqrt{\mathcal{L}_1} d\varphi \rightarrow \varphi$), приходим к стандартной форме лагранжевой плотности скалярного поля:

$$\mathcal{L} = \frac{w^2}{2} - V(\varphi), \quad \gamma = \frac{3w^2}{w^2 + 2V}, \quad (9.5.3)$$

где $V = V(\varphi)$ — потенциальная энергия. Состояние поля φ зависит от соотношения между его кинетической и потенциальной энергиями. Так, при доминировании кинетического члена получаем предельно жесткое уравнение состояния $p \simeq \varepsilon$ ($\gamma = 3$); на стадии частиц имеем паритет между кинетической и потенциальной энергиями ($\gamma = 3/2$, $\langle p \rangle = 0$), а в режиме инфляции $w^2 < V$.

Релаксация подобных полей рассмотрена в разделе 7.4.

Для массивного поля $V = m^2 \varphi^2 / 2$. Прямым обобщением такого потенциала является степенная функция

$$V = \frac{\lambda}{2p} \varphi^{2p}, \quad (9.5.4)$$

где p — натуральное число, λ — положительная постоянная.

На адиабатическом этапе инфляционной стадии:

$$\gamma = \frac{p}{2N} \quad \text{и} \quad U = \frac{2\omega^2}{\kappa^2}.$$

Следовательно, выполняются все формулы (9.3.10)–(9.3.14) и спектры генерируемых возмущений имеют вид:

$$q_{k\lambda} \simeq \frac{H}{M_P}, \quad (9.5.5)$$

$$q_k \simeq \frac{H}{M_P} \sqrt{\frac{\mathcal{N}}{p}}, \quad (9.5.6)$$

$$r = \frac{2p}{N} = -2n_T = -\frac{2p}{p+1} n_S, \quad (9.5.7)$$

где

$$H \propto \mathcal{N}^{p/2}, \quad \mathcal{N} = \frac{2\pi G \varphi^2}{p} = \ln \left(\frac{k_m}{k} \right) + \frac{p}{2} \ln \mathcal{N}.$$

¹⁾ При адиабатической скорости звука $\beta^{-2} \simeq \text{const} \geq 1$ второй исчезающий член разложения в правой части (9.5.2) пропорционален $w^{1+\beta^{-2}}$.

Полученные спектры слабокрасные, однако они более круто спадают с ростом k при $p > 1$ по сравнению с $p = 1$. Это обстоятельство сильно ограничивает возможность реализации степенных потенциалов в ранней Вселенной.

Принимая во внимание наблюдательные ограничения: $r < 0,1$ и $n_S > -0,05$ в области масштабов $k \sim 0,01 \text{ Мпк}^{-1}$ ($\mathcal{N} \lesssim 60$), получаем жесткое ограничение на показатель степенной инфляции:

$$p < 2, \quad (9.5.8)$$

которое оставляет жизнеспособным только инфляцию на массивном поле (подробнее см. раздел 10.5).

Этот пример показывает, что потенциалы $V(\varphi)$, достаточно круто растущие с ростом инфлатона (в среднем быстрее, чем φ^2), не могли быть причиной рождения космологических возмущений плотности. Однако остается свобода в выборе более пологих потенциалов, генерирующих практически любые формы S-спектров: ведь мы имеем возможность определить не просто отдельные параметры теории, а целую функцию $V(\varphi)$, хотя и слабо ограниченную.

Мы уже отмечали, что тензорная мода более фундаментальна, чем скалярная, поскольку зависит от меньшего количества параметров (см. (8.3.14)). На инфляционной сепаратрисе, независимо от того, адиабатическая она или нет, всегда справедливо решение (9.3.10)–(9.3.15), дающее общий вид T-спектра:

$$T = \frac{4H^2}{\pi M_P^2} \simeq \frac{10V}{M_P^4}, \quad n_T = -2\gamma, \quad (9.5.9)$$

где

$$\frac{k}{k_m} = \frac{\bar{H}}{\bar{H}_m} = \sqrt{\frac{V}{V_m}} \exp\left(-\frac{8\pi}{M_P^2} \int_m^{\varphi} \frac{V d\varphi}{V_{,\varphi}}\right),$$

$V_m = V(\varphi_m)$. Мы видим, что амплитуда и наклон T-спектра прямо связаны соответственно с хаббловской и γ -функцией в эпоху генерации возмущений.

Общий рецепт для определения S-спектра можно дать только при адиабатическом поведении функций β и γ . Тогда задача сводится к предыдущей, и мы приходим к следующим выражениям:

$$S = \frac{H^2}{\pi\beta\gamma M_P^2} \simeq \frac{425 V^3}{\beta M_P^6 V_{,\varphi}^2}, \quad (9.5.10)$$

$$r = 4\gamma\beta, \quad (9.5.11)$$

$$n_S = 2\epsilon - 6\gamma + \sqrt{\frac{\gamma}{4\pi}} \frac{M_P \beta_{,\varphi}}{\beta}, \quad (9.5.12)$$

где

$$\sqrt{\pi\gamma} = \frac{M_P}{4V} \frac{dV}{d\varphi}, \quad \epsilon = \frac{M_P^2}{8\pi V} \frac{d^2V}{d\varphi^2}. \quad (9.5.13)$$

При $\beta = 1$ S-спектр зависит от *двух* параметров ранней Вселенной (H и γ) а его наклон — от *трех* (H , γ и ϵ). Зная только амплитуду S-моды, мы имеем одно алгебраическое соотношение на два параметра Большого взрыва (H и γ). Чтобы решить проблему ранней Вселенной, надо знать как минимум амплитуду T-моды, тогда число уравнений совпадет с числом неизвестных.

Полученные S-спектры характерны для поздней динамики инфлатона. T-спектр всегда красный, тогда как S-спектр может быть и красным, и строго плоским, и голубым. Мы уже встречали S-спектры в форме горы (рис. 9.6). Приведем пример другой формы спектра — в виде ложбины.

Речь пойдет о потенциале $V(\varphi)$ с положительной остаточной энергией $V_0 > 0$ в точке локального минимума $\varphi = 0$ (ср. (9.5.4)):

$$V = V_0 + \frac{\lambda}{2p} \varphi^{2p}. \quad (9.5.14)$$

Этот вариант носит название «гибридной» инфляции или Λ -инфляции, поскольку речь идет о дополнительном параметре энергии, напоминающем вакуумный Λ -член. Как и сам инфлатон, константа V_0 метастабильна, однако она вполне могла существовать какое-то время и задержать релаксацию инфлатона. Нам интересно знать, какие характерные следы такой задержки останутся в спектрах космологических возмущений.

Для решения этой задачи воспользуемся формулами (9.5.9), (9.5.10). Поскольку все эффекты можно пояснить на простом примере, например для $p = 2$, ограничимся рассмотрением потенциала (см. (9.5.14)):

$$V = V_0 \left(1 + \frac{\varphi^4}{c^4} \right),$$

где при значении поля $\varphi = c \equiv (4V_0/\lambda)^{1/4}$ его потенциальная энергия становится равной энергии метастабильного вакуума. Вся задача характеризуется одной безразмерной постоянной

$$\epsilon_c = \frac{M_P^2}{2\pi c^2}. \quad (9.5.15)$$

Условие медленного скатывания (7.4.9) ограничивает константы теории:

$$\epsilon_c < 1, \quad c > \frac{M_P}{\sqrt{2\pi}}. \quad (9.5.16)$$

В этом случае справедливы формулы (9.5.9), (9.5.10), и спектры возмущений имеют вид:

$$S = \frac{4V_0}{3\epsilon_c M_P^4} \left(\frac{c^2}{\varphi^2} + \frac{\varphi^2}{c^2} \right)^3, \quad (9.5.17)$$

$$T = \frac{32V_0}{3M_P^4} \left(1 + \frac{\varphi^4}{c^4} \right), \quad (9.5.18)$$

$$r = 8\epsilon_c \left(\frac{c\varphi^3}{c^4 + \varphi^4} \right)^2 \leq 2,5\epsilon_c, \quad (9.5.19)$$

где поле связано с волновым числом следующим образом:

$$\frac{\varphi^2}{c^2} - \frac{c^2}{\varphi^2} = 2\epsilon_c \ln \left(\frac{k_c}{k} \sqrt{\frac{1}{2} + \frac{\varphi^4}{2c^4}} \right); \quad (9.5.20)$$

$k_c = \text{const}$ соответствует значению $\varphi = c$. Решая это уравнение, получаем явный вид спектра возмущений плотности:

$$S = \frac{32V_0}{3\epsilon_c M_P^4} \left[1 + \epsilon_c^2 \ln^2 \left(\frac{k}{k_c} \right) \right]^{3/2}, \quad (9.5.21)$$

$$n_S = \frac{3\epsilon_c^2 \ln \left(\frac{k}{k_c} \right)}{1 + \epsilon_c^2 \ln^2 \left(\frac{k}{k_c} \right)}, \quad (9.5.22)$$

$$r_S = 3\epsilon_c^2 \frac{1 - \epsilon_c^2 \ln^2 \left(\frac{k}{k_c} \right)}{\left[1 + \epsilon_c^2 \ln^2 \left(\frac{k}{k_c} \right) \right]^2}. \quad (9.5.23)$$

Оба спектра S- и T-мод представлены на рис. 9.7.

S-спектр имеет форму «ласточки» с минимумом в точке $k = k_c$ ($\varphi = c$), где энергия поля и Λ -члена становятся равными между собой. Параметр r в этой области близок к своему максимуму, который порядка единицы при $c \sim 1$ и уменьшается обратно пропорционально c^2 (см. (9.5.15), (9.5.19)). Наиболее интересна голубая часть S-спектра с умеренным наклоном $n_S > 0$ и небольшим фактором r . Такие свойства спектра не противоречат имеющимся наблюдательным ограничениям и подлежат экспериментальной проверке (подробнее см. раздел 10.5).

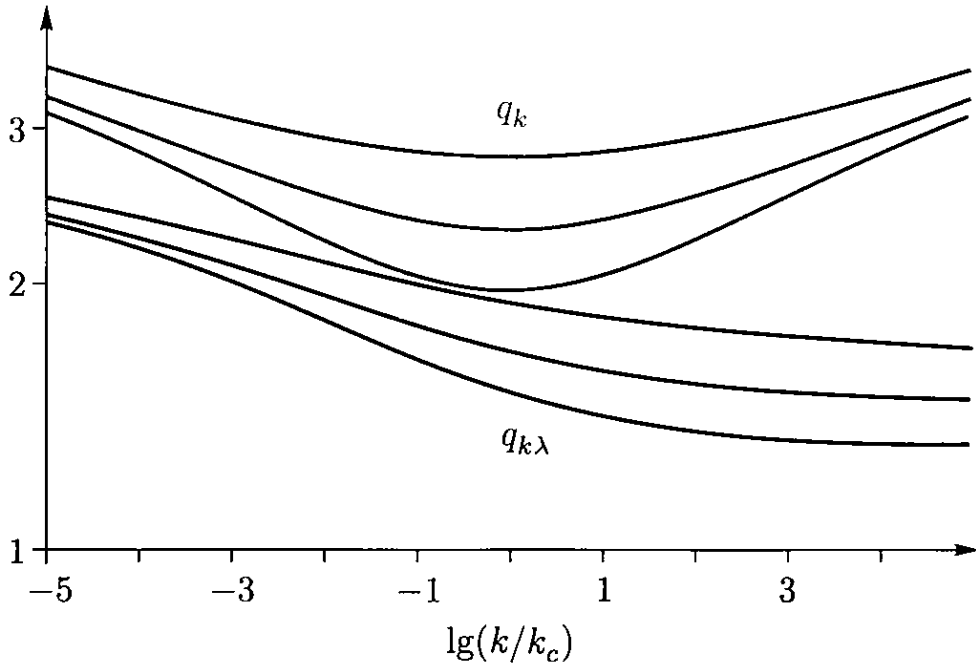


Рис. 9.7. Спектры скалярных возмущений (красные линии) и гравитационных волн (черные линии) в модели Λ -инфляции в произвольной нормировке для значений $c/M_P = 5, 6, 7$ (снизу вверх)

9.6. Создание квазиизотропной геометрии

Для того чтобы запустить инфляцию, необходимо придать полю большое значение, но заведомо меньшее, чем то, что соответствует квантовой границе:

$$1 < \varphi < m^{-1};$$

здесь для оценок мы используем массивное поле, все величины нормированы на планковскую массу.

Напомним, что по определению (см. (9.3.9))

$$E^4 = \frac{\dot{\varphi}^2}{2} + \frac{m^2 \varphi^2}{2}, \quad E < 1.$$

В этой области генерируемые тензорные неоднородности кривизны малы ($q_\lambda \sim m\varphi < 1$), однако скалярные неоднородности оказываются малыми лишь при значительно меньших энергиях:

$$E < E_c \equiv m^{1/4}, \quad |\varphi| < m^{-1/2}.$$

В этой зоне $q \sim m\varphi^2 < 1$ и можно корректно определить классическую часть поля и инфляционную сепаратрису (рис. 7.3).

Инфляционных стадий в ранней Вселенной могло быть несколько в зависимости от количества полей, их параметров и общей структуры потенциала $V(\varphi_1, \varphi_2, \dots)$. Для наблюдательной космологии важна заключительная стадия инфляции

(с параметрами m и \mathcal{N}_0), обеспечившая требуемую величину S -моды возмущений в видимой Вселенной.

Что касается глобальной квазифридмановской области, то в ее образовании участвовали *все* стадии инфляции, начиная от энергии E_c и заканчивая $E_m \equiv \sqrt{m}$ (см. раздел 7.4). В ходе инфляционной стадии энергия полей затухала, приводя к падению амплитуды рождающихся неоднородностей. В результате в масштабах (см. (9.3.11))

$$\bar{H}_c < k < \bar{H} < k_m \quad (9.6.1)$$

формировался локально изотропный квазиоднородный поток материи с пространственными неоднородностями кривизны, растущими с увеличением сопутствующего размера ($q \sim m\mathcal{N}$).

Процесс создания квазифридмановской области прекратился с окончанием последней инфляции, и ее размер оказался большим, но конечным:

$$\ell_F = \bar{H}_c^{-1} \sim H_0^{-1} e^{\mathcal{N}_\Sigma - 50} \gg H_0^{-1}, \quad (9.6.2)$$

где H_0^{-1} — хаббловский радиус современной Вселенной, \mathcal{N}_Σ — суммарное количество хаббловских эпох на всех стадиях инфляции (включая \mathcal{N}_0), фактор 50 в показателе экспоненты связан с наблюдаемой амплитудой космологических возмущений плотности $q \sim 10^{-5} \sim 50m$. Из (9.6.2) следует, что мы наблюдаем лишь часть квазифридмановской геометрии в пределах хаббловского радиуса. При этом величина \mathcal{N}_0 никак не ограничена и подлежит экспериментальному определению.

Квазифридмановская область представляет собой расщепление $1 + 3$ с глобально однородной пространственноподобной гиперповерхностью $t = \text{const}$, на которой скаляры H , γ , ϵ и p зависят только от времени (с точностью до членов $O(k^2/\bar{H}^2)$), а масштабный фактор имеет явную зависимость от координат (см. (2.7.11)):

$$a = a(t) \cdot \bar{a}(\mathbf{x}). \quad (9.6.3)$$

Застывший гравитационный потенциал $\bar{a}(\mathbf{x})$ описывает одну из реализаций случайного поля рожденных неоднородностей. Здесь под «точкой» \mathbf{x} надо понимать окрестность *конечного* размера, поэтому значение функции $\bar{a}(\mathbf{x})$ в любой точке зависит от того, по какому размеру мы произвели сглаживание. Для малых амплитуд неоднородностей вокруг точки наблюдателя имеем:

$$\bar{a}(\mathbf{x}) = \bar{a}_0 \cdot (1 - q(\mathbf{x})), \quad \bar{a}_0 = \text{const}. \quad (9.6.4)$$

Рассмотрим процесс создания классической квазиизотропной гиперповерхности (9.6.3), (9.6.4).

В каждый момент времени она определена в масштабе больше текущего инфляционного хаббловского радиуса:

$$r \equiv |\mathbf{x}_1 - \mathbf{x}_2| \geq \sigma \equiv \frac{c}{H} \simeq -c\eta, \quad (9.6.5)$$

где численный фактор

$$c \sim 10 \gg 1.$$

Поэтому для ее описания нам потребуется крупномасштабная часть поля q .

Разобьем полное поле q на две части — крупномасштабную Q и все остальное:

$$q = Q + \text{остаток}. \quad (9.6.6)$$

Поле Q мы будем считать классическим, а остаток — квантовым.

Процедура разбиения квантового оператора на крупномасштабную часть и квантовый остаток хорошо известна. Для этого сгладим поле q и импульс \dot{q} с помощью пространственного гауссова фильтра, характеризуемого σ :

$$Q(t, \mathbf{x}) = \int K_\sigma(|\mathbf{x} - \mathbf{x}'|) q(t, \mathbf{x}') d\mathbf{x}', \quad (9.6.7)$$

$$P(t, \mathbf{x}) = \int K_\sigma(|\mathbf{x} - \mathbf{x}'|) \dot{q}(t, \mathbf{x}') d\mathbf{x}', \quad (9.6.8)$$

где

$$K_\sigma(r) = (2\pi)^{-3/2} \sigma^{-3} \exp\left(-\frac{r^2}{2\sigma^2}\right).$$

Мы получили крупномасштабную часть кривизны. Что же остается в остатке?

Чтобы ответить на этот вопрос, перейдем к представлению вторичного квантования (см. главу 8). Из (8.2.6) получаем:

$$Q = \int W q_{\mathbf{k}} e^{i\mathbf{k}\mathbf{x}} \frac{d\mathbf{k}}{(2\pi)^{3/2}}, \quad P = \int W \dot{q}_{\mathbf{k}} e^{i\mathbf{k}\mathbf{x}} \frac{d\mathbf{k}}{(2\pi)^{3/2}}, \quad (9.6.9)$$

где $q_{\mathbf{k}} = a_{\mathbf{k}} \nu_{\mathbf{k}} + a_{\mathbf{k}}^\dagger \nu_{\mathbf{k}}^*$ — операторы элементарных осцилляторов,

$$W = W(k\sigma) \equiv \exp\left(-\frac{k^2\sigma^2}{2}\right) = \int K_\sigma(\mathbf{r}) e^{i\mathbf{k}\mathbf{r}} d\mathbf{r}$$

— гауссов фильтр (функция окна) (рис. 9.8). В фурье-представлении гауссиана остается гауссианой и сглаживание означает, что мы оставили только гармоники с $k \leq \sigma^{-1}$ или, эквивалентно:

$$|k\eta| \leq c^{-1} \ll 1. \quad (9.6.10)$$

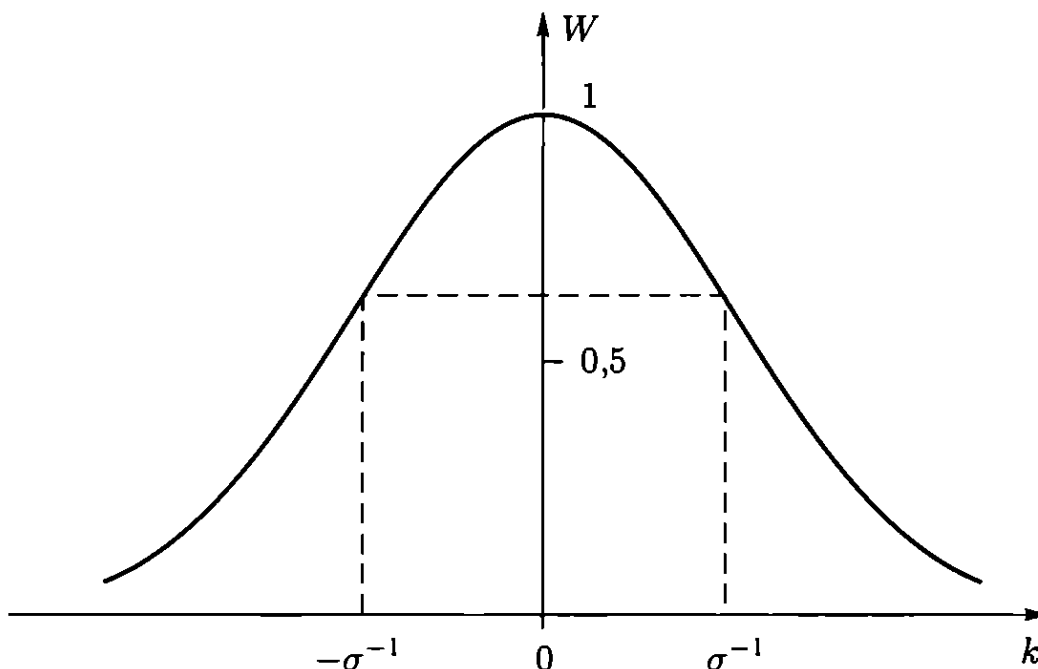


Рис. 9.8. Гауссов фильтр крупномасштабных полей Q и P

На адиабатической стадии инфляции уравнение поля q имеет простой вид:

$$\ddot{q} + 3H\dot{q} - a^{-2}\Delta q = 0. \quad (9.6.11)$$

Перепишем его в виде системы уравнений для сглаженных полевых компонент:

$$\dot{Q} - P = f, \quad (9.6.12)$$

$$\frac{\dot{P}}{H} + 3P = g - c^{-2}f, \quad (9.6.13)$$

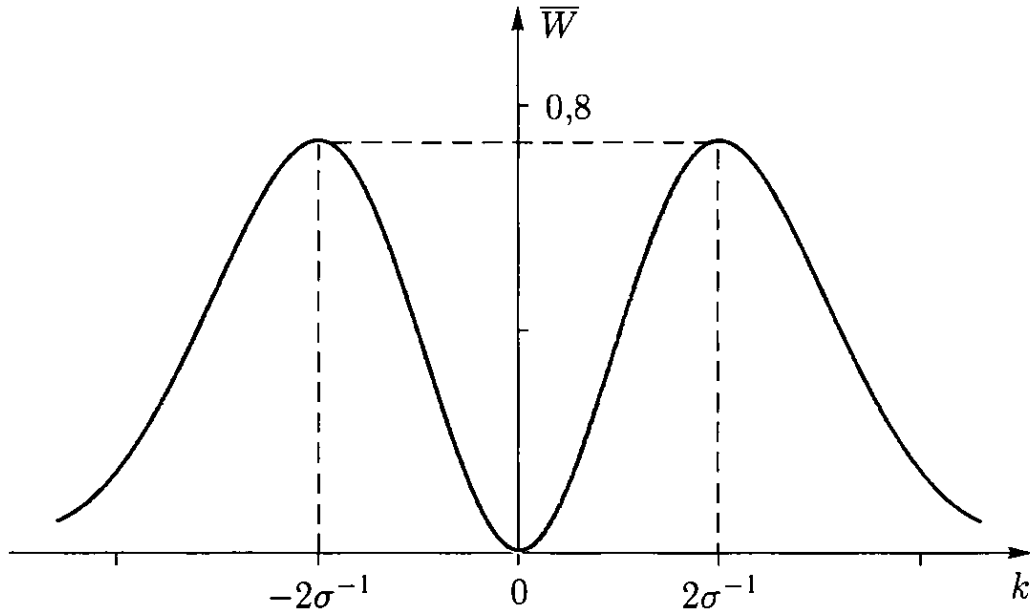
где в правых частях

$$f = H \int \overline{W} q_{\mathbf{k}} e^{i\mathbf{k}\mathbf{x}} \frac{d\mathbf{k}}{(2\pi)^{3/2}}, \quad g = \int \overline{W} \dot{q}_{\mathbf{k}} e^{i\mathbf{k}\mathbf{x}} \frac{d\mathbf{k}}{(2\pi)^{3/2}}$$

— гауссовы случайные генераторы, управляемые квантовыми флуктуациями кривизны в области $r \leq \sigma$.

Поля f и g имеют эффективный фильтр, представленный на рис. 9.9:

$$\overline{W} = \overline{W}(k\sigma) \equiv k^2\sigma^2 \exp\left(-\frac{k^2\sigma^2}{2}\right),$$


 Рис. 9.9. Эффективный фильтр мелкомасштабных полей f и g

который «убирает» гармоники с малыми k . Таким образом, мы можем заключить, что на масштабе σ происходит переход от начальных микроскопических квантовых возмущений в квази-классический режим.

Корреляторы случайных полей находятся путем усреднения квадратов соответствующих операторов по состоянию (8.3.11) и экстраполяции временных функций ν_k и ν'_k к границе классической области (9.6.10). Например, для f имеем: $\langle f \rangle = 0$ и

$$\langle q_{\mathbf{k}_1} q_{\mathbf{k}_2}^\dagger \rangle = |\nu_k|^2 \delta(\mathbf{k}_1 - \mathbf{k}_2), \quad (9.6.14)$$

$$\langle f_1 f_2 \rangle = \frac{H^2 \sigma_1^2 \sigma_2^2}{2\pi^2} \int_0^\infty W \frac{\sin kr}{kr} |\nu_k|^2 k^6 dk, \quad (9.6.15)$$

где $\sigma^2 = \sigma_1^2 + \sigma_2^2$. Мы видим, что причинно-несвязанные области пространства ($r > \sigma$) не коррелированы между собой, а в каждой из локальных областей размером $r \leq \sigma$ мы имеем дело с временными корреляциями:

$$\langle f(t_1) f(t_2) \rangle = \frac{H^2 \sigma_1^2 \sigma_2^2}{2\pi^2} \int_0^\infty W |\nu_k|^2 k^6 dk. \quad (9.6.16)$$

С помощью линейной комбинации

$$\tilde{Q} = Q + \frac{1}{3H} P \quad (9.6.17)$$

уравнение (9.6.12) можно привести к нормальному виду:

$$\ddot{\tilde{Q}} = \tilde{f}, \quad (9.6.18)$$

где

$$\tilde{f} = \left(1 - \frac{1}{3c^2}\right) f + \frac{1}{3} g,$$

\tilde{Q} и P являются нормальными модами уравнений (9.6.12), (9.6.13). В этом представлении мы можем отдельно изучать эволюцию классического поля и его производной.

Для определения коэффициентов корреляций обратимся к точному решению уравнения (9.6.11) в локальной переходной области $\eta \sim -k^{-1}$ и перейдем к пределу (9.6.10):

$$\begin{aligned} \nu_k &= \frac{H}{\alpha\sqrt{2k}} \left(i\eta + \frac{1}{k}\right) e^{-ik\eta} \rightarrow \frac{H}{M_P} \sqrt{\frac{2\pi}{\gamma k^3}} (1 + o(c^{-2})), \\ \nu'_k &= \frac{H}{\alpha\sqrt{2k}} k\eta e^{-ik\eta}, \quad \frac{|\nu'_k|}{\nu_k H} \rightarrow o(c^{-2}). \end{aligned} \quad (9.6.19)$$

Пренебрегая членами $\sim c^{-2}$, получаем, что из всех квадратичных корреляторов не исчезают в нулевом порядке по c^{-2} только следующие:

$$\langle f_1 f_2 \rangle = \langle \tilde{f}_1 \tilde{f}_2 \rangle = D \delta^{(1)}(t_1 - t_2), \quad (9.6.20)$$

где коэффициент диффузии определяется выражением:

$$D = \frac{H^3}{2\gamma M_P^2}. \quad (9.6.21)$$

При выводе (9.6.20) мы заменили функцию $\frac{4\sigma_1^2\sigma_2^2}{\pi\sigma^4}$ на

$$\delta^{(1)}\left(1 - \frac{\sigma_2}{\sigma_1}\right) = H^{-1}\delta^{(1)}(t_1 - t_2),$$

поскольку размер ее локализации не превышает σ .

При условии (9.6.20) уравнение (9.6.18) напоминает уравнение Ланжевена, описывающее дрейф броуновской частицы под действием некоррелированных соударений малых частиц. Также в каждой точке x под действием квантовых флуктуаций дрейфует классическое поле Q . Интегрируя (9.6.18) и усредняя квадрат по ансамблю (9.6.20), получаем расплывание вариации поля со временем:

$$\langle \tilde{Q} \rangle = 0, \quad \langle \tilde{Q}^2 \rangle = \int_0^t D dt. \quad (9.6.22)$$

Физически это означает, что в процессе космологической инфляции постоянно рождаются мелкомасштабные поля. Рождение конкретной моды k происходит в момент, когда она пересекает горизонт. В этот момент поле Q «получает удар» и его импульс приобретает малое изменение. Потом из-под горизонта появляется меньшая длина волны, и опять происходит переопределение среднего поля. Дрейф среднего импульса P , как видно из (9.6.13), подавлен фактором c^{-2} : за горизонтом поля квазистатичны. В точности до этого же фактора c^{-2} знак тильды у функции \tilde{Q} можно опустить.

Аналогично стандартной технике для решений уравнения Ланжевена можно ввести понятие вероятности $dP = \Pi dQ$ получить поле с величиной Q в момент времени t . Плотность вероятности распределения поля $\Pi = \Pi(Q, t)$ описывается диффузионным уравнением Фоккера–Планка:

$$\frac{\partial \Pi}{\partial t} = \frac{D}{2} \frac{\partial^2 \Pi}{\partial Q^2}, \quad (9.6.23)$$

решение которого, нормированное на $\Pi(Q, t_0) = \delta^{(1)}(Q)$ в начальный момент времени t_0 , имеет вид:

$$\Pi = \frac{1}{\sqrt{2\pi D}} \exp\left(-\frac{Q^2}{2D}\right). \quad (9.6.24)$$

Подставляя (9.6.24) в (9.6.23), получаем растущую со временем вариацию $D \equiv \langle Q^2 \rangle = \int D dt$ (ср. (9.6.22)).

Мы видим, что в ходе инфляции идет дрейф поля кривизны квазифридмановской гиперповерхности. Поле $Q(t, \mathbf{x})$ эволюционирует независимо в каждой точке \mathbf{x} как в положительную сторону, так и в отрицательную, и его вариация пропорциональна времени: $D \propto t$. С окончанием процесса инфляции квантовые генераторы f и g исчезают и Q перестает зависеть от времени, определяя «застывшую» классическую реализацию начального потенциала возмущений кривизны: $Q = q(\mathbf{x})$.

9.7. Множественные миры

По своей природе поле q является гравитационным скалярным потенциалом Вселенной. Сама по себе величина Q в данном месте пространства значения не имеет. Важны ее градиенты — разность потенциалов в разнесенных местах пространства. Дрейф поля Q идет независимо в каждой точке, однако процесс

расплывания поля степенной, тогда как расстояние между точками увеличивается экспоненциально.

Таким образом, если амплитуда поля на каждом шаге рождения меньше единицы ($D < H$), то пространство *изотропизуется*. Уменьшение со временем локальных градиентов Q означает падение пекулярной скорости и затухание (рассасывание) неоднородностей плотности и давления. Пространственная гиперповерхность как бы разворачивается, распрямляется, выглаживается, оставаясь при этом конечной и глобально искривленной (закрепленной) в своих границах, но стремящейся к однородности и изотропии внутри границ. Параллельно с течением времени происходит монотонное «скатывание» инфлатона к минимуму своего потенциала, и амплитуда новых рождающихся k -мод поля q тоже монотонно уменьшается. Процесс выглаживания прогрессирует: идет строительство упорядоченного хаббловского потока.

Если бы этот процесс построения продолжался вплоть до $q \rightarrow 0$, мы бы получили асимптотически фридмановскую симметрию. Однако формирование хаббловского потока прерывается в момент окончания инфляции и остаются застывшие малые возмущения кривизны. В результате мы имеем квазифридмановскую геометрию.

Так протекает инфляция при малых $|q| < 1$. При больших полях Q (соответственно при $H > H_c$ или $D > H$) теория возмущений формально перестает быть применимой, однако в этой области параметр γ оказывается малым ($\gamma < H^2/M_P^2$, см. (9.6.21)), и мы имеем дело с флуктуациями в модели де Ситтера (см. раздел 4.2). В этой ситуации возмущения де-ситтеровского метрического тензора продолжают оставаться малыми, а большие поля q (также как и поля v_i , см. сноску на с. 137) в каждом акте рождения ($\Delta\mathcal{N} \sim 1$) означают нелинейный переход (флуктуацию) к очередному *неоднородному* расщеплению геометрии де Ситтера. Мы можем только предполагать, что при столь интенсивном и быстропеременном гравитационном поле пространственных неоднородных расщеплений на масштабах $\sim H^{-1}$ происходит перераспределение энергии между полями, подобными q , рождаются новые физические поля, возможно, меняются константы взаимодействий, нарушаются и создаются новые симметрии и т. д. Главной причиной этого стохастического процесса является гравитация.

На этом пути мы приходим к идее множественности миров или к «мультиверсу» (от англ. multiverse). Во флуктуирующей метрике случайным образом могут возникать геометрические конфигурации конечных размеров с плотностью ниже критической ($E < E_c$), где последующие флуктуации уже не столь боль-

шие и не способны кардинально изменить и перекрыть данную область пространства-времени (см., напр., (4.2.10)). При достаточно гладком начальном распределении кривизны и материального поля:

$$|q| < 1, \quad \left| \frac{\nabla\varphi}{\varphi} \right| < m, \quad \left| \frac{\dot{\varphi}}{\varphi} \right| < m, \quad (9.7.1)$$

эти области подхватываются инфляцией и «раздуваются» в квазифридмановские миры с низкой плотностью энергии¹⁾.

Образующиеся таким образом миры находятся в причинно-несвязанных областях пространства-времени, и их разделяют высокоэнергетические пограничные области, где доминируют флуктуации метрики. Каждый из миров эволюционирует независимо от другого, у них разная физика, разные поля q , по-разному протекают процессы динамики и образования структуры. По-видимому, Природа таким образом «кидает кости», испытывая разные варианты физических конфигураций и существуя в бесконечном многообразии состояний. Саму эту возможность ей предоставляет гравитация.

Заметим, что сегодняшние рассуждения о множественных мирах не подкреплены строгими расчетами. Мы можем лишь догадываться о нелинейных квантовых процессах с участием гравитации, эта наука только начала развиваться.

Таким образом, кажется логически обоснованным, что мы находимся в одном из таких квазифридмановских миров, где условия случайным образом оказались подходящими для образования *структуры* и для существования *долгоживущих звезд*. В этом кардинальную роль сыграли параметры темной энергии и темной материи (подробнее см. разделы 10.6, 10.7). С точки зрения парадигмы мультиверса наша часть Вселенной не является типичным или наиболее вероятным результатом эволюции (принцип Коперника).

¹⁾ Обращая неравенства (9.7.1), получаем минимальный начальный размер сглаженной области пространства-времени, в которой выполняются стартовые условия запуска инфляции: $\ell > \ell_C \equiv m^{-1}$. Это комптоновский (квантовый) размер массивного поля φ : область локализации его неприводимой энергии mc^2 . Поясним также топологию раздувания и отсутствие влияния границ. Часто возникающий вопрос — куда же все это расширяется и почему не «расталкивается» соседняя материя? — связан с артефактом евклидовой геометрии. Реальная же природа многомерна, и ее топология нетривиальна. В этой ситуации существует масса возможностей, при которых инфляция отдельных измерений не вызывает никакого противодействия кривизны и не приводит к столкновениям. Простейший пример: расширение цилиндрической поверхности с *нулевой* внутренней кривизной — плоский лист, свернутый в трубку, — во внешнее пустое пространство.

ЭКСПЕРИМЕНТАЛЬНЫЕ ОСНОВАНИЯ СТАНДАРТНОЙ МОДЕЛИ ВСЕЛЕННОЙ

Наши представления о физических процессах в ранней Вселенной опираются на три основания:

- *наблюдения и эксперименты* в области сравнительно низких энергий, доступных ускорителям;
- внутренне непротиворечивую физическую космологическую *теорию*, построенную на экспериментально-наблюдательном базисе;
- *экстраполяцию* известных физических законов в область высоких энергий ранней Вселенной.

Наиболее важным является первое основание — наблюдательная база космологии, которая включает накопленные человечеством экспериментальные и наблюдательные результаты и факты, число и точность которых возрастает благодаря использованию все более совершенных технологий, появлению новых инструментов и методов исследований, прогрессу вычислительной техники и обратной связи с развивающейся теорией.

Теория очень ранней Вселенной смыкается с построением единой модели мироздания, поскольку именно процессы, происходившие при высоких энергиях, привели к тому миру, который мы наблюдаем и частью которого являемся сами. Сегодня у нас есть две отдельные теории: микромира (стандартная модель элементарных частиц) и макромира (космологическая стандартная модель, КСМ). Обе они развиваются, достраиваются и сближаются, подходя с двух сторон к построению будущей единой теории мира.

В этой главе мы обсудим экспериментальные и наблюдательные основы КСМ.

10.1. Вперед в прошлое

Основная цель теории ранней Вселенной — объяснить структуру и развитие окружающего нас мира, геометрию пространства-времени и свойства материи.

Сегодня ни для кого не секрет, что Вселенная, в которой мы живем, нестационарна, то есть возникла эволюционно и продол-

жает динамически развиваться. То, что мы видим вокруг, — это лишь текущий этап ее существования, одна из ее «фотографий». Поэтому, чтобы восстановить всю биографию Вселенной, необходимо проводить экстраполяцию за границы прямого опыта. Рельсами экстраполяции является теория. Основываясь на ОТО и физике элементарных частиц, мы экстраполируем наши знания в прошлое и к большим энергиям.

Вехами на пути физической экстраполяции, по которым сверяется ее достоверность и точность, являются знания, полученные астрономией, атомной физикой, ядерной физикой, физикой элементарных частиц и непрямыми (астрофизическими) экспериментами.

Энергии, достижимые в прямых экспериментах и ускорителях на Земле, не превышают 100 ТэВ, что покрывает масштабы сильных (~ 100 МэВ) и электрослабых (~ 1 ТэВ) взаимодействий.

Информация о физике частиц и взаимодействиях при больших энергиях приходит из астрофизических наблюдений: ее источниками служат сверхновые звезды, гамма-излучение далеких источников, проявления нейтронных звезд и черных дыр, космические лучи высоких энергий, нейтринные и гравитационно-волновые детекторы. Помимо этого, важными направлениями исследований являются поиски продуктов распада темной материи, исследование динамических проявлений темной энергии, геометрии Вселенной, поляризации реликтового излучения и многие другие. Существенным дополнением к этим направлениям являются лабораторные эксперименты по распаду протона, упругому рассеянию частиц темной материи, осцилляциям нейтрино, поиску магнитных монополей и т. п.

Рисунок 10.1 схематично представляет области экспериментальных исследований с характерными для них энергиями E , соответствующими радиусу кривизны ¹⁾ Вселенной H^{-1} :

$$E \equiv \epsilon^{1/4} \simeq 0,6 \sqrt{M_P H}. \quad (10.1.1)$$

Рассмотрим основные результаты и уроки экстраполяции, двигаясь от малых энергий к большим.

¹⁾ Вместо временной шкалы t мы используем геометрический масштаб H^{-1} — радиус внешней кривизны, или хаббловский радиус Вселенной в момент времени t . Для степенных функций $a \propto t^{1/\gamma}$: $t = (\gamma H)^{-1}$. Современный возраст Вселенной $H_0^{-1} \simeq t_0 \simeq 13,7$ млрд лет.



Рис. 10.1. Экспериментальные основания теории ранней Вселенной. Слева — радиус кривизны H^{-1} , выраженный в секундах, справа — энергия E

10.2. Астрономия и атомная физика

Важнейшая информация о геометрических свойствах и составе Вселенной при энергиях $E \in (10^{-3}, 1)$ эВ приходит к нам с квантами света. Это всеволновая астрономия, исследующая на разных длинах волн структуры разных масштабов — от единичных объектов (таких как карликовые галактики размером ~ 1 кпк) до сверхскоплений галактик и всей видимой Вселенной, о которых мы можем судить по обширным галактическим каталогам, охватывающим седьмую часть современного хаббловского радиуса (~ 600 Мпк), и картам реликтового излучения.

Геометрическая структура Вселенной проста, и ее развитие детерминировано. На больших масштабах мы наблюдаем квазихаббловский разлет материи, на малых — нелинейные гравитационно-связанные структуры (гало материи). Последние образуются эволюционно в результате роста контраста плотности, индуцированного гравитационной неустойчивостью темной материи и наличием малых геометрических отклонений от пространственной однородности и изотропии (начальные возмущения). Сначала возникают гало малых размеров, а затем все более крупные.

В настоящее время средний по Вселенной масштаб неоднородности, разделяющий нелинейные и квазилинейные области распределения материи, близок к 10 Мпк (рис. 10.2), что соответствует охватываемой массе $\sim 2 \cdot 10^{14} M_{\odot}$. Но это, как уже было сказано, средний масштаб. Локальный масштаб неоднородности варьируется в зависимости от величины плотности

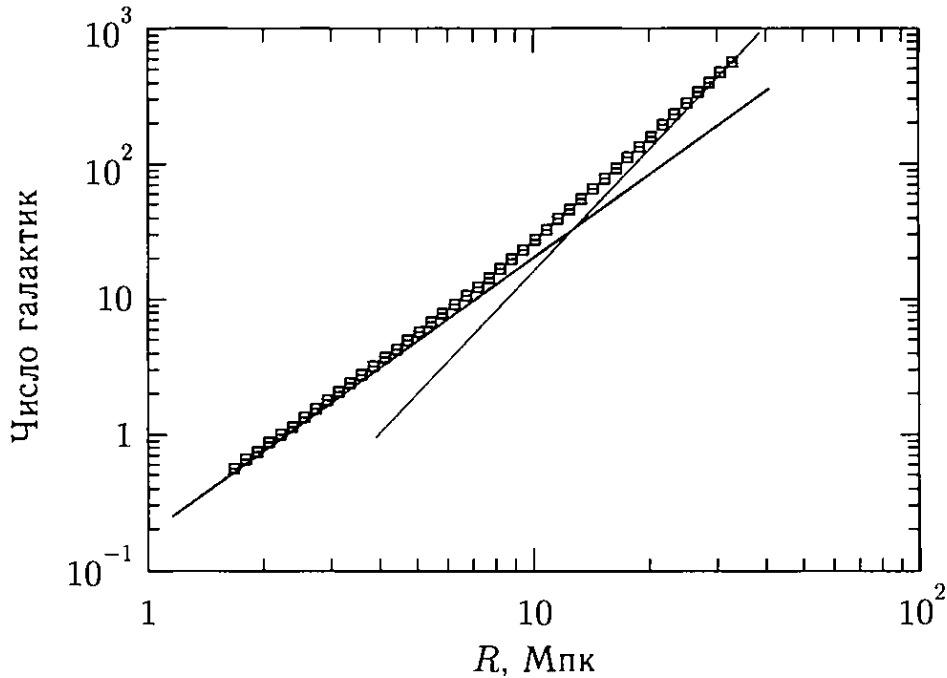


Рис. 10.2. Число галактик $N(R)$ в сфере радиуса R при усреднении по местоположению в пространстве (по данным [Sylos Labini et al., 2009]). Точность степенных асимптотик $\sim 10\%$: красная линия $N \propto R^3$ — однородное распределение материи, синяя линия $N \propto R^2$ — нелинейная структура Вселенной

вещества в рассматриваемом объеме. Так, самыми большими «возмутителями спокойствия» являются богатые скопления галактик, масса которых достигает $10^{15} M_{\odot}$. Однако для Местной группы галактик, находящейся вдали от больших скоплений вещества, масштаб неоднородности — всего около двух Мпк, что соответствует массе Местной группы $\sim 2 \cdot 10^{12} M_{\odot}$.

Материя, заполняющая Вселенную, состоит из многих компонент. Первые пять из них в порядке значимости вклада в общую плотность Вселенной — это скрытая, или темная, энергия ($\Omega_E \simeq 0,7$), темная материя ($\Omega_M \simeq 0,25$), обычное вещество ($\Omega_b \simeq 0,05$), массивные нейтрино ($\Omega_\nu \sim 10^{-2}$) и излучение ($\Omega_\gamma \simeq 5 \cdot 10^{-5}$). Сумма этих пяти компонент равна единице, причем точность определения общей плотности ($\sim 1\%$) выше, чем темной энергии и темной материи по отдельности ($\sim 10\%$). Причина этого кроется в том, что именно сумма всех компонент определяет геометрию и пространственную кривизну Вселенной. Также нам хорошо известны плотность излучения и плотность барионов (точнее, комбинация $\Omega_b H_0^2$, см. раздел 10.6).

Эти количественные характеристики состава и распределения материи во Вселенной лежат в фундаменте КСМ. Рассмотрим последовательно их наблюдательные и экспериментальные обоснования.

10.3. Закон Хаббла

Начнем с больших масштабов.

Мы наблюдаем состояние Вселенной спустя миллиарды лет после Большого взрыва. Однако современное понимание Большого взрыва и соответствующей организации движения (разлета) вещества сильно отличается от обыденных представлений о взрыве.

У космологического взрыва нет центра, а имеется конечная в пространстве и во времени область расширяющейся материальной среды, часть которой и есть *видимая* Вселенная. Можно сказать, что в пределах этой области «центр» везде, поскольку относительная скорость удаления *любой* пары соседних точек среды δV пропорциональна собственному расстоянию между ними δr :

$$\delta V \equiv \frac{\partial \delta r}{\partial t} = H \delta r. \quad (10.3.1)$$

Коэффициент пропорциональности $H = H(t)$ не зависит от взаимного расположения точек среды (лагранжевой координаты x), но зависит лишь от собственного времени t . Такой закон расширения материи мы называем законом Хаббла: он не делает различия между пространственными точками среды и, соответственно, *бесструктурен*.

Если бы закон (10.3.1) выполнялся для любых собственных расстояний δr от наблюдателя до источника, то он свидетельствовал бы о глобальной однородности и изотропии пространственных сечений $t = \text{const}$. Однако таких свидетельств в реальности не существует. Во-первых, любые наблюдения ограничены горизонтом (конечным размером области наблюдений). А во-вторых, мы имеем конечную точность наблюдательных данных: при увеличении точности, как правило, обнаруживаются отклонения от закона (10.3.1). В такой ситуации, даже если H не зависит от x , описываемая (10.3.1) метрика оказывается однородной и изотропной лишь *локально*¹⁾, то есть *зависит* от координат x .

Остановимся на этом подробнее.

Если закон (10.3.1) выполняется для наблюдателя, находящегося в *одной* точке среды, то этот закон будет справедлив и для остальных точек с тем же коэффициентом H . Иллюстрацией

¹⁾ Максимальный размер области однородности может быть достаточно большим, его определение — это вопрос точности наблюдательных данных. Для нашей Вселенной он превышает хаббловский радиус (см. (10.3.8), (10.3.9)).

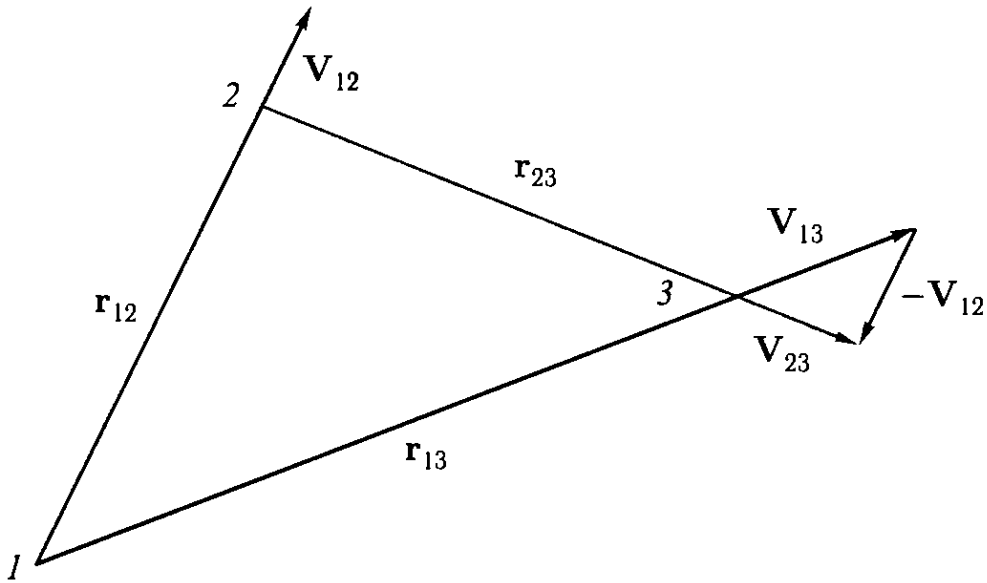


Рис. 10.3. Иллюстрация к закону Хаббла

этого утверждения служит рис. 10.3. На нем цифрами 1, 2 и 3 обозначены галактики хаббловского поля. Пусть мы находимся в точке 1 и наблюдаем удаление галактик 2 и 3. Тогда их скорости относительно галактики 1 будут определяться выражениями:

$$\mathbf{V}_{12} = H\mathbf{r}_{12}, \quad \mathbf{V}_{13} = H\mathbf{r}_{13}. \quad (10.3.2)$$

Перейдя в систему отсчета галактики 2, мы увидим ту же картину:

$$\mathbf{V}_{23} = \mathbf{V}_{13} - \mathbf{V}_{12} = H(\mathbf{r}_{13} - \mathbf{r}_{12}) = H\mathbf{r}_{23}. \quad (10.3.3)$$

Итак, если коэффициент H в (10.3.1) не зависит от направления, то он не зависит и от \mathbf{x} . Это, однако, *не доказывает* однородность пространства: мы имеем дело с тавтологией, поскольку *изотропная* гиперповерхность закладывается руками в форме линейного сложения расстояний и скоростей, которое мы использовали в (10.3.3).

Геометрически хаббловский поток в форме (10.3.1) можно представить себе как равномерно растягивающуюся со временем трехмерную гиперповерхность $t = \text{const}$, причем растяжение локально-изотропно во *всех* точках среды с фактором локального растяжения $a_{ij} = a(t) \cdot a_{ij}(\mathbf{x})$, где $a = a(t)$ и $a_{ij}(\mathbf{x})$ — гладкие непрерывные функции класса C^2 .

Действительно, полагая по определению:

$$H_i^j \equiv \dot{a}_{ik} a^{kj},$$

где a_{ij} и a^{ij} — прямая и обратная симметричные положительно определенные матрицы, получаем из (10.3.1):

$$H_i^j = H \delta_i^j: \quad a_{ij} = a(t) \cdot a_{ij}(\mathbf{x}),$$

$$a = \exp\left(\int H dt\right), \quad (10.3.4)$$

что соответствует квадрату интервала между точками среды в следующем виде:

$$ds^2 = dt^2 - a^2 g_{ij} dx^i dx^j,$$

$$g_{ij} = g_{ij}(\mathbf{x}) = a_{ik} a_{lj} \delta^{kl} \quad (10.3.5)$$

(здесь символ Кронекера появился из уравнения (10.3.4)).

Верно и обратное. Предполагая (10.3.5), получаем, что собственное расстояние между близкими точками среды, разделенными координатами δx^i на гиперповерхности $t = \text{const}$, равно $\delta r_i = a(t) a_{ij}(\mathbf{x}) \delta x^j$. Дифференцируя по времени, получаем закон расширения в форме (10.3.1), где $H = \dot{a}/a$. Таким образом, формулы (10.3.1) и (10.3.5) эквивалентны. Они описывают локально изотропное и однородное пространство с *однородным* распределением материи и хаббловским течением, в конечной окрестности любой точки которого выполняется *линейный* закон сложения скоростей. При переходе из одной области пространства в другую надо переопределять правило сложения расстояний с учетом функций g_{ij} .

Действительно, пусть g_{ij} — это аналитические функции класса C^ω , которые можно разложить в ряд Тейлора в конечной окрестности произвольной точки:

$$g_{ij}(\mathbf{x}) = c_{ij} + k_0 c_{ijk} x^k + \frac{1}{2} k_0^2 c_{ijkl} x^k x^l + \dots, \quad (10.3.6)$$

где коэффициенты $k_0, c_{ij} \dots$ зависят от выбора точки (наблюдателя), которую в данном случае мы совместили с $\mathbf{x} = 0$, а k_0^{-1} — масштаб вариации следа g_{ii} , или радиус сходимости $k_0 |\mathbf{x}| < 1$, в пределах которого первый член ряда c_{ij} больше всей остальной суммы. Поскольку речь идет о форме $g_{ij} dx^i dx^j$, то с помощью надлежащего выбора координат всегда можно привести c_{ij} к виду $c_{ij} = \delta_{ij}$ и оставить лишь не исчезающие коэффициенты $c_{ijk} \dots$. Тогда выражение (10.3.6) сводится к

$$g_{ij}(\mathbf{x}) = e^{-2q_0} (\delta_{ij} - 2\mathfrak{G}_{ij}), \quad (10.3.7)$$

где неприводимые скаляр $q_0 = q(\mathbf{x})$ и тензор $\mathfrak{S}_0 = \mathfrak{S}_{ij}(\mathbf{x})$ зависят от \mathbf{x} и малы в окрестности наблюдателя, индекс «0» относится к поздней Вселенной ¹⁾.

Назовем окрестность $|\mathbf{x}| < k_0^{-1}$ с хаббловским течением материи фридмановской областью или *миром*: внутри этой области выполняется линейный закон сложения скоростей. Согласно (10.3.7), величина $q(\mathbf{x})$ определена с точностью до аддитивной постоянной, однако разность ее значений между удаленными точками может быть сколь угодно большой. Обычно полагают, что средняя величина q_0 в наблюдаемом объеме равна нулю. Тогда отклонения $q(\mathbf{x})$ от нуля возрастают с увеличением расстояния от наблюдателя, и границы фридмановского мира определяются условием:

$$|q(k_0^{-1})| \sim 1. \quad (10.3.8)$$

Этот размер заведомо больше радиуса внешней кривизны ($k_0 \ll H_0$), поскольку, согласно астрономическим данным по крупномасштабной анизотропии реликтового излучения, на хаббловском расстоянии

$$q_H = |q(H_0^{-1})| \sim 10^{-4}. \quad (10.3.9)$$

Таким образом, из наблюдений следует, что видимая Вселенная находится внутри фридмановского мира и $|q(\mathbf{x})| \ll 1$ в масштабах наблюдательной космологии ²⁾.

¹⁾ При последовательном вычислении в правую часть (10.3.7) следует также включить скалярное $B_{,ij}$ и векторное $\xi_{(i,j)}$ слагаемые, зависящие от \mathbf{x} . Однако оба эти члена нефизичны и могут быть убраны с помощью координатных преобразований (см. (4.5.1)). Очевидно, векторная мода, как и падающие ветви скалярной и тензорной мод возмущений, несовместима с изотропным расширением (10.3.4). Напротив, растущие ветви S- и T-мод геометрических неоднородностей консервируются в форме (10.3.5) на больших масштабах и не нарушают хаббловский закон расширения (10.3.4) в главном порядке по q и \mathfrak{S} . Именно эти «замороженные» гравитационные потенциалы сохраняются в ранней Вселенной и являются зародышами космологической структуры, тогда как векторная и падающие моды затухают и становятся пренебрежимо малыми к началу эпохи образования галактик (см. также разделы 5.5, 5.6, 7.2).

²⁾ Для получения соотношения (10.3.9) воспользуемся формулой Сакса–Вольфа для крупномасштабной анизотропии реликтового излучения (см. (11.6.2), (11.6.3)):

$$\delta_{TSW} = q_{HZ}/5 \simeq 10^{-5},$$

и оценим q_H с помощью спектрального интеграла (см. (8.7.1)):

$$\langle q_H^2 \rangle = \int_{H_0}^{\infty} q_k^2 \frac{dk}{k} \sim q_{HZ}^2 \ln \left(\frac{k_{eq}}{H_0} \right) \simeq (2 q_{HZ})^2 \simeq 10^{-8},$$

где $q_{HZ} = q_{k \sim H_0}$ — крупномасштабный спектр Харрисона–Зельдовича, $k_{eq} = \bar{H}_{eq} \simeq 0,01 \text{ Мпк}^{-1}$ — масштаб горизонта на момент равенства плотностей излучения и вещества (ср. (6.6.9), рис. 3.12).

В функциях $q(\mathbf{x})$ и $\mathfrak{S}(\mathbf{x})$ закодирована информация о геометрии и структуре Вселенной: анизотропии реликтового излучения (q_0 и \mathfrak{S}_0) и зародышах будущих галактик (только q_0). Градиент и лапласиан поля q_0 определяют соответственно локальные отклонения движения вещества от хаббловского потока (вектор смещения) и пространственную кривизну мира (возмущение плотности). Мы видим, что наблюдаемая Вселенная *детерминирована*, а движения в ней могут рассматриваться в рамках задачи Коши. Имея динамические уравнения, мы получаем причинно-следственную цепочку событий, полностью вытекающую из *начальных космологических условий*. Это и обуславливает космологическую *стрелу времени*¹⁾.

10.4. Почему образуются галактики

Рассмотрим подробнее эволюцию вещества в конечных масштабах.

Исходя из наблюдаемого закона расширения материи (10.3.1) и понятия метрического пространства, мы получили интервал (10.3.5), который содержит информацию о зародышах крупномасштабной структуры Вселенной. Для того чтобы понять, каким образом эволюционирует данная метрика, как она образовалась и к чему ведет, необходимо обратиться к уравнениям ОТО.

Из динамического уравнения (6.3.8) следует, что метрика (10.3.5)–(10.3.7) представляет собой главный член разложения *точного* решения по малому параметру

$$\frac{\beta k}{H} < 1. \quad (10.4.1)$$

Пекулярная скорость и возмущение полной плотности связаны с полем q_0 следующим выражением (см. (5.5.3)):

$$\mathbf{v}_{\text{pec}} = -\nu \nabla q_0, \quad \delta_\epsilon = g \Delta q_0, \quad (10.4.2)$$

где факторы роста $\nu = \nu(a)$ и $g = g(a)$ зависят только от масштабного фактора a . При уменьшении эффективной скорости звука решение $q = q_0$ продолжается в причинно-связанную область (см. (10.4.1)), а при условии $|\beta| \ll 1$ функции (10.4.2) неограниченно нарастают в коротких длинах волн.

¹⁾ Однако в отдельных областях пространства-времени, где важны *нелинейные* и квантовые эффекты, стрела времени может разрушаться, например в черных дырах или кротовых норах.

Для определения минимального масштаба хаббловского потока (10.3.1) выпишем первые члены разложения метрики в лагранжевых координатах $t \equiv t_c$, $\mathbf{x} \equiv \mathbf{x}_c$ (см. (5.1.4)):

$$ds^2 = (1 - 2\beta^2\delta_\varepsilon) dt^2 - \delta\mathbf{r}^2, \quad (10.4.3)$$

$$\delta\mathbf{r} \equiv \delta r_i = a e^{-q_0} (1 - \bar{b} \Delta q_0) (\delta_{ij} - b q_{0,ij}) dx^j. \quad (10.4.4)$$

Здесь функция \mathfrak{S}_0 опущена по причине тривиальности и для сокращения записи. Функция q_0 есть замороженная часть поля q хаббловского потока:

$$q = q_0 + \bar{b} \Delta q_0 + \dots \quad (10.4.5)$$

Теперь мы можем записать закон Хаббла в более корректной форме:

$$\delta V_i \equiv \frac{\partial \delta r_i}{(1 - \beta^2 \delta_\varepsilon) \partial t} = H_{ij} \delta r^j, \quad (10.4.6)$$

где

$$H_{ij} = H (\delta_{ij} - h q_{0,ij}). \quad (10.4.7)$$

Функции роста h , b и \bar{b} зависят от масштабного фактора a и связаны следующими соотношениями с ν и g (см. (5.5.4)):

$$h = \frac{\nu}{H} = \frac{db}{d \ln a} \equiv \frac{c}{H^2}, \quad (10.4.8)$$

$$g = \frac{d\bar{b}}{\beta^2 d \ln a} \equiv \frac{\bar{c}}{H^2}, \quad (10.4.9)$$

$$c = \frac{H}{a} \int_0^a \frac{da}{H}, \quad (10.4.10)$$

$$\bar{c} = \frac{H}{\gamma a} \int_0^a \frac{\gamma da}{H} = \frac{H}{\gamma a} \int_0^a a d\left(\frac{1}{H}\right) = \frac{1-c}{\gamma}. \quad (10.4.11)$$

Ключевым фактором здесь является функция $c = c(a)$. Перечислим ее наиболее важные значения:

- для $\gamma \geq 0$ (отсутствие «фантомной» материи) $c \leq 1$;
- для $\gamma < 3$ (нет предельной жесткости) $c > 1/4$;
- для $\gamma = 1$ ($\bar{H} = \text{const}$) $c = 1/2$;
- для $\gamma < 1$ (инфляция) $c > 1/2$;
- для $\gamma > 1$ (замедленное расширение) $c < 1/2$;
- для $\gamma = \text{const}$ (степенная эволюция) $c = \bar{c} = (\gamma + 1)^{-1}$.

Таким образом, при инфляции (функция $\bar{H}(a)$ монотонно растет) факторы h и ν затухают и фридмановская геометрия восстанавливается. Наоборот, при замедляющемся расширении (функция $\bar{H}(a)$ монотонно падает) факторы h , ν и g нарастают и могут разрушить фридмановскую геометрию. Эти выводы опираются на решение уравнений гравитации (10.4.3)–(10.4.11) и имеют самую общую применимость в рамках ОТО.

Основные выводы таковы.

- Инфляция выстраивает хаббловский поток и делает невозможным развитие неоднородных структур.
- При замедленном космологическом расширении развивается гравитационная неустойчивость и возникают необходимые условия для образования галактик.
- Пренебрегая эффективной скоростью звука, на нижней границе хаббловского потока k_1 имеем:

$$\delta_\epsilon \sim 1, \quad v_{\text{pec}} \sim \frac{\bar{H}}{k_1}, \quad q_0 \sim \frac{\bar{H}^2}{k_1^2} \ll 1. \quad (10.4.12)$$

Учитывая средний масштаб неоднородности в нашей Вселенной $k_1 \sim 0,1 \text{ Мпк}^{-1}$ (рис. 6.2, 10.2), получаем спектральную амплитуду потенциала кривизны на нижней границе современного хаббловского потока:

$$q_0(k_1) \sim 5 \cdot 10^{-6} \quad (10.4.13)$$

в соответствии с требуемым значением переходной функции $q_0/q_k = T(k_1) \sim 0,1$ (см. (6.6.5), (6.6.8), (6.6.9)).

Каким образом текущий масштаб неоднородности соотносится с минимальным масштабом квазифридмановской геометрии в ранней Вселенной? Каковы условия для запуска инфляции или коллапса, и как они возникают?

Чтобы ответить на эти и связанные с ними вопросы, нужно знать состояние среды и исследовать решение уравнений гравитации в зависимости от начальных условий.

При локально ускоренном расширении материи возникает пространственноподобная гиперповерхность однородности в причинно-несвязанной области. Что определяет ее ориентацию в пространстве-времени (расслоение t, \mathbf{x})?

Все дело в том, что сама гиперповерхность никуда не распространяется: она изначально ограничена в сопутствующих координатах $|\delta \mathbf{x}| < k_0^{-1}$, и ее положение в пространстве-времени задается условиями на момент запуска инфляции. Для того чтобы инициировать процесс инфляционного расширения в неко-

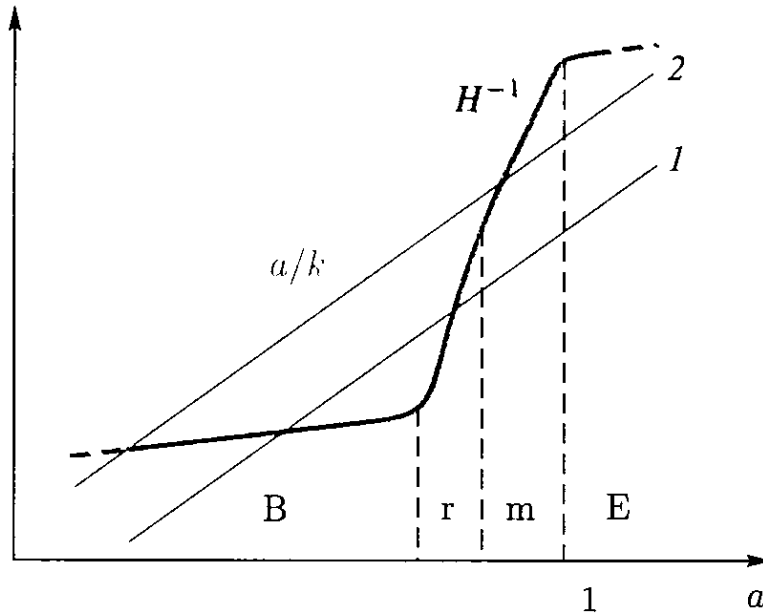


Рис. 10.4. Хаббловский радиус H^{-1} (синяя линия) и размеры возмущений (красные линии: 1 — галактика, 2 — сверхскопление) в виде функций a ; обозначения эпох см. на рис. 10.8

торый момент времени t_{i0} , необходимо иметь пространственную область *конечного* размера более хаббловского ($k_0 < \bar{H}_{i0}$) с *ограниченными* кривизной $|q| < 1$ и контрастом плотности $|\delta_\epsilon| < 1$ и обеспечить локальное *состояние* материи с $\gamma < 1$. Тогда эта область будет «подхвачена» инфляционной динамикой, а не сколлапсирует (для запуска коллапса нужны другие начальные условия), так как пекулярная скорость частиц среды локально подавлена по сравнению с хаббловской и в дальнейшем только уменьшается, в то время как хаббловская компонента скорости монотонно растет:

$$|\mathbf{v}_{\text{pec}}| \sim \frac{k|q|}{\bar{H}} < \frac{\bar{H}}{k} \sim |\mathbf{V}_H|. \quad (10.4.14)$$

В ходе инфляции происходит *растяжение* уже подготовленной пространственноподобной области размера k_0^{-1} и *достройка* функции $q(\mathbf{x})$ со стороны *малых* масштабов $k > k_0$: они приходят в зону $k < \bar{H}(t)$ из микроскопической области при ускоренном растяжении $\dot{\bar{H}} > 0$. Масштабы, растущие пропорционально $a(t)$, увеличиваются со временем быстрее хаббловского радиуса $H^{-1}(t)$ (рис. 10.4), и сам процесс растяжения идет таким образом, что все складки и неровности начальной гиперповерхности внутри k_0^{-1} с течением времени разглаживаются, а вновь *появляющиеся* малые масштабы встраиваются в уже готовый крупномасштабный каркас изотропизирующейся гиперповерхности (см. раздел 5.6).

Этот процесс управляется гравитацией и позволяет за динамически короткий период времени создать квазиоднородную, изотропную и плоскую пространственноподобную гиперповерхность конечного размера, простирающуюся в причинно-несвязанную область и являющуюся основой фридмановского мира¹⁾.

В качестве примера рассмотрим *заключительную* инфляционную стадию Большого взрыва, определившую геометрию нашей Вселенной в диапазоне масштабов

$$k_0 < k < k_m. \quad (10.4.15)$$

Максимальный размер фридмановской гиперповерхности k_0^{-1} определяется стартовыми условиями инфляции и, как следует из наблюдений, больше сегодняшнего горизонта ($k_0 < 2 \times 10^{-4} \text{ Мпк}^{-1}$, см. (10.3.8), (10.3.9)). Минимальный размер k_m^{-1} связан с окончанием инфляции, он заведомо меньше размеров наблюдаемой структуры Вселенной и может быть сопоставим с длиной волны реликтового фотона.

Очень интересен вопрос об экспериментальном определении размера k_0^{-1} фридмановской «подложки» нашей Вселенной. В принципе, такое определение возможно. Дело в том, что масштаб k_0^{-1} не обязан быть очень большим. Поскольку он связан с *последним* этапом ускоренного расширения в цепочке инфляционных стадий Большого взрыва²⁾, а этот этап по продолжительности мог быть относительно коротким, k_0^{-1} может, в принципе, не сильно превышать радиус внешней кривизны H_0^{-1} . В этом случае полный потенциал кривизны $q(\mathbf{x})$, сгенерированный к концу инфляционного взрыва и задающий геометрию наблюдаемого мира, будет складываться из *двух* составляющих: *негауссовой* сильнокоррелированной гладкой части размера k_0^{-1} , связанной с глобальной доинфляционной геометрией, и мелко-масштабной (по сравнению с k_0^{-1}) ряби гауссова поля рожденных неоднородностей.

¹⁾ Такое расширение 3-пространства из малого в большое можно представить себе как эволюцию топологически конечного трехмерного тора в n -мерном n -пространстве (где $n > 3$): внешне он никуда не расширяется (как горловина полузамкнутого мира), а внутри ничто не препятствует его расширению, поскольку внутренняя кривизна равна нулю (в отличие от полузамкнутого мира).

²⁾ Инфляционные стадии чередовались со стадиями доминирования вещества, образующегося в результате распада *промежуточных* короткоживущих инфлатонов (см. разделы 7.4, 9.3).

Негауссова компонента возмущений плотности растет с увеличением расстояния от наблюдателя и проявляет себя в квадрупольной анизотропии реликтового фона. Гауссовы возмущения не затухают с уменьшением масштаба и ответственны за формирование галактик.

Таким образом, наблюдаемая крупномасштабная анизотропия реликтового излучения индуцирована двумя физическими источниками — гауссовым и негауссовым, — которые линейно складываются и неразличимы в фазовом секторе, поскольку мы имеем дело с *единственной* реализацией гауссова поля. Однако эти составляющие можно различить по амплитуде, экстраполируя известный спектр гауссовых возмущений в длинноволновую область (то есть предполагая $k_0 = 0$ *a priori*) и определяя разницу с наблюдательными данными.

Результирующая амплитуда квадрупольной анизотропии реликтового излучения может оказаться как ниже, так и выше значения, ожидаемого при такой экстраполяции коротковолнового спектра значения. Этот эффект способен объяснить наблюдаемую аномально низкую квадрупольную анизотропию реликтового излучения (если, конечно, данная аномалия подтвердится будущими наблюдениями и улучшенным анализом данных).

Обратимся теперь к постинфляционной материально-доминированной динамике.

На стадии замедленного расширения, сопровождаемой уменьшением скорости звука, на малых масштабах возникают условия для коллапса: открывается «окно» гравитационной неустойчивости. В отличие от инфляции, где начальные условия *забываются*, для реализации коллапса необходимы *затравочные* неоднородности кривизны, которые определяют области, куда материя притекает или из которых оттекает.

К современному моменту времени область масштабов

$$k_1 < k < k_m$$

принадлежит нелинейным структурам, а в масштабах $k < k_1$ (в среднем во Вселенной) все еще продолжается неразрушенное, но уже искаженное квазихаббловское течение материи.

Мы видим, что для гравитации в равной мере характерны как отталкивание (инфляция), так и притяжение (коллапс). Все зависит от начальных условий и уравнения состояния материи — источника гравитации. Оба присущих ОТО процесса — инфляция и коллапс — возникают при достаточно общих начальных

распределениях и приводят к появлению *упорядоченных* геометрических конфигураций разных масштабов — хаббловских потоков и нелинейных гало материи.

10.5. Ранняя Вселенная

Наблюдаемая структура Вселенной, образуемая галактиками и их распределением в пространстве и времени, является результатом совместного действия двух факторов: наличия затравочного поля возмущений кривизны и роста неоднородностей плотности во времени.

В настоящее время имеется целый ряд независимых свидетельств того, что галактики возникли из малых неоднородностей плотности. В первую очередь, это измерения реликтового излучения, позволяющие реконструировать распределение барионов в эпоху рекомбинации, и данные по первичному нуклеосинтезу, обеспечившему распространённость легких химических элементов в согласии с однородной моделью Фридмана (см. раздел 10.8). Из карты реликтового излучения (рис. 10.5) видно, что в ранней Вселенной не было галактик и возмущения плотности составляли величину $\sim 10^{-5}$.

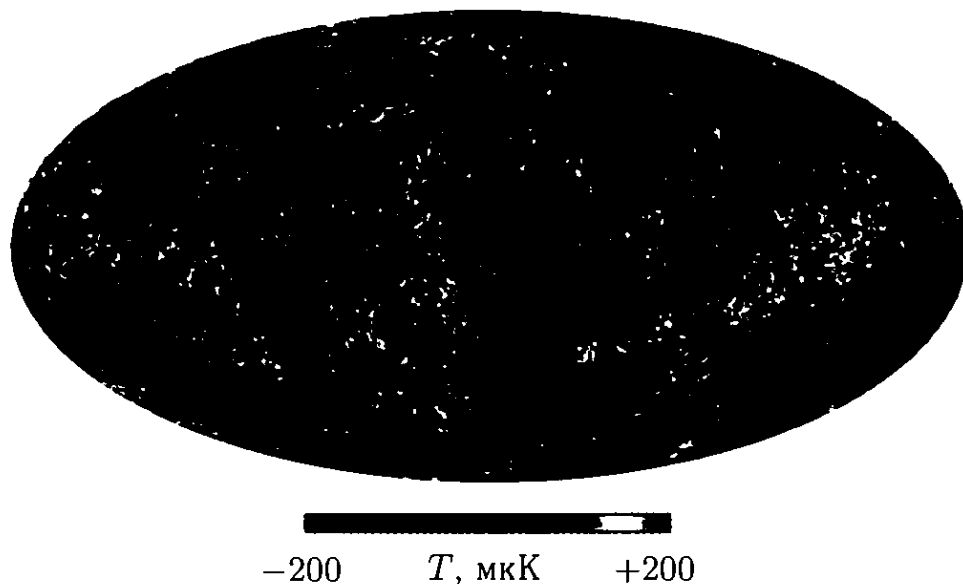


Рис. 10.5. Карта анизотропии температуры реликтового излучения на сфере последнего рассеяния $z=1100$ (по данным [Hinshaw et al., 2008])

Совместный анализ наблюдательных данных по распределению галактик и анизотропии реликтового излучения позволил отделить информацию о начальных условиях и условиях развития. Это положило начало *независимым* исследованиям физики ранней и поздней Вселенной.

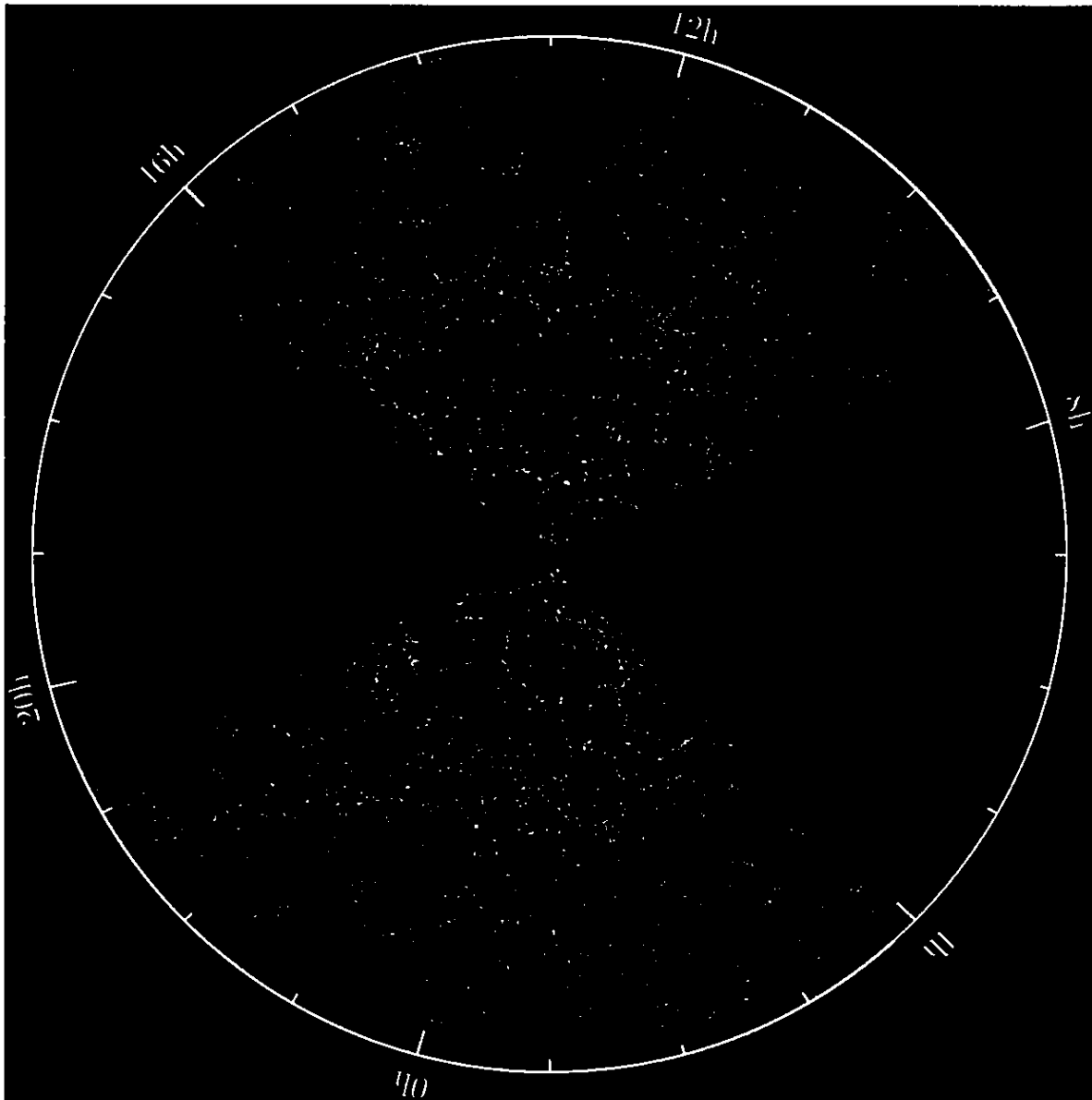


Рис. 10.6. Верный срез распределения галактик в пространстве по данным каталога SDSS (<http://www.sdss.org>). Глубина галактического каталога 600 Мпк, ширина среза $2,5^\circ$; на внешней стороне окружности обозначено прямое восхождение галактик

Действительно, мы располагаем *двумя* «фотографиями» поля космологических неоднородностей в разные эпохи его существования: догалактической (карта анизотропии реликтового излучения, рис. 10.5) и современной (трехмерное распределение галактик, например в каталоге SDSS, рис. 10.6). Опираясь на эти данные и зная законы развития гравитационной неустойчивости, мы можем восстановить спектр начальных возмущений в ранней Вселенной и вычислить значения космологических параметров, определивших динамику роста контраста плотности в поздней Вселенной. На рис. 10.7 представлен восстановленный таким образом спектр возмущений плотности, экстраполированный к $z = 0$ в соответствии с линейной эволюцией возмущений

плотности. Численные значения параметров ранней и поздней Вселенной приведены в таблице 10.1.

Таблица 10.1. Значения космологических параметров в порядке убывания точности их определения (от долей процента до 10%)

Температура реликтового излучения, T [K]	2,726
Полная плотность Вселенной, Ω	1
Плотность барионов, $\Omega_b h^2$	0,023
Плотность темной материи, $\Omega_M h^2$	0,12
Наклон S-спектра, n (при $T/S = 0$)	0,96
Постоянная Хаббла, $h = H_0/100$ [км/(с · Мпк)]	0,7
Плотность материи, $\Omega_m = \Omega_b + \Omega_M$	0,28
Плотность темной энергии, Ω_E	0,72
Оптическая толща реионизации, $\tau_{\text{Г}}$	0,08
Отношение спектров мощности, T/S	$< 0,1$
Плотность нейтрино, Ω_ν	$< 0,02$

Геометрические параметры ранней Вселенной находятся путем сравнения теоретически вычисленных спектров космологических возмущений с наблюдательными данными. Для описания геометрии можно пользоваться теорией возмущений, поскольку имеется малый параметр — начальная амплитуда космологических неоднородностей ($\sim 10^{-5}$). В нулевом порядке по этому параметру Вселенная фридмановская и описывается единственной функцией времени — масштабным фактором $a(t)$. Первый порядок устроен несколько сложнее: возмущения метрического тензора распадаются на сумму трех независимых мод — скалярной $S(k)$, тензорной $T(k)$ и векторной $V(k)$, каждая из которых характеризуется спектральной функцией волнового числа k . В последующих порядках новых свободных функций не возникает, если не будут обнаружены признаки негауссовости и/или изометрические добавки в начальном поле неоднородностей (см. раздел 6.2).

Таким образом, вся геометрия Вселенной может быть описана с помощью четырех функций — $a(t)$, $S(k)$, $T(k)$ и $V(k)$, из которых нам сегодня известны первые две (в некоторых областях определения). Теория предсказывает заметную амплитуду

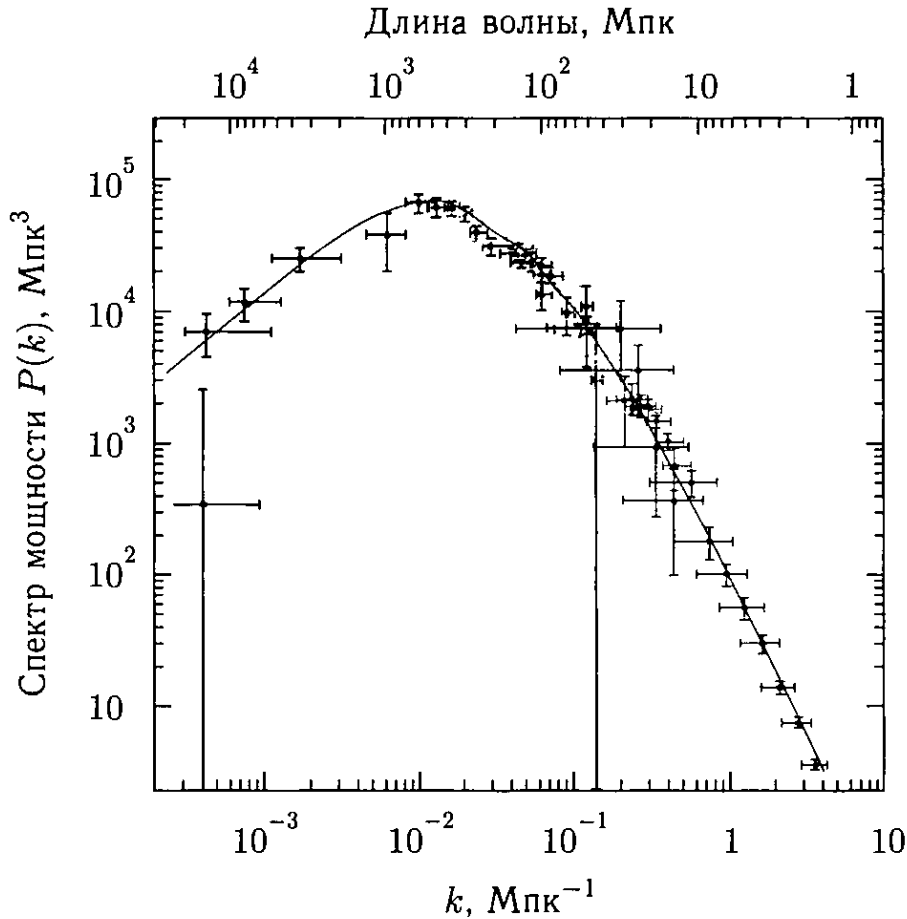


Рис. 10.7. Спектр $P(k)$ линейных возмущений плотности в послерекомбинационной Вселенной, восстановленный из наблюдательных данных и экстраполированный к современной эпохе (по работе [Tegmark et al., 2002])

тензорной моды и ничтожно малую векторную моду, что также требует экспериментальной проверки.

Рассмотрим нулевой порядок геометрии. Рисунок 10.8 демонстрирует историю масштабного фактора Вселенной как функцию радиуса внешней кривизны на всем интервале изменения $H_0 \leq H < M_P$. Степенной наклон функции a по аргументу H^{-1} равен

$$\gamma^{-1} \equiv -\frac{d \ln a}{d \ln H},$$

что легко позволяет различать участки с инфляционной ($\gamma < 1$) или замедленной ($\gamma > 1$) динамикой расширения.

За все время эволюции по законам радиационной динамики ($\gamma = 2$) изменение масштабного фактора не превысило 30 порядков. Это значит, что наблюдаемый размер Вселенной, характеризуемый большим безразмерным числом $(H_0 \ell_P)^{-1} \sim 10^{60}$, не может быть объяснен, если не предполагать наличия инфляционной стадии Большого взрыва «В» ($\gamma < 1$), предшествовавшей радиационно-доминированному этапу расширения. В начале

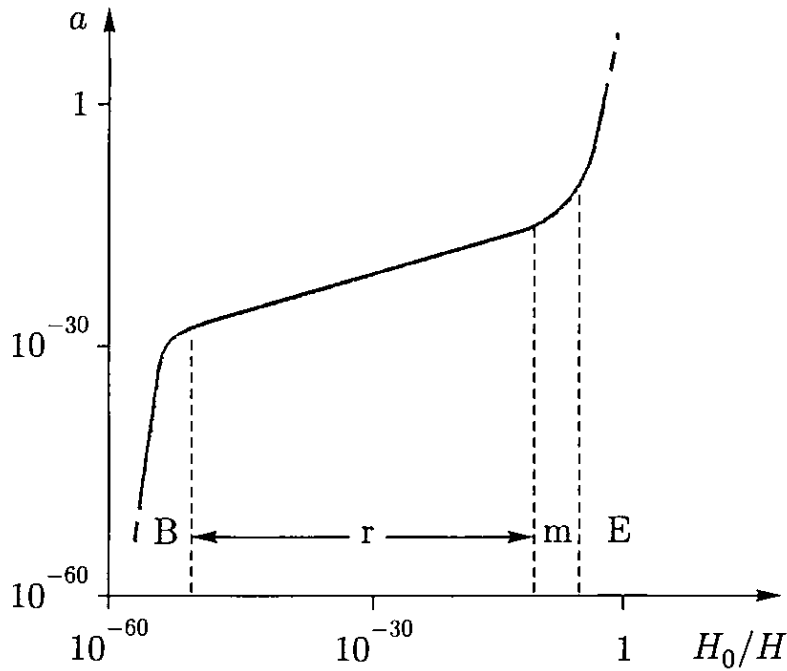


Рис. 10.8. Масштабный фактор Вселенной a как функция нормированного хаббловского радиуса H_0/H . Буквами отмечены эпохи Большого взрыва (В), доминирования радиации (γ), материи (м) и темной энергии (Е). Степенной наклон $\gamma^{-1} \equiv -d \ln a / d \ln H$ больше единицы на стадиях инфляции В и Е и меньше единицы на радиационной γ и материальной м стадиях

γ -стадии космологические масштабы структуры находились в причинно-несвязанных областях пространства¹⁾ — справа от кривой на рис. 10.8. В этом легко убедиться, если поменять местами координатные оси (рис. 10.4) и рассмотреть историю соревнования двух масштабов a/k и H^{-1} в пространстве красных смещений.

Перейдем к первому порядку геометрии. Под действием внешнего переменного потенциала $a(t)$ возмущения кривизны

¹⁾ Удобно исследовать эволюцию масштабов в конформных координатах (η, x) , где мировые линии света наклонены под углом 45° к оси абсцисс η :

$$x = \eta + \text{const}, \quad \eta = \int \frac{dt}{a} = \eta_0 - \int_0^z \frac{dz}{H}$$

(x — координатное расстояние, проходимое светом). При $\gamma = 1$ ($a \propto t$) аффинный параметр η изменяется от $-\infty$ до $+\infty$ (интеграл расходится на обоих пределах) и световые геодезические тождественны световым в пространстве Минковского. При замедленном расширении ($\gamma > 1$) интеграл сходится на нижнем пределе и световой конус прошлого ограничен: точки материи $x = \text{const}$ причинно не связаны друг с другом в прошлом (так называемая проблема горизонта прошлого). При ускоренном расширении ($\gamma < 1$) интеграл сходится на верхнем пределе и ограничен световой конус будущего: точки материи $x = \text{const}$ оказываются со временем в причинно-несвязанных областях пространства.

спонтанно рождаются параметрическим образом из вакуумных флуктуаций метрики (см. главу 8). Наблюдательные данные прямо указывают на квантово-гравитационный эффект рождения космологических возмущений¹⁾. Возникающие таким образом поля имеют четыре характерных признака:

- случайная пространственная фаза;
- гауссова статистика;
- выделенная временная фаза (растущая ветвь);
- ненулевая амплитуда гравитационных волн.

Первые три из них уже подтверждены наблюдениями (в пределах экспериментальной точности). Четвертый — на повестке дня.

Температурный спектр анизотропии реликтового излучения (рис. 10.9) свидетельствует о том, что затравочные возмущения кривизны для образования крупномасштабной структуры Вселенной находились в эпоху рекомбинации в *причинно-несвязанной* зоне и были *синхронизованы* во времени. Модуляционная составляющая спектра реликтового излучения с процентной точностью соответствует растущей эволюционной ветви S-моды в *плоской* модели Фридмана (см. раздел 8.6)²⁾:

$$q \propto \frac{\sin \kappa}{\kappa}, \quad \kappa \simeq \frac{k \eta_{\text{rec}}}{\sqrt{3}}. \quad (10.5.1)$$

Анализ мелкомасштабной зоны карты реликтового излучения (рис 10.5) подтверждает ее гауссовость на высоком уровне достоверности (доля возможной негауссовой примеси ограничена сверху десятой долей процента).

Четвертый пункт списка имеет решающее значение для полного и окончательного подтверждения теории параметрического происхождения мод возмущений (см. главы 5–9). В рамках теории величина T-моды ничем не дискриминирована: во многих моделях предсказываются *значительные* амплитуды гравитационных волн, которые могут быть обнаружены с помощью уже имеющихся инструментальных возможностей. Подчеркнем, что речь здесь идет о наблюдательном подтверждении

¹⁾ Тем самым крупномасштабная структура Вселенной является примером решения проблемы измеримости в квантовой теории поля.

²⁾ Так как $\delta_b \propto \dot{q}$, то масштабы первых пиков в спектре реликтового излучения определяются из условий $\kappa_p \simeq \pi p$, откуда получаем их угловую проекцию на сфере последнего рассеяния, мультипольный индекс $l_p = k \eta_0 \simeq 200p$, где $p = 1, 2, \dots$

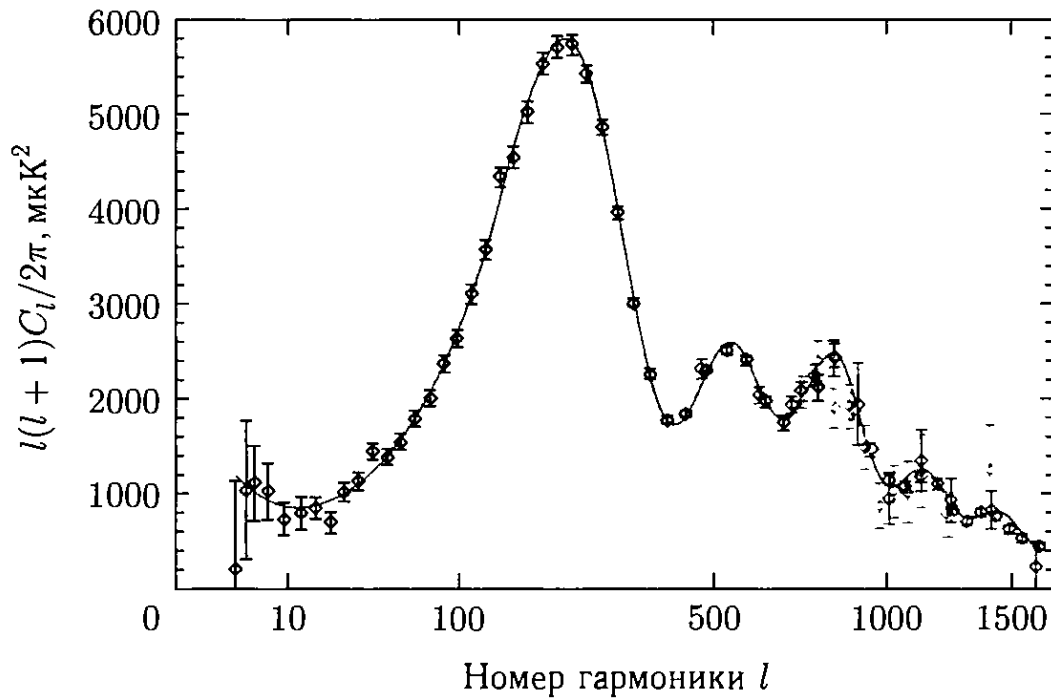


Рис. 10.9. Спектр анизотропии реликтового излучения [Nolta et al., 2008]

именно *квантово-гравитационной теории* происхождения возмущений, а не о моделях инфляции.

Галактики и другие астрономические объекты образовались в результате развития *скалярной* моды возмущений. Анизотропия и поляризация реликтового излучения возникли задолго до галактик в результате воздействия на распределение фотонов *всех трех мод* космологических возмущений. Так как сумма $S + T + V \simeq 10^{-10}$ фиксирована величиной анизотропии реликтового излучения (значение V дается на момент рекомбинации), получаем наблюдательное ограничение величины T -моды в нашей Вселенной ¹⁾:

$$r \equiv \frac{T}{S} < 0,1. \quad (10.5.2)$$

Если бы это неравенство было нарушено, величина возмущений плотности оказалась бы недостаточной для образования галактик. Сам факт их существования означает, что в нашем мире моды T и V подавлены.

Как теория отвечает на поставленный вопрос?

¹⁾ Отметим, что в ряде обзоров и монографий по историческим причинам буквой r обозначена величина, в четыре раза превышающая определение T/S , принятое в нашей книге (см. (8.7.1)–(8.7.5)).

Теория параметрического эффекта предсказывает универсальный результат (см. главу 9):

$$T \simeq \frac{|2 - \gamma| H^2}{M_P^2}, \quad r = 4\gamma \simeq -2n_T \quad (10.5.3)$$

(здесь $\beta = 1$, фактор $2/\pi$ в T-моду опускаем), откуда с учетом (6.6.9) и (10.5.2) получаем ограничения на параметры Вселенной в эпоху генерации возмущений:

$$H < 10^{14} \text{ ГэВ}, \quad (10.5.4)$$

$$E \simeq \sqrt{M_P H} < 10^{16} \text{ ГэВ}, \quad (10.5.5)$$

$$\gamma < 0,02. \quad (10.5.6)$$

Последнее равенство свидетельствует о том, что Большой взрыв носил инфляционный характер ($\gamma < 1$). Этот аргумент в пользу инфляционного Большого взрыва совершенно независим от аргументов, связанных с геометрией нулевого порядка (большая Вселенная, непринципиальность и др.).

Очевидно, одного знания S-моды и ограничения (10.5.2) еще недостаточно для определения *двух* спектральных параметров H и γ в эпоху генерации возмущений. Только обнаружение T-моды даст информацию об энергетическом масштабе и темпе его изменения в эпоху Большого взрыва.

Для более эффективного ограничения спектральных параметров необходимо привлекать априорную информацию, например рассмотреть простейшие динамические модели ранней Вселенной и сопоставить их спектры с наблюдаемыми. В практическом плане достаточно ограничиться моделями, в которых в качестве инфлатона выступает скалярное поле с потенциальной энергией $V(\varphi) \geq 0$, и классифицировать разные формы этого потенциала (см. разделы 7.4, 9.3).

Рисунок 10.10 иллюстрирует классификацию типов инфляций по соответствующим им классам потенциалов:

- инфляция на больших полях:

$$V'' > 0, \quad V(0) = 0 \quad (10.5.7)$$

(характерный пример: $V(\varphi) \propto \varphi^p$);

- инфляция на малых полях:

$$V'' < 0 \quad (10.5.8)$$

(характерный пример: $V(\varphi) = V_0[1 - (\varphi/\varphi_c)^p]$);

- Λ -инфляция:

$$V'' > 0, \quad V(0) > 0 \quad (10.5.9)$$

(характерный пример: $V(\varphi) = V_0[1 + (\varphi/\varphi_c)^p]$).

Имеющиеся наблюдательные данные позволяют существенно ограничить спектральные параметры n и γ (см. раздел 8.7). Что касается величин n_T и τ_S , то они находятся пока в зоне погрешности измерений. На рис. 10.11 представлены теоретически рассчитанные значения n и γ для спектров космологических возмущений, генерируемых в вышеперечисленных классах инфляций. Линия между зеленой и синей зонами соответствует спектрам экспоненциальной инфляции ($V \propto \exp(p\varphi)$), а линия

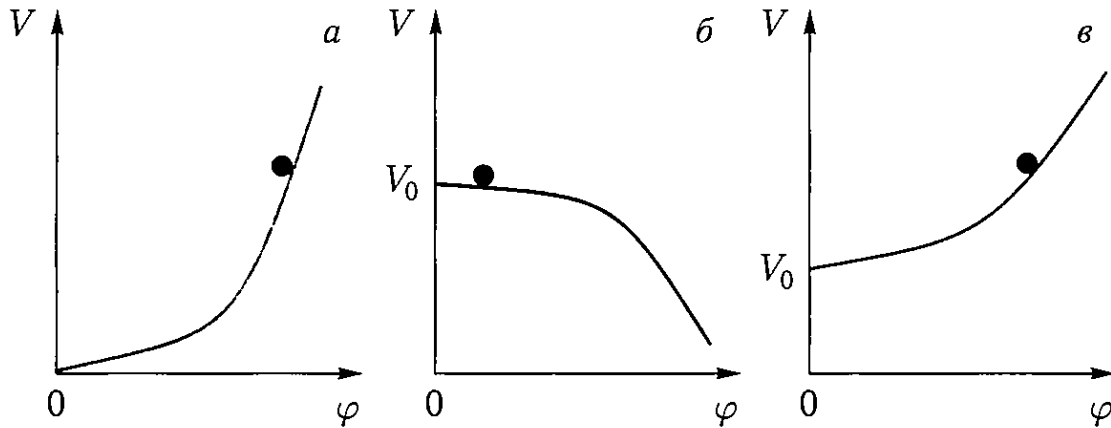


Рис. 10.10. Три класса инфляционных потенциалов $V(\varphi)$: а — инфляция на больших полях, б — инфляция на малых полях, в — Λ -инфляция

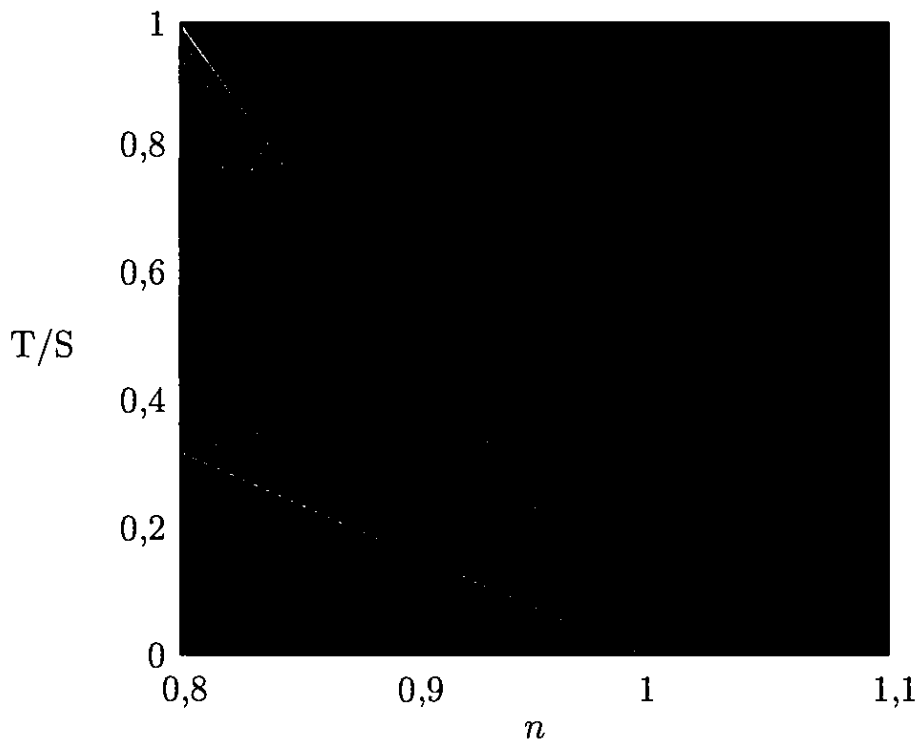


Рис. 10.11. Соотношения между амплитудой первичных гравитационных волн T/S и наклоном спектра возмущений плотности n для трех основных классов инфляционных потенциалов (см. (10.5.7)–(10.5.9)). Зеленая область — инфляция на больших полях, красная — на малых, синяя — Λ -инфляция

между красной и зеленой областями — инфляции на линейном потенциале ($V \propto \varphi$).

Рисунок 10.12 иллюстрирует наблюдательные ограничения в плоскости $n - r$. Белые и черные точки соответствуют предсказаниям инфляционных моделей на массивном скалярном поле ($V \propto \varphi^2$) и поле с самодействием ($V \propto \varphi^4$, зеленая область на рис. 10.11). Для больших точек $\mathcal{N} = 50$, а для малых $\mathcal{N} = 60$ (см. (9.3.7)). Мы видим, что единственный не закрытый пока пример степенной инфляции — это массивное скалярное поле с потенциалом $V = m^2 \varphi^2 / 2$, тогда как модели с более высокими показателями $p > 1$ не удовлетворяют наблюдениям.

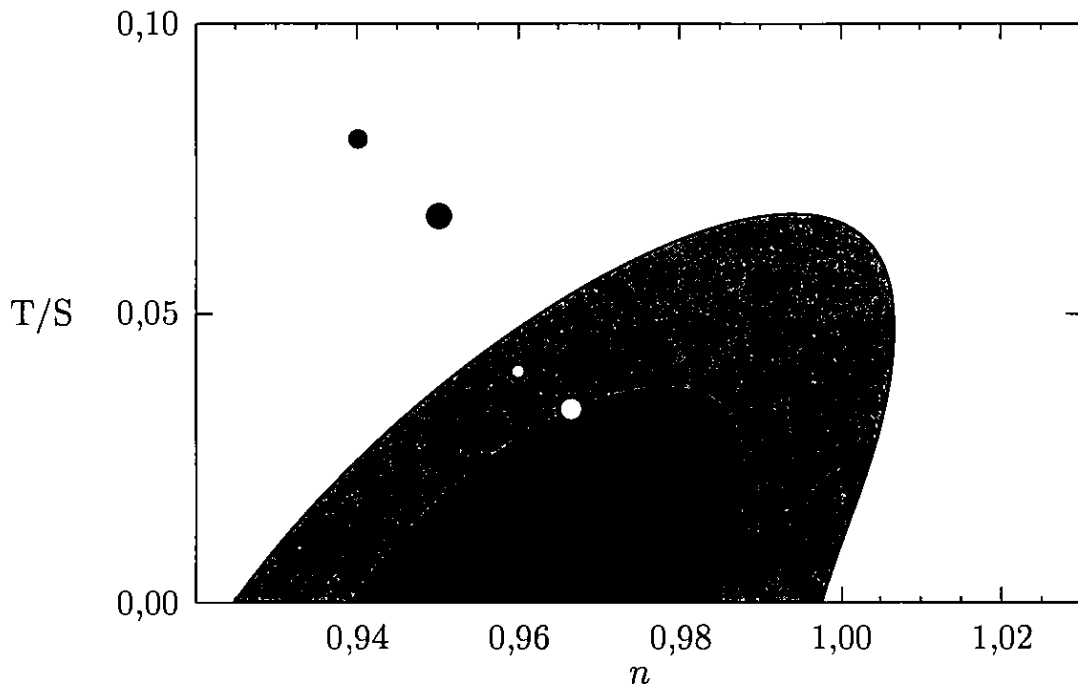


Рис. 10.12. Разрешенные значения T/S и n (цветная область) по данным наблюдений WMAP [Komatsu et al., 2008]. Внешняя граница соответствует уровню достоверности 95%, белые точки — предсказания инфляционной модели на массивном скалярном поле

Остаются еще разрешенные области и у других эволюционных сценариев (ср. рис. 10.11 и рис. 10.12). При этом возможны пока любые наклоны генерируемых спектров возмущений плотности, как красные ($n < 1$), так и синие ($n > 1$). Тем не менее, допустимые значения n оказываются сильно коррелированными с величиной параметра r .

10.6. Темная материя

Наши знания о поздней Вселенной достаточно полны: мы имеем подтвержденную наблюдениями теорию образования и развития структуры, знаем состав материи и основные космологические

параметры (табл. 10.1). Однако имеющиеся данные не позволяют пока однозначно судить о происхождении компонент материи (барионной асимметрии, темной материи, темной энергии).

Знание состава материи во Вселенной — это без преувеличения блестящее достижение астрономии второй половины XX века. Где и в каких формах находится материя во Вселенной?

Светящаяся компонента вещества наблюдается в виде звезд, собранных в галактики разных масс и светимостей, в виде горячего газа в скоплениях галактик (температура газа достигает нескольких кэВ в зависимости от массы скопления), а также в виде облаков молекулярного и ионизованного газа в галактиках (здесь температура варьируется от сотых долей эВ до нескольких эВ). Однако доля всех барионов, попавших в видимые звезды, не превышает 10%, а газовая фракция барионов, наблюдаемая в дисках галактик, их группах и скоплениях, достигает 10–20%.

Основная часть барионов (не менее 50%) находится в состоянии разреженного межгалактического газа с температурой 0,01–0,1 кэВ, а также (до 10–15%) в форме массивных компактных объектов — черных дыр, нейтронных звезд, белых, красных, коричневых и др. карликов, планет и астероидов различных масс. Поскольку эти барионные системы обладают низкой светимостью, за ними закрепилось название *темные барионы*.

В центральных областях ядер галактик находятся сверхмассивные черные дыры, масса которых составляет тысячную долю массы ядра (балджа) хозяйской галактики:

$$M_{\text{bh}} \simeq 10^8 \left(\frac{\sigma}{200 \text{ км/с}} \right)^4 M_{\odot}, \quad (10.6.1)$$

где $\sigma = (\overline{v^2})^{1/2}$ — дисперсия скоростей звезд в галактическом балдже (см. разделы 3.3, 3.5). Рисунок 10.13 демонстрирует эту зависимость в диапазоне масс черных дыр от 10^3 до $10^{10} M_{\odot}$.

Самые маленькие из числа массивных¹⁾ черных дыр ($M_{\text{bh}} \sim 10^2 - 10^3 M_{\odot}$) существуют в центрах шаровых звездных скопле-

¹⁾ Альтернативой массивным дырам являются черные дыры звездных масс ($M_{\text{bh}} < 100 M_{\odot}$), образующиеся при коллапсе ядер массивных звезд на конечной стадии их эволюции. Более массивные черные дыры являются результатом слияний менее массивных дыр и аккреции вещества на них в плотных компактных галактических подсистемах, например в шаровых или центральных звездных скоплениях. Вопрос о происхождении сверхмассивных черных дыр отсылает нас к эпохе образования первых звезд во Вселенной ($z \sim 10-20$) и до конца пока не прояснен.

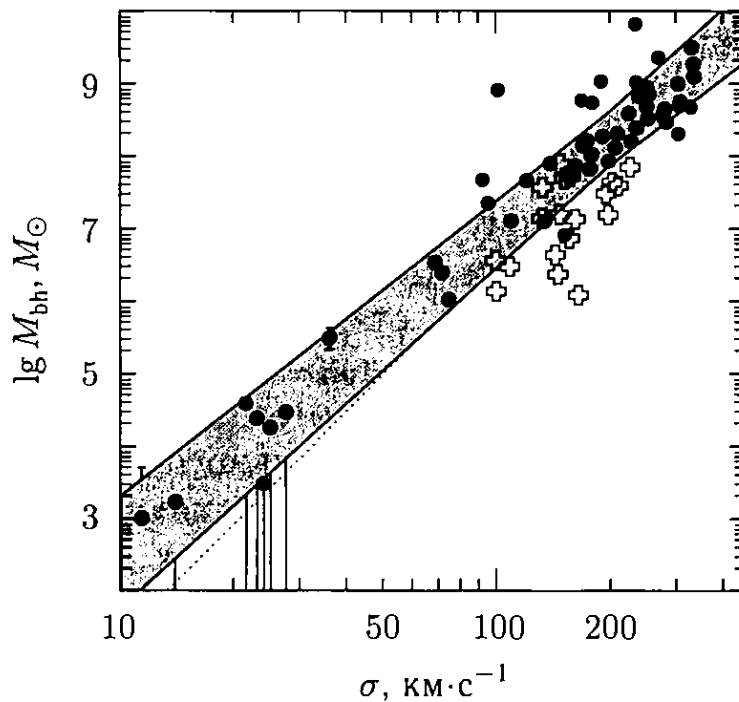


Рис. 10.13. Масса центральной черной дыры M_{bh} как функция дисперсии скорости σ звезд галактического балджа (по данным [Graham, 2008])

ний. Характерные массы черных дыр, находящихся в центральных областях больших галактик¹⁾, порядка $10^7 - 10^9 M_{\odot}$. Самые большие черные дыры расположены в ядрах старых эллиптических галактик, их массы могут достигать $\sim 10^{10} M_{\odot}$. Еще более массивные дыры во Вселенной образоваться не успели. Повсеместно в галактиках существует много малых черных дыр (остатков массивных звезд); их совокупная масса оценивается менее надежно и близка к 0,1% от полной барионной массы галактики.

Количество компактных темных объектов в гало нашей Галактики также невелико. С помощью экспериментов по микролинзированию удалось выяснить, что полная масса МАЧО-объектов²⁾ не превышает 15% всей массы гало (рис. 2.14). Остальное приходится на частицы темной материи небарионной природы.

О существовании небарионной темной материи (ТМ) имеется множество свидетельств, начиная с работ 30-х годов XX века по исследованию дисперсии скоростей галактик в скоплениях. Восстановленная по величине дисперсии скоростей (в предположении о гидростатическом равновесии галактик) масса скопле-

¹⁾ В центре нашей Галактики находится черная дыра умеренной массы $4 \cdot 10^6 M_{\odot}$.

²⁾ МАЧО — от англ. *MACHO* — *M*Assive *C*ompact *H*alo *O*bject.

ний оказывалась в десятки раз больше, чем суммарная масса звезд галактик. Дальнейшие исследования выявили высокие peculiar скорости галактик в сверхскоплениях, которые также требуют огромных масс в сверхскоплениях, в десятки и сотни раз превышающих суммарную массу звезд.

Следующим по времени (но не по важности) аргументом в пользу существования ТМ стало открытие плоских кривых вращения у большинства спиральных галактик (рис. 3.4, 3.5). Длительное существование диска и его спирального узора возможно только при наличии сфероидального гало галактики, стабилизирующего погруженный в него вращающийся диск.

Наиболее надежное определение полных масс скоплений галактик произошло в 80-х годах XX века. Измеренные различными способами (по температуре рентгеновского газа в скоплениях и по гравитационному линзированию удаленных галактик фона близлежащими скоплениями) величины определяемой массы скопления галактик совпали друг с другом, в результате чего скептики, сомневавшиеся в существовании ТМ, сдали свои позиции.

По итогам проведенных в XX веке исследований на сакараментальный вопрос о том, где во Вселенной находится небарионная ТМ, был получен ясный и недвусмысленный ответ: *небарионная темная материя находится в гравитационно-связанных системах (гало)*.

Как и барионы, частицы ТМ — это *нерелятивистские* частицы, их скорость движения много меньше скорости света. Основное различие между темной и обычной материей состоит в том, что частицы ТМ являются *слабовзаимодействующими* и не могут диссипировать, как барионы. В результате, барионы радиационно остывают и оседают к центрам галактических гало, достигая там вращательного равновесия и образуя диски с характерными размерами ~ 10 кпк. Частицы ТМ не могут так быстро терять свою энергию и остаются в гало с типичным масштабом ~ 200 кпк, распределенные вокруг видимого вещества галактик. Наблюдения показывают, что вся ТМ входит в состав гравитационно-связанных гало и основная доля массы всех гало находится в диапазоне $5 \cdot 10^{11} - 5 \cdot 10^{13} M_{\odot}$, что соответствует массивным галактикам и группам галактик (см. разделы 12.6, 12.7). Численное моделирование процессов формирования нелинейной структуры Вселенной подтверждает эти выводы.

Интересен вопрос, существует ли фракция рассеянной ТМ, по аналогии с межгалактическим газом. Ответ известен. «Неучтенного» темного вещества нет: практически вся ТМ (более 90%) входит в гравитационно-связанные системы.

Эту ситуацию можно сопоставить с определением полного количества барионов. Мы знаем, сколько всего материи во Вселенной и сколько барионов по спектру анизотропии реликтового излучения (а именно, по положению, глубине и форме модуляционной части спектра, рис. 10.9). Далее сопоставляем определенное по реликтовому излучению полное количество материи с наблюдаемой динамической массой, входящей в нелинейную структуру Вселенной. Эти две оценки для ТМ совпадают ¹⁾ с текущей точностью лучше 10%. Отметим, что для барионного вещества такое сравнение не проходит, поскольку оно динамически пассивно: его в 5 раз меньше, чем ТМ. Для барионов независимую экспертизу обеспечивает первичный нуклеосинтез (см. раздел 10.8).

Итак, развитие астрономии увенчалось блестящим доказательством того, что полная плотность нерелятивистской материи во Вселенной не превышает трети критической плотности Вселенной:

$$\Omega_m \equiv \frac{\rho_m}{\rho_c} \lesssim 0,3. \quad (10.6.2)$$

При этом доля барионов составляет только шестую часть полного «бюджета» нерелятивистской материи.

В стандартной модели элементарных частиц нет подходящих по свойствам кандидатов в частицы ТМ. Современной физике известны три поколения лептонов (электрон, мюон, тау-лептон и их нейтрино), три пары кварков (u, d, c, s, t и b), античастицы к каждой из частиц и переносчики взаимодействий

¹⁾ В точной формулировке следует говорить об определении по анизотропии реликтового излучения комбинаций параметров $\Omega_m H_0^2$ и $\Omega_b H_0^2$, входящих в правую часть уравнения Фридмана. Анализируя галактические каталоги, мы определяем усеченную комбинацию $\Omega_m H_0$, задающую характерный масштаб пространственного распределения галактик: $k_{eq} \propto \Omega_m H_0^2$ [Мпк] $\propto \Omega_m H_0$ [км·с⁻¹]. Одна степень H_0 уходит в размерность, поскольку *радиальные расстояния* в космологии измеряются в единицах *скорости* с помощью красных смещений галактик. Сопоставление двух измеряемых величин $\Omega_m h^2 \simeq 0,14$ и $\Omega_m h \simeq 0,2$ дает «канонические» значения космологических параметров стандартной модели $\Omega_m \simeq 0,28$ и $H_0 \simeq 70$ км/с·Мпк⁻¹ (точность порядка 10%, см. табл. 10.1). Для дополнительной проверки параметра плотности Ω_m необходимо привлекать независимые данные по измерению постоянной Хаббла H_0 . Это часто встречаемая в космологии ситуация называется проблемой вырождения параметров: общее пространство космологических параметров, даже если нас интересует только один из них, можно эффективно ограничить только при достаточно большом количестве имеющихся физически различных и независимых экспериментальных данных (как минимум, более числа самих параметров, а в стандартной модели их более десятка, см. табл. 10.1).

(фотоны, W^\pm - и Z -бозоны, глюоны). Частицы ТМ следует искать в расширенных (гипотетических) моделях элементарных частиц. Список возможных кандидатов приведен в табл. 10.2 в порядке возрастания массы.

Таблица 10.2. Кандидаты в частицы темной материи

Кандидат	Масса
Гравитоны	10^{-21} эВ
Аксионы	10^{-5} эВ
Стерильные нейтрино	10 кэВ
Зеркальное вещество	1 ГэВ
Массивные частицы	100 ГэВ
Сверхмассивные частицы	10^{13} ГэВ
Монополи и «дефекты»	10^{19} ГэВ
Первичные черные дыры	10^{-16} – $10^{-7} M_\odot$

В силу принципа эквивалентности частицы движутся в гравитационном поле независимо от их массы. Поэтому масса частицы ТМ остается пока неизвестной¹⁾. Тем не менее мы знаем, что частица ТМ — стабильная, электрически нейтральная и слабо-взаимодействующая.

Большинство перечисленных в табл. 10.2 кандидатов никогда не находились в тепловом равновесии с излучением в ранней Вселенной. Исключением является нейтралино — гипотетическая частица с массой ~ 100 ГэВ, возникающая в минимальной суперсимметричной теории. Ее можно назвать *фотоном* суперсимметричного блока частиц. Вычисляемая концентрация нейтралинного гало Галактики на радиусе орбиты Солнца соответствует одной частице на объем чашки чая.

Поиск частиц ТМ ведется как на больших ускорителях, так и в подземных детекторах по регистрации событий их упругого рассеяния на ядрах вещества. Интересным следствием существования частиц ТМ является возможность наблюдения продуктов их аннигиляции (в том случае, если частица тождественна своей античастице, как фотон) в местах повышенной плотности, например

¹⁾ Тогда как плотность ТМ хорошо известна из астрономических данных. Например, в окрестности Солнечной системы средняя плотность ТМ порядка плотности барионного вещества и составляет величину $0,3 \text{ ГэВ}\cdot\text{см}^{-3}$, что на пять порядков превышает критическую плотность Вселенной.

в ядре Галактики. Частицы типа нейтралино могут образовывать минигало с массами вплоть до массы Земли и размером с Солнечную систему. Существование таких гипотетических минигало определяется малоизученной коротковолновой частью спектра мощностей возмущений плотности $P(k)$ (рис. 3.12, 10.7).

Динамику минигало ТМ в Галактике легко рассчитать: они частично разрушаются приливными силами из-за столкновений со звездами, однако большая их часть выживает и перемещается в галактическом гало подобно компактным объектам. Это означает, что Земля может оказаться в остатках подобных минигало. Оценки показывают, что вероятность такого события достаточно высока и может достигать десятка процентов. При этом фактор возрастания плотности частиц ТМ в окрестности Земли по отношению к среднему галактическому фону достигает нескольких десятков, что сильно повышает возможность детектирования ТМ в экспериментах как по упругому рассеянию, так и по аннигиляции. Все это вселяет надежду на возможность обнаружения природы ТМ уже в ближайшее время.

Загадка происхождения ТМ пока не раскрыта, однако у нас есть два косвенных аргумента в пользу того, что ее рождение в ранней Вселенной было связано с возникновением барионной асимметрии.

Первый состоит в том, что отношение барионной плотности к полной плотности ТМ $\sim 1/5$ и не зависит от времени. Это означает, что таким же это соотношение было и в эпоху образования ТМ и барионной асимметрии.

Второй аргумент связан с совпадающими масштабами пространственных распределений барионов и ТМ (см. (6.5.6)). Барионный масштаб — это масштаб первого звукового колебания на сфере последнего рассеяния, совпадающий со звуковым горизонтом на начало рекомбинации k_{rec} . Масштаб ТМ — это хаббловский радиус на момент равенства плотностей излучения и материи k_{eq} . Удивительный экспериментальный факт состоит в том, что два этих масштаба, обусловленные совершенно разными физическими процессами в нашей Вселенной, совпадают.

Все это свидетельствует о том, что обе компоненты материи — темная и барионная — были каким-то образом связаны с излучением и друг с другом в эпоху их генерации в ранней Вселенной.

10.7. Темная энергия

Неравенство (10.6.2) говорит о том, что более двух третей гравитирующей плотности Вселенной находится в форме, не прини-

мавшей участия в гравитационном сгущивании материи:

$$1 - \Omega_m \gtrsim 0,7. \quad (10.7.1)$$

Что же это за компонента, обладающая ничтожным взаимодействием и с плотностью энергии, не зависящей от пространственной координаты (в пределах ошибок измерений)?

Довольно распространенная в течение длительного времени интерпретация состояла в том, что Вселенная обладает *ненулевой* пространственной кривизной. Действительно, если не вводить в рассмотрение другие формы материи, кроме (10.6.2), то мы приходим к открытой космологической модели, для которой уравнение Фридмана в современную эпоху имеет вид:

$$\rho_c = \rho_m + \frac{3}{8\pi G R_x^2}, \quad (10.7.2)$$

где R_x — радиус пространственной кривизны. В соответствии с (10.7.1) кривизна (последнее слагаемое в (10.7.2)) должна доминировать, составляя не менее $0,7\rho_c$.

Однако такая интерпретация (10.6.2) сталкивалась с рядом трудностей. Во-первых, заметную пространственную кривизну крайне трудно совместить с идеей инфляционного Большого взрыва: без специальной подгонки параметров инфляционные модели приводят к чрезвычайно малым значениям пространственной кривизны. Во-вторых, возраст Вселенной в открытой модели (10.7.2) составляет всего 11 млрд лет, тогда как известны шаровые звездные скопления с возрастом 12–13 млрд лет. Имеются и другие аргументы против открытой модели с большой пространственной кривизной.

Если пространственная кривизна отсутствует, то не менее 70% полной плотности Вселенной обусловлено наличием такой формы материи, которая не может быть возмущена гравитационными полями структур и остается *нескученной* в ходе космологической эволюции. Эти свойства означают, что эффективное давление такой материи отрицательно¹⁾ и достаточно велико по

¹⁾ В принципе, если давление среды сопоставимо по модулю с ее плотностью энергии (при $\epsilon + p > 0$), то такая среда эластична и противодействует кластеризации. Примером является излучение, свободно пронизывающее галактики и скопления (скорость света больше второй космической скорости при $r > r_g$). Однако плотность среды с *положительным* давлением будет убывать со временем быстрее, чем плотность нерелятивистской материи, что не может обеспечить ее наблюдаемое доминирование. При условии (10.7.1) эффективное давление среды должно быть *отрицательным*.

абсолютной величине, то есть $p \approx -\rho$. Такая форма материи получила название темной энергии (ТЭ).

Модель с пространственной кривизной была окончательно отвергнута после определения положения первого пика в угловом спектре анизотропии, наиболее чувствительного к значению пространственной кривизны (рис. 10.9). Стало ясно, что трехмерное пространство с высокой точностью евклидово (то есть $R_{\text{кр}}^{-1}$ близко к нулю). Ключевую роль здесь сыграли аэростатные эксперименты BOOMERANG и MAXIMA. Впоследствии и другие эксперименты подтвердили этот результат. Таким образом, суммарная плотность энергии всех видов материи действительно совпадает с критическим значением и, следовательно, ТЭ действительно имеется в природе.

Структурный аргумент (10.6.2), основанный на комбинации результатов измерения анизотропии реликтового излучения и крупномасштабной структуры Вселенной, является однозначным свидетельством существования ТЭ. Другими перспективными методами определения свойств ТЭ являются наблюдаемый интегральный эффект Сакса–Вольфа и геометрические способы измерения расстояний, основанные на наблюдениях далеких сверхновых звезд, ультракомпактных радиоисточников и других ярких объектов (см. раздел 12.3).

Понимание природы ТЭ стоит на первом месте среди проблем космологии, а может быть и всего современного естествознания. Простейшая гипотеза — постоянная «плотность вакуума» — не может быть признана удовлетворительной, поскольку имеет внутренние противоречия. Перечислим некоторые из них.

Первое связано с наблюдаемой иерархией энергетических масштабов фундаментальных взаимодействий: они *возрастают* по мере *уменьшения* константы взаимодействия, эти масштабы приближенно равны 1 ГэВ (сильное взаимодействие), 100 ГэВ (электрослабое взаимодействие) и 10^{19} ГэВ (гравитационное взаимодействие). До открытия ТЭ не было никаких оснований считать, что в природе имеется еще один, гораздо *меньший* масштаб энергий, отвечающий за «новое» *сверхслабое взаимодействие*. Однако характерный масштаб ТЭ составляет $E \sim 10^{-3}$ эВ (рис. 10.1), что никак не соответствует современному пониманию иерархии физических взаимодействий.

Второе противоречие может быть сформулировано как «проблема совпадений» — приблизительное равенство энергий компонент материи в современном мире (в пределах одного порядка величины):

$$\rho_E \approx \rho_M \approx \rho_\nu. \quad (10.7.3)$$

Отношение ρ_b/ρ_M постоянно во времени (от наших дней в прошлое) вплоть до образования барионной асимметрии и генерации ТМ, поскольку ρ_M и ρ_b одинаковым образом уменьшаются при расширении Вселенной (см. раздел 9.6). С другой стороны, ρ_E очень слабо зависит от времени, так что первое из равенств (10.7.3) справедливо только в современную эпоху.

Наконец, само понятие «постоянной положительной плотности вакуума», впервые введенное Э.Б. Глинером, не свободно от внутренних противоречий. С одной стороны — это Λ -член Эйнштейна, допускающий наиболее полную группу движений, свойственную *вакуумной симметрии*. Однако при этом отсутствует другое важное свойства вакуума как *основного энергетического состояния* физических полей: всегда можно взять меньший Λ -член, и он будет описывать состояние с меньшей энергией.

Эта несостоятельность «вакуумного» подхода к проблеме ТЭ явным образом проявляется при рассмотрении квантово-гравитационных эффектов в модели де Ситтера, описывающей полное пространство *постоянной положительной кривизны* $\lambda = \text{const}$ (см. раздел 4.2). В этом случае возникает температура Гиббонса–Хокинга $T \simeq H/(2\pi)$ как реакция вакуума физических полей на гравитационное воздействие положительной плотности $\rho_\lambda \equiv 3\lambda/(8\pi G)$. Данную ситуацию можно сравнить с реакцией вакуума на сильное магнитное поле (например, вблизи нейтронной звезды), где он приобретает свойства среды с коэффициентом преломления $\neq 1$. Если постоянное магнитное поле никто не назовет «вакуумом», то можно ли считать «вакуумом» Λ -член Эйнштейна? Тем более, что это состояние описывается ненулевой температурой.

Гораздо проще представить себе, что «энергия вакуума» путем релаксации упала до нуля уже в начале эволюции Вселенной и задолго до известных стадий. В ходе последующих стадий динамика космологического расширения определялась последовательно разными компонентами материи, доминирующими каждая в свою очередь в силу того, что плотности компонент изменяются со временем по-разному в зависимости от собственных (парциальных) эффективных давлений. На этом пути мы приходим к выводу, что ТЭ — это состояние сверхслабого поля ¹⁾,

¹⁾ В литературе рассматриваются и другие гипотетические причины ускоренного расширения Вселенной, например такие, как отклонение на больших расстояниях и космологических временах реальной гравитации от ньютоновской или проявление дополнительных пространственных измерений и пр. Эти

дождавшегося своей очереди доминирования через 10 млрд лет после Большого взрыва и находящегося сейчас в начальной стадии медленного скатывания. В разделе 12.3 мы рассмотрим способы определения уравнения состояния ТЭ методами наблюдательной космологии, а сейчас вернемся к обсуждению основных свойств ТЭ.

При моделировании ТЭ сверхслабым полем (массивным скалярным полем, «квинтэссенцией», «фантомом» и др.) не возникает проблем с энергетическими масштабами и загадками совпадений, поскольку последовательность стадий доминирования связана исключительно с внутренними параметрами материальных полей (их массой, плотностью и пр.), что ставит ТЭ в один ряд с ТМ, барионами, радиацией.

Факт существования крупномасштабной структуры является ключевым в понимании первого равенства (10.7.3).

Соотношения между плотностями компонент материи

$$\rho_r \lesssim \rho_M, \quad \rho_E \lesssim \rho_M \quad (10.7.4)$$

являются необходимым условием для возможности генерации структуры во Вселенной. Гравитационная неустойчивость развивается в период доминирования нерелятивистской материи и невозможна ни на радиационно-доминированной стадии, ни на стадии доминирования ТЭ. Однако для образования галактик необходимо выполнение еще одно условия: начальная амплитуда возмущений должна быть именно такой, чтобы неоднородности успели вырасти в эпоху открытия «окна» гравитационной неустойчивости (10.7.4).

Итак, крупномасштабная структура Вселенной сформировалась в результате совместного действия двух совершенно различных физических факторов. Эффект квантово-гравитационного рождения безмассовых полей обеспечил затравочные возмущения плотности, а гравитационная неустойчивость превратила их в галактики. Последнее происходит только в том случае, если во Вселенной некоторое время доминирует нерелятивистская материя.

ТЭ подавляет процесс роста возмущений и генерации структуры (см. разделы 12.1, 12.2). Таким образом, для образования галактик необходимо сочетание двух условий: достаточно высокой амплитуды начальных возмущений и достаточно продолжительной материально-доминированной стадии эволюции.

и подобные варианты мы здесь не обсуждаем, поскольку они выходят за рамки стандартной космологии, базирующейся на ОТО.

В нашей Вселенной оба эти условия выполнены: начальные возмущения ($\sim 10^{-5}$) успевают подрасти и образовать крупномасштабную структуру Вселенной в период открытия «окна», который длится от 400 тыс. до 20 млрд лет. Подчеркнем, что выполнение этих условий никак не связано с антропным принципом, поскольку в неравенства (10.7.4) не входят барионы.

Условия формирования звезд в нелинейных гало ТМ выглядят совсем по-иному:

$$\rho_R \ll \rho_b \lesssim \rho_M. \quad (10.7.5)$$

Левое неравенство обеспечивает остывание газа и формирование звезд, а правое — стабилизацию галактических дисков и существование долгоживущих планетных систем.

10.8. Ядерная физика и частицы

Двигаясь назад по времени, мы приходим к эпохе первичного нуклеосинтеза. Характерная температура плазмы в то время составляла ~ 100 кэВ и была близка к массе электрона.

При энергиях больших, чем несколько МэВ, количество электронов (и позитронов) во Вселенной сопоставимо с числом фотонов. Пока температура плазмы выше массы электрона, электроны и позитроны можно считать «безмассовыми» — процессы рождения и аннигиляции электрон-позитронных пар идут в одинаковом темпе:



При температуре ниже 511 кэВ рождение пар подавлено пороговым условием. В результате ниже порога большая часть энтропии электронов и позитронов перекачивается в излучение и его полная энтропия почти утраивается¹⁾:

$$T_\gamma = \left(\frac{11}{4}\right)^{1/3} T_\nu, \quad n_\gamma = \frac{11}{3} n_\nu = 5,5 \left(\frac{T_\gamma}{K}\right)^3. \quad (10.8.2)$$

Одновременно с перекачкой энтропии идет процесс нуклеосинтеза, аналогичный тому, что протекает в звездах. Специфика *первичного* нуклеосинтеза состоит в том, что он идет в условиях убывающей температуры и падающей концентрации частиц: $T \propto a^{-1}$, $\rho_b \propto a^{-3}$.

¹⁾ С учетом ферми-статистики электронов и позитронов энтропия излучения увеличилась в $1 + 2 \cdot (7/8) = 11/4$ раза. Каждая из этих частиц, как и левосторонние нейтрино, имеет по две степени свободы. Напомним, что первоначально температуры излучения и нейтрино были одинаковы.

Почему эти процессы преобразования вещества и излучения и первичный нуклеосинтез идут одновременно? Причина в том, что разница между массами нейтрона и протона и дефект масс образующихся ядер того же порядка, что и масса электрона. Поэтому, как только температура становится меньше нескольких МэВ, равновесная концентрация нейтронов по отношению к протонам падает $\sim \exp(-\Delta m/T)$, где $\Delta \equiv m_n - m_p = 1,3$ эВ. Удивительно, но в это же время происходит *закалка* числа нейтронов относительно протонов: они перестают взаимодействовать с другими частицами и становятся свободными (рис. 10.14). Это происходит потому, что концентрация частиц уменьшается настолько, что нейтрон не успевает найти необходимого партнера, чтобы превратиться обратно в протон.

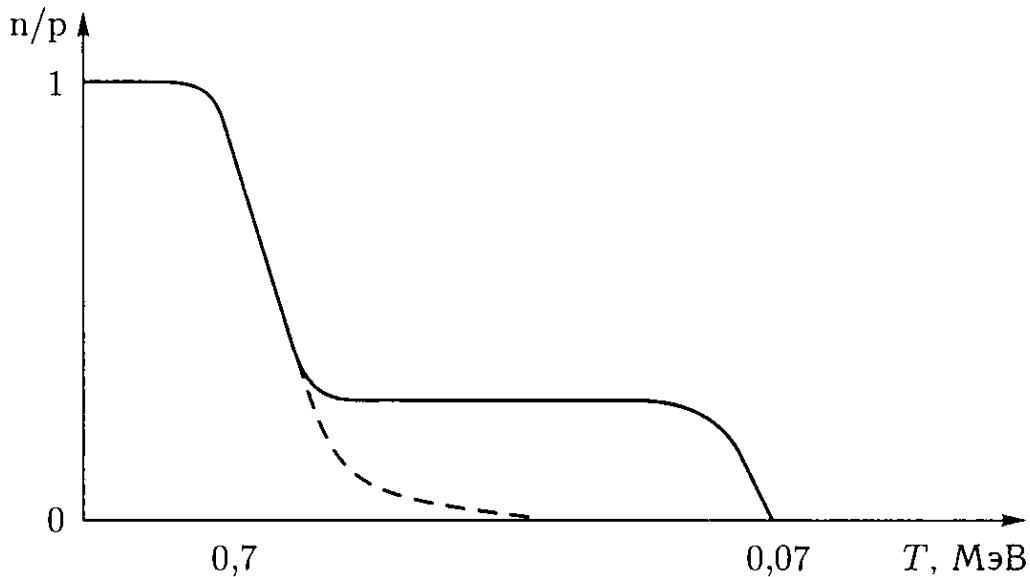
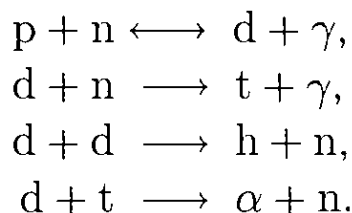


Рис. 10.14. Отношение числа нейтронов к числу протонов в зависимости от температуры плазмы в расширяющейся Вселенной. Штриховая линия — равновесная концентрация, плато — закалка нейтронов, падение — переход в легкие ядра

Как известно, свободный нейтрон живет 886 секунд и затем распадается на более легкие частицы — протон, электрон и антинейтрино. Возраст Вселенной к моменту закалки нейтронов в десятки раз меньше, то есть они живут в свободном состоянии несколько хаббловских эпох и за это время встречаются с протонами, образуя с ними связанные состояния — ядра легких элементов. Цепочка синтеза ядер выглядит следующим образом:



Все начинается с ядра водорода (протона p). Путем захвата нейтрона n образуются тяжелые изотопы водорода (ядра дейтерия d и трития t с атомными массами соответственно 2 и 3), далее — ядра ${}^3\text{He}$ (h), ${}^4\text{He}$ (α -частица) и ${}^7\text{Li}$. На этом процесс прекращается. Для генерации более массивных ядер не хватает времени, потому что среда быстро расширяется и плотность барионов уменьшается. Если бы плотность по-прежнему оставалась большой, как, например, в звездах, то синтезировались бы и более тяжелые ядра, однако в ранней Вселенной этого не происходит (рис. 10.15).

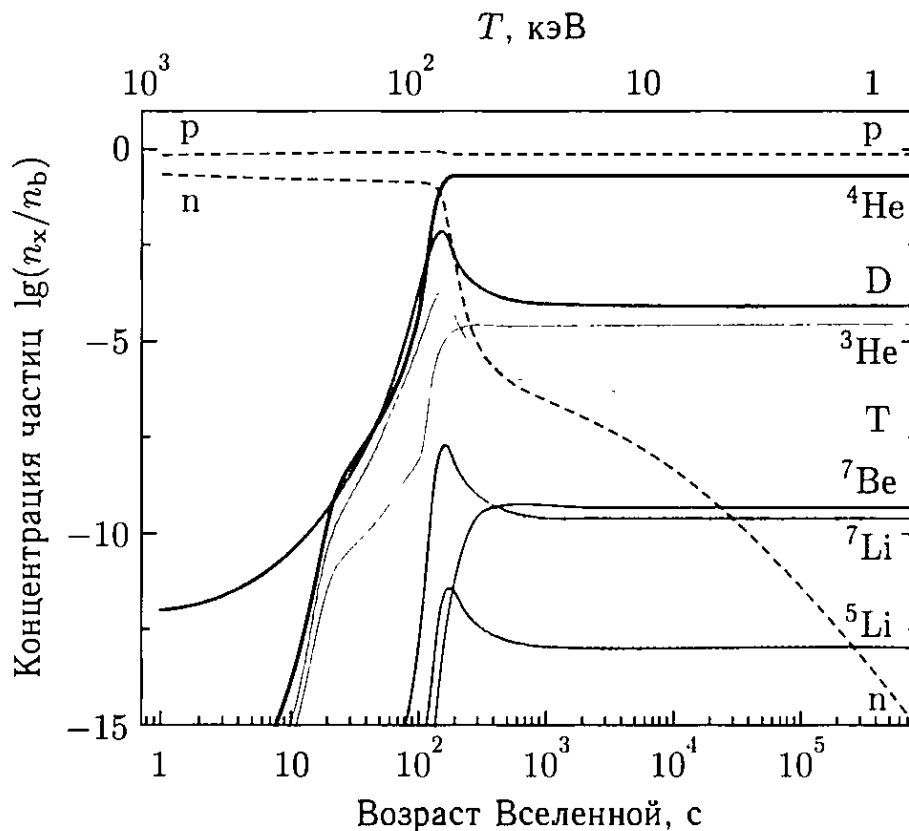


Рис. 10.15. Процесс первичного нуклеосинтеза в космологической стандартной модели [Ivanchik et al., 2001]. Концентрация барионов $n_b = n_p + n_n$

Концентрация ядер на «выходе» чувствительна к двум космологическим параметрам: полной плотности барионов $\Omega_b H_0^2$ при заданной температуре и полной плотности энергии *всех* релятивистских частиц (так как через уравнение Фридмана они совместно с излучением определяют темп расширения радиационно-доминированной Вселенной и, следовательно, изменение температуры плазмы со временем в эпоху нуклеосинтеза). Поскольку температура излучения известна, все сводится к двум параметрам: количеству барионов, приходящихся на один фотон во Вселенной (мера барионной асимметрии), и отношению

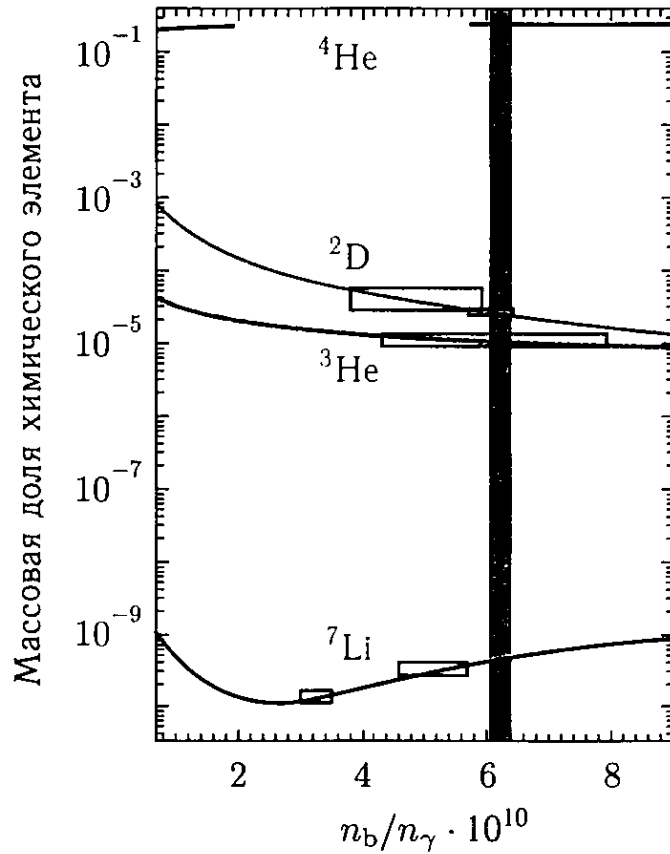


Рис. 10.16. Массовые доли элементов по отношению к водороду при $\kappa_T = 1$. Линии — результаты вычислений, прямоугольники — данные измерений, вертикальная полоса — ограничение по анизотропии реликтового излучения

полной плотности релятивистских частиц в эпоху первичного нуклеосинтеза к полной плотности фотонов и нейтрино и антинейтрино трех ароматов.

Сравнение результатов расчета с наблюдениями дают следующие численные значения этих параметров в нашей Вселенной (рис. 10.16):

$$\eta \equiv \frac{n_b}{n_\gamma} \simeq 6 \cdot 10^{-10}, \quad \kappa_T \equiv \frac{\varepsilon_{\text{rel}}}{\varepsilon_{\gamma+\nu}} \simeq 1. \quad (10.8.3)$$

Для их экспериментального определения наиболее продуктивными в настоящее время являются измерения концентраций атомарного и молекулярного дейтерия в ранней Вселенной по линиям поглощения в спектрах далеких квазаров. Результаты исследований согласуются в пределах ошибок с данными по анизотропии реликтового излучения, что является независимым подтверждением правильности наших выводов об эволюции и составе Вселенной.

Температура фотонов и плотность барионов сегодня хорошо известны. Важнейшим результатом исследования нуклеосинтеза является ограничение на параметр κ_T — полное количество

релятивистских частиц в ранней Вселенной. Он задает суммарную плотность релятивистской материи, которая, в свою очередь, определяет темп изменения сопутствующего объема со временем, а значит, и количественный состав легких химических элементов.

Результат состоит в том, что в ранней Вселенной не было дополнительных релятивистских частиц (в пределах погрешности измерений), кроме известных фотонов и трех ароматов нейтрино левой спиральности (и антинейтрино правой) в полном соответствии со стандартной моделью элементарных частиц. Их плотности и температуры в современной Вселенной равны соответственно для фотонов:

$$n_\gamma = 411 \text{ см}^{-3}, \quad T_\gamma = 2,726 \text{ К},$$

и для нейтрино + антинейтрино каждого аромата:

$$n_\nu = 112 \text{ см}^{-3}, \quad T_\nu = 1,946 \text{ К}.$$

Нерешенным остается пока вопрос о массе нейтрино, влияющей на форму переходной функции $T(k)$ и на процесс образования структуры (см. разделы 6.5, 6.6). В космологии важна сумма масс всех трех ароматов нейтрино, так как она задает *полную* плотность нейтринной компоненты:

$$\rho_\nu = 112 \frac{\sum m_\nu}{\text{см}^3}, \quad \sum m_\nu = 13 f_\nu \text{ эВ},$$

где $f_\nu \equiv \Omega_\nu / \Omega_m$ — доля нейтрино в плотности материи. Имеющиеся наблюдательные данные по крупномасштабной структуре Вселенной и анизотропии реликтового излучения позволяют получить *верхние* пределы на параметры ¹⁾:

$$\sum m_\nu < 1 \text{ эВ}, \quad f_\nu < 0,07. \quad (10.8.4)$$

Эти ограничения намного сильнее пределов, полученных в земных лабораториях. Для сравнения напомним, что $f_b \simeq 0,17$.

Из экспериментов по осцилляциям атмосферных и солнечных нейтрино следует *нижний* предел на сумму масс нейтрино: $\sum m_\nu \geq 0,05 \text{ эВ}$. Остающийся интервал всего в один порядок величины между нижним и верхним пределами внушает оптимизм в том, что масса нейтрино будет открыта космологически-

¹⁾ Неравенство (10.8.4) можно усилить, привлекая наблюдательные данные по распределению галактик и абсорбционных систем L_α : $\sum m_\nu < 0,3 \text{ эВ}$ (уровень статистической достоверности 95 %). Однако такой результат не будет свободен от систематических погрешностей, связанных с нелинейными соотношениями между распределениями светящейся и темной материи во Вселенной.

ми методами с увеличением точности наблюдательных данных. Массовая плотность нейтрино и мала и велика одновременно. Динамически ее влияние незначительно, хотя его и придется учитывать при обработке будущих более точных экспериментальных данных. Однако не следует забывать, что по абсолютной величине космологическая плотность нейтрино сопоставима с плотностью всех звезд во Вселенной: $f_\nu > 4 \cdot 10^{-3}$.

Дальнейшая экстраполяция к началу Вселенной приводит нас к энергиям кварк-глюонного перехода (~ 100 МэВ), когда из тепловой бани частиц конденсировались протоны и нейтроны, и электрослабого фазового перехода (~ 1 ТэВ), когда разделились электромагнитные и слабые взаимодействия. К физике таких энергий у нас остается пока много вопросов, однако принципиальных затруднений здесь нет, так как указанный интервал энергий доступен для исследований на ускорителях частиц на встречных пучках.

Еще более далекая экстраполяция в область высоких энергий приводит нас к окутанной тайнами эпохе появления барионной асимметрии и ТМ. Для решения этих проблем мы можем использовать данные только не прямых экспериментов (см. раздел 10.1), а также более точные результаты исследований в низкоэнергетическом пределе. Однако и здесь есть повод для оптимизма: все эти «временные трудности» *микроскопической* экстраполяции в область высоких энергий существуют только на уровне микропроцессов взаимопревращений материи. И они, бесспорно, требуют расширения стандартной модели частиц. Что же касается *геометрической* экстраполяции в прошлое, то здесь проблем нет: мы можем продолжать движение к энергиям вплоть до планковских, не входя в противоречие с ОТО.

10.9. Уроки экстраполяции

Подведем итог главным урокам геометрической экстраполяции, следующим из ОТО и наблюдательной космологии.

Начнем с геометрии в нулевом порядке. Она адекватно описывается моделью Фридмана и фиксирована значениями космологических параметров, полученными из анализа астрономических данных (табл. 10.1). Подстановка этих параметров в уравнения Фридмана дает хаббловскую функцию $H = H(a)$ и ее производную $\gamma(a)$, допускающие экстраполяцию по масштабности

му фактору a и в прошлое, и в будущее:

$$\frac{H}{H_0} = 10^{61} \frac{H}{M_P} \simeq \left(\frac{10^{-4}}{a^4} + \frac{0,3}{a^3} + 0,7 \right)^{1/2}, \quad (10.9.1)$$

$$\gamma = -\frac{d \ln H}{d \ln a} = \frac{3(\varepsilon + p)}{2\varepsilon} \simeq \frac{2(10^{-4} + 0,2a)}{10^{-4} + 0,3a + 0,7a^4}. \quad (10.9.2)$$

Напомним, что $H_0^{-1} \simeq 14$ млрд лет = 4,3 Гпк = 10^{33} эВ⁻¹ — хаббловский размер Вселенной в разных единицах измерений, $M_P = \ell_P^{-1} = 10^{19}$ ГэВ = 10^{33} см⁻¹ — обратный планковский масштаб. Функция γ описывает темп изменения хаббловского радиуса по отношению к $a \equiv (1+z)^{-1}$. Члены в круглых скобках (10.9.1) соответствуют радиации, нерелятивистской материи и ТЭ.

В этих уравнениях нет свободных параметров. Строгость временной экстраполяции ограничена только точностью численных коэффициентов и скрытыми параметрами в погрешностях наблюдений.

Какие уроки мы можем извлечь из эволюции, описываемой этими простыми уравнениями?

Из уравнения (10.9.1) следует, что известные нам динамические переходы от доминирования радиации к нерелятивистской материи и далее к ТЭ происходили при малых энергиях, хорошо известных атомной физике:

$$T_r \simeq 2,5 \cdot 10^{-4} a^{-1} \text{ эВ.}$$

При дальнейшей экстраполяции

$$H \rightarrow \frac{H_0}{(10a)^2} \simeq \frac{M_P}{(10^{31}a)^2}, \quad \gamma \rightarrow 2 \quad (10.9.3)$$

мы приходим к следующим выводам о свойствах Вселенной.

Прежде всего, Вселенная большая: ее физический размер в начале радиационно-доминированной эпохи расширения, как минимум, в $\sim 10^{30}$ раз превосходил планковский масштаб¹⁾.

¹⁾ Начальный физический размер видимой Вселенной $a_P H_0^{-1} \sim 10^{-3}$ см совпадает по порядку величины с сегодняшней длиной реликтового кванта, который, в свою очередь, сопоставим с сопутствующим планковским масштабом $\ell_P a_P^{-1} \sim 10^{-2}$ см, где $a_P^{-1} \sim 10^{31}$ — максимальный фактор расширения Вселенной в режиме (10.9.1) за все время жизни Вселенной $(H_0 \ell_P)^{-1} = 10^{61}$. Эта оценка вселяет надежду, что в наблюдательных данных может содержаться информация о начале инфляционной стадии, ответственной за наблюдаемую часть мира.

Этот большой фактор можно объяснить существованием короткой, предшествовавшей (10.9.3) инфляционной стадии с числом хаббловских времен $\mathcal{N} \sim \ln 10^{30} = 70$.

Космологические возмущения, индуцировавшие образование структуры, «пришли» из причинно-несвязанной области: масштабы «входят» внутрь горизонта при $\gamma > 1$. Уравнение (10.9.2) описывает *распад* γ от 2 до 0,4. Чтобы объяснить, каким образом начальные возмущения оказались в не причинной зоне, необходим предшествовавший (10.9.3) период расширения с *растущей* функцией γ , начиная от значений меньше единицы.

Мы привели два *прямых* аргумента, указывающих на наличие в истории ранней Вселенной периода с инфляционной динамикой. К ним можно добавить теоретический результат о неизбежном образовании хаббловского потока из первоначально неоднородного распределения гравитирующей среды при ее локально-инфляционном расширении (см. разделы 5.6, 9.6, 10.3). Если первые два аргумента вытекают из экстраполяции ОТО к большим энергиям, то последний является следствием (внутренним свойством) теории гравитации и никак не ограничен масштабами энергий¹⁾. Существует также ряд *косвенных* свидетельств инфляционного характера расширения: плоскостность трехмерного пространства, существование моря реликтовых фотонов, оставшихся после распада инфлатона и др. (см. раздел 10.5).

Важнейшие динамические свойства Вселенной мы узнаем из описания геометрии в первом порядке. Говоря о Вселенной, мы подразумеваем хаббловский поток. Создание Вселенной — это *создание* хаббловского потока. Противоположный ему процесс *разрушения* хаббловского потока мы называем гравитационной неустойчивостью и, в конечном счете, образованием структуры во Вселенной.

Образование галактик и скоплений галактик в причинно-связанных областях хаббловского потока требует выполнения двух условий: наличия затравочных возмущений плотности и гравитационной неустойчивости для роста контраста плотности до нелинейных значений. Оба условия физически независимы и удивительным образом выполняются в нашей Вселенной.

Гравитационная неустойчивость может динамически развиваться только в нерелятивистской среде, состоящей из слабо-взаимодействующих частиц, поскольку наличие давления препятствует росту неоднородностей плотности. В нашей Вселенной

¹⁾ Напомним, что восстановление хаббловского потока происходит в наши дни (см. главу 12).

роль такой среды выполняет ТМ. Однако логарифмический рост контраста плотности ТМ на стадии доминирования радиации недостаточен для развития структур (см. раздел 6.5). Для ускорения темпа роста возмущений необходимо доминирование самой ТМ, что неизбежно наступает в некоторый момент времени в силу того, что плотность ТМ уменьшается медленнее, чем плотность радиации.

Однако, если во Вселенной существуют другие формы материи с эффективным отрицательным давлением, то их плотность энергии изменяется со временем еще медленнее, чем у ТМ, и они рано или поздно прекращают стадию доминирования ТМ и останавливают развитие гравитационной неустойчивости. Именно так и происходит в нашей Вселенной. Три с половиной миллиарда лет назад, когда плотность ТЭ начала динамически доминировать во Вселенной¹⁾, началось неспешное (по меркам текущего хаббловского радиуса) закрытие «окна» гравитационной неустойчивости. Сам процесс подавления роста линейных неоднородностей развивается за время динамического влияния ТЭ, которое составляет $H_E^{-1} \simeq 17$ млрд лет (см. (7.4.1)).

Таким образом, наблюдаемые значения параметров плотности обеспечили процесс гравитационной неустойчивости на конечном интервале времени от нескольких сотен тысяч до двух десятков миллиардов лет (см. разделы 10.7, 12.1). Однако этого еще недостаточно для образования галактик: затравочная амплитуда возмущений должна быть именно такой, чтобы неоднородности *успели* вырасти и распасться на гравитационно-связанные системы в период открытия «окна» неустойчивости. Если условия (10.7.4) связаны с составом Вселенной, то затравочные возмущения относятся к ее геометрии.

Квантово-гравитационная теория генерации космологических возмущений плотности была создана в 1980 году и выдержала испытание временем. Изложение этой теории составляет основу книги. Гейзенберговское состояние элементарных гармонических осцилляторов с импульсом \mathbf{k} , описывающих эволюцию скалярных степеней свободы возмущений кривизны в геометрии Фридмана, соответствует начальным вакуумным флуктуациям плотности однородной гравитирующей среды. Причиной рождения является гравитационное поле хаббловского потока расширяющейся материи, которое параметрически воздействует на элементарные осцилляторы. Это общий эффект рождения физических безмас-

¹⁾ При этом параметр γ пересек уровень $\gamma = 1$ несколько ранее, около 7 млрд лет назад

совых степеней свободы с конечной длиной волны во внешнем нестационарном гравитационном поле.

Неоднородности кривизны в геометрии Фридмана описываются теорией безмассового скалярного поля (см. разделы 5.3, 5.4 и 7.1) и представляют тем самым огромный интерес, поскольку, с одной стороны, имеется хорошо понимаемая физика их происхождения, а с другой — они послужили затравками для образования галактик и наблюдаемой крупномасштабной структуры Вселенной. Рожденные параметрическим образом космологические поля возмущений обладают целым рядом характерных признаков и свойств, которые подтверждаются наблюдениями.

Что это за свойства?

Прежде всего — это наблюдаемые *случайная пространственная фаза* и *гауссова статистика* догалактических возмущений поля плотности. Данное свойство изначально присуще скалярным вакуумным флуктуациям, и оно *линейным* образом передается рожденным неоднородностям кривизны и связанной с ней плотности (см. главу 8).

Другой факт — *выделенная временная фаза* скалярных неоднородностей, соответствующая одной (растущей) из двух эволюционных ветвей решения. Это общее свойство затравочных возмущений плотности для образования космологической структуры, было, по-видимому, впервые предсказано А. Д. Сахаровым в 1965 году. Однако долгое время оставался открытым вопрос, как физически образуется именно *растущая* мода линейных возмущений кривизны с достаточно большой амплитудой ($\sim 10^{-5}$) из первоначально *случайных по временной фазе* и ничтожных по величине вакуумных флуктуаций плотности. Данное свойство внутренне присуще самому параметрическому эффекту, усиливающему *только* растущее решение (см. разделы 7.4, 8.6) до практически *произвольной* амплитуды (но меньше единицы) в зависимости от космологических параметров хаббловского потока (см. разделы 8.7, 8.8). Это важнейшее свойство теории подтверждается наблюдаемыми звуковыми модуляциями соответствующих спектров мощности (см. разделы 6.6, 10.5, рис. 6.3, 10.9).

Часто перечисленные свойства связывают с инфляцией. Однако это не так. *Случайная* пространственная и *детерминированная* временная фазы рожденных неоднородностей являются *визитной карточкой* квантово-гравитационного эффекта рождения безмассовых полей, к которым принадлежат S- и T-моды космологических возмущений. Эти два свойства присущи *любым*

динамическим режимам хаббловского расширения, включая, в частности, и инфляционные.

Какие свойства геометрии первого порядка свидетельствуют об инфляционном Большом взрыве?

Наблюдаемое верхнее ограничение на амплитуду космологических гравитационных волн, указывающее на малость параметра $\gamma < 1$ в эпоху генерации возмущений (см. (10.5.2)), можно считать косвенным аргументом в пользу инфляционной ранней Вселенной. Окончательное доказательство первичной инфляции по первому порядку геометрии станет возможным только после обнаружения Т-моды космологических возмущений и проверки теоретически предсказанной связи показателя наклона Т-спектра с величиной параметра $r \simeq -2n_T$ (см. (10.5.3)). Это может произойти уже в ближайшее время при обработке данных космических экспериментов по измерению анизотропии и поляризации реликтового излучения ¹⁾.

Часто в качестве доказательства инфляции приводят наблюдаемую плоскую форму S-спектра ($n_S \simeq 0$). Напомним, что такой аргумент некорректен: инфляционные модели позволяют получить *произвольный* наклон S-спектра в зависимости от значения параметров (см. главу 9). Другое дело Т-спектр: он всегда близок к плоскому в инфляционных сценариях, и его наклон жестко связан с r -параметром. Поэтому обнаружение Т-моды послужило бы прямым доказательством (или опровержением) первичной инфляции ²⁾.

Предельно возможная геометрическая экстраполяция к моменту окончания инфляционного взрыва и/или началу радиационно-доминированной стадии расширения достигается при условии, что амплитуда космологических гравитационных волн находится вблизи наблюдаемого верхнего ограничения (10.5.5):

$$E \rightarrow 10^{16} \text{ ГэВ} \gtrsim 10^{-4} a^{-1} \text{ эВ}. \quad (10.9.4)$$

Отсюда получаем оценку минимального масштабного фактора Вселенной:

$$a_{\min} \sim 10^{-29},$$

¹⁾ Напомним в этой связи, что простейшая модель инфляции на массивном скалярном поле предсказывает величину r всего лишь в четыре раза меньше современного верхнего ограничения (см. раздел 9.4, рис. 10.12).

²⁾ Напомним, что гарантии детектирования Т-моды не существует: некоторые сценарии предсказывают ее незначительную амплитуду (см. (10.5.8), рис. 10.11 и 10.12).

и минимального хаббловского (сопутствующего) радиуса Вселенной к моменту окончания инфляции:

$$k_{\min}^{-1} \equiv \overline{H}_{\min}^{-1} \sim 1 \text{ м.}$$

Таким образом, наблюдательная космология дает уникальную информацию не только о физике ранней Вселенной и высоких энергий, недоступных ускорительным экспериментам. С ее помощью проверяется теория гравитации на 30 (!) порядков по масштабному фактору и энергетической шкале, что является беспрецедентным опытом науки.

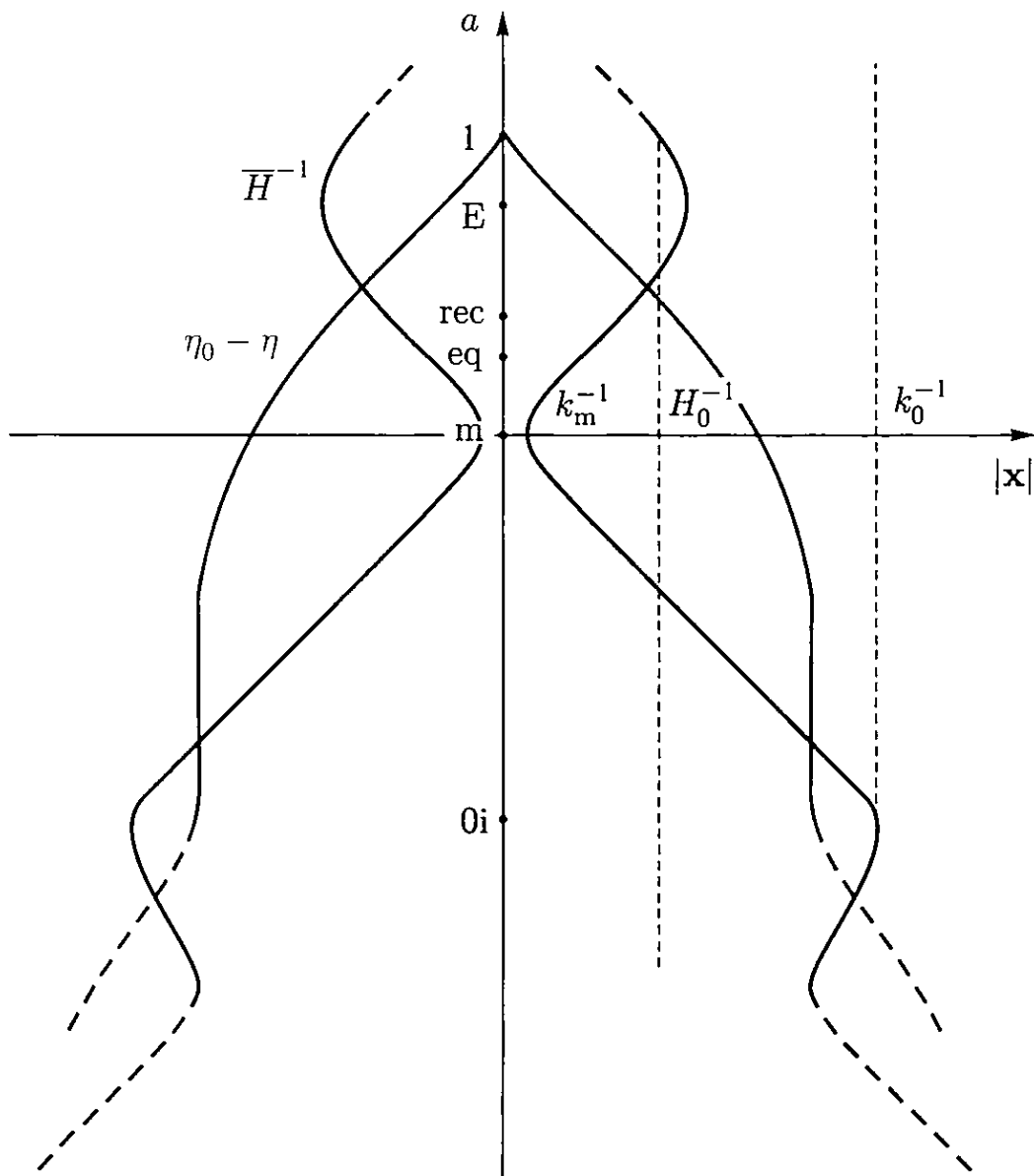


Рис. 10.17. Окрестность наблюдателя в координатах $(a, |x|)$. Синяя линия — хаббловский радиус, красная — световой конус прошлого, m — конец первичной инфляции, E — начало новой инфляции, k_0^{-1} — размер фридмановского мира

Наблюдательных оснований для дальнейшей экстраполяции к более высоким энергиям, чем (10.9.4), у нас пока нет. Мы не

знаем, сколько длилась последняя стадия инфляционного взрыва (их могло быть несколько, см. раздел 7.4), которая сгенерировала космологические возмущения для образования наблюдаемой структуры Вселенной. Если продолжительность этой стадии была не очень большой ($N_{0i} < 70$), то следы доинфляционной динамики должны проявиться в аномалиях крупномасштабного спектра неоднородностей, в частности в квадрупольной аномалии (см. раздел 9.4). Точная космология способна будет прояснить и этот вопрос. Рисунок 10.17 иллюстрирует область экстраполяции наблюдательной космологии.

Уроки экстраполяции обогащают нас знаниями об истории Вселенной и могут быть применены для экстраполяции в будущее. Существование инфляционной стадии в ранней Вселенной подсказывает разгадку физической природы ТЭ.

Действительно, за 13,7 млрд лет своего существования Вселенная прошла по крайней мере *одну* радиационную ($\gamma = 2$) и *одну* материальную ($\gamma = 3/2$) стадии, но *дважды* находилась в состоянии инфляции — первичной (Большой взрыв) и новой (ТЭ). Стадий с $\gamma < 1$ могло быть и больше, и физические причины их возникновения и распада могли быть схожими.

Анализируя крупномасштабную структуру Вселенной и продукты распада первичной инфляции (реликтовые фотоны, барионы и др.), мы изучаем физику *конца* инфляционной стадии Большого взрыва. В то же время мы все свидетели *начала* новой стадии инфляции, управляемой ТЭ. Предполагая схожие физические причины обеих стадий, мы приходим к выводу, что каждая из них имела свое начало и свой конец.

Проиллюстрируем это на примере *слабовзаимодействующих массивных полей* с разными начальными значениями и массами (см. раздел 7.4).

Каждая инфляционная стадия начинается с доминирования очередного массивного поля, протекает в режиме медленного скатывания и заканчивается переходом в стадию осцилляций (доминирования нерелятивистских частиц данного поля) или распадом на маломассивные частицы. К началу очередной инфляционной стадии соответствующее поле «заморожено» и его плотность практически неотличима от космологического Λ -члена. В ходе доминирования поле «ощущает» воздействие собственной гравитации, приходит в движение и дрейфует к точке равновесия. Этот динамический переход от состояния Λ -члена к медленному скатыванию происходит в течение одной хаббловской эпохи в начале данной стадии инфляции.

В этом примере история Вселенной представляет собой историю возникновения и распада (релаксации) массивных полей. ТЭ может быть одним из таких массивных полей. Она принципиально не отличается от инфлатона в ранней Вселенной, имея по сравнению с ним лишь другие значения параметров. Отметим, что первая хаббловская эпоха доминирования ТЭ еще не прошла, она только наступила и будет длиться еще 13 млрд лет. Тем не менее, найти наблюдательные свидетельства эволюции ТЭ представляется возможным (если начальное поле φ_{0E} не очень велико). Мы вернемся к этому вопросу в главе 11.

Наиболее точная наблюдательная информация о геометрии Вселенной связана с исследованием реликтового излучения. Удивительный подарок природы состоит в том, что в настоящее время в спектре реликтового излучения мы можем воочию наблюдать возмущения, бывшие когда-то квантовыми. Это, кроме всего прочего, доказывает, что наши физические представления о микромире остаются справедливыми вплоть до периода, когда величина H была большой и происходила интенсивная генерация возмущений. Если бы человечество появилось несколько позже (по космологическим меркам), когда реликтовое излучение остыло, например, до 0,03 К, то наши представления о Вселенной были бы гораздо беднее. И в современную эпоху реликтовое излучение было открыто на уровне шума радиометров. Всем нам очень повезло, что мы не только его наблюдаем, но и с помощью точных приборов можем исследовать его анизотропию и поляризацию.

ГЕНЕРАЦИЯ АНИЗОТРОПИИ РЕЛИКТОВОГО ИЗЛУЧЕНИЯ

Через 400 лет после окончания инфляции горячая плазма рекомбинирует и превращается в нейтральный водород. Реликтовое излучение (РИ), генетически связанное с плазмой из-за томсоновского рассеяния фотонов на свободных электронах, отсоединяется от вещества. Вселенная становится прозрачной для фотонов, и большая их часть приходит к нам без рассеяния, принося прямую информацию о пространственном распределении барионов и возмущениях кривизны в ту раннюю эпоху, когда еще не было галактик и существовали лишь их прародители — первичные космологические неоднородности. Эти неоднородности мы наблюдаем сегодня в виде угловых вариаций температуры и поляризации РИ на небесной сфере.

В данной главе мы покажем, каким образом поле начальных космологических возмущений трансформируется в анизотропию РИ.

11.1. Характерные масштабы анизотропии

Анизотропия реликтового излучения формируется в эпоху космологической *рекомбинации* водорода. В этот период ($z_{\text{rec}} \simeq 1100$) температура плазмы падает до 3000 К и свободные электроны захватываются протонами, образуя атомы нейтрального водорода.

Из-за уменьшения пространственной концентрации электронов длина свободного пробега фотонов возрастает и, начиная с некоторого момента ($z \simeq 1000$), становится больше хаббловского радиуса Вселенной: вещество и излучение разъединяются. Кванты РИ перестают рассеиваться на электронах, поскольку они их больше не находят, и далее распространяются свободно.

Реликтовые фотоны приносят информацию о плотности и движении барионов в момент *последнего рассеяния* и о метрике на всем пути следования. Основная часть этой информации — пространственная кривизна Вселенной и структура поля скалярных неоднородностей, однако определенный вклад в возмущения метрики дают и гравитационные волны. Наблюдая РИ в разных

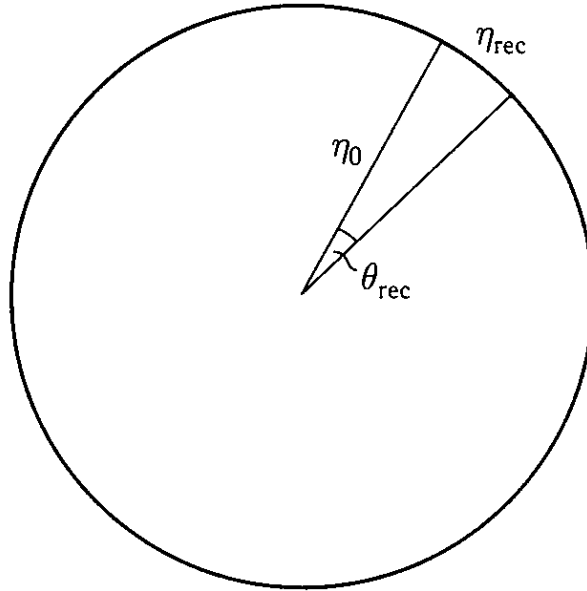


Рис. 11.1. Угловой масштаб рекомбинации на сфере последнего рассеяния

направлениях на небесной сфере, мы изучаем геометрию ранней Вселенной и поле начальных космологических возмущений.

Физический масштаб, «отпечатанный» на карте РИ, — это космологический горизонт, или хаббловский размер Вселенной в момент отрыва излучения от вещества (рис. 11.1). Угол, под которым мы видим этот размер на сфере последнего рассеяния в евклидовом пространстве, равен

$$\theta_{\text{rec}} = \frac{\eta_{\text{rec}}}{\eta_0 - \eta_{\text{rec}}} \simeq \frac{\eta_{\text{rec}}}{\eta_0} \simeq z_{\text{rec}}^{-1/2} \simeq 2^\circ. \quad (11.1.1)$$

Поскольку главную роль в генерации анизотропии РИ играют скалярные возмущения (их амплитуда больше амплитуды первичных гравитационных волн (10.5.2)), то характерный масштаб анизотропии — это звуковой горизонт, или характерная длина волны звуковых колебаний плазмы непосредственно перед отсоединением фотонов. В то время плотность энергии излучения была больше космологической плотности барионов (хотя уже и меньше плотности темной материи) и скорость звука составляла половину скорости света. Отсюда получаем угловой размер звукового горизонта в эпоху рекомбинации:

$$\theta_{\text{srec}} \simeq \beta_{\text{R}} \theta_{\text{rec}} \simeq 1^\circ, \quad (11.1.2)$$

где $\beta_{\text{R}} = [3(1 + 3f_{\text{b}} y_{\text{rec}})]^{-1/2} \simeq \sqrt{8/33} \simeq 0,5$ (см. раздел 6.5).

Имеется еще и третий характерный угловой размер, наблюдаемый на сфере последнего рассеяния: это угол, под которым мы видим собственный физический масштаб спектра скалярных возмущений η_{eq} (горизонт в момент равенства излучения и ве-

щества). Как мы уже отмечали ранее (см. (6.5.6)), он удивительным образом совпадает со звуковым горизонтом (11.1.2):

$$\theta_{\text{eq}} = \frac{\eta_{\text{eq}}}{\eta_0} \simeq 1^\circ. \quad (11.1.3)$$

Все три угла одного порядка, они задают типичный угловой масштаб неоднородностей РИ. На реальной карте РИ (рис. 10.5) мы видим эти характерные холодные и горячие пятна размером порядка 1° . Связь между пространственной частотой k в эпоху отрыва фотонов от плазмы и двумерной частотой на небесной сфере l (индекс мультиполя сферической гармоник) имеет следующий вид:

$$l = k\eta_0 \simeq \frac{\pi}{\theta}. \quad (11.1.4)$$

Дипольная гармоника ($l = 1$) соответствует угловому масштабу 180° , квадрупольная ($l = 2$) — углу 90° . Для звукового горизонта в момент рекомбинации $l \simeq 200$. Таких причинно-несвязанных областей на сфере последнего рассеяния несколько десятков тысяч.

Рекомбинация — это не событие, а процесс, протекающий во времени. Степень ионизации плазмы (доля свободных электронов) уменьшается вдвое за характерный промежуток времени, называемой *полушириной рекомбинации*. В стандартной модели она составляет десятые доли текущего космологического возраста, то есть рекомбинация протекает достаточно быстро, за время меньше хаббловского. По отношению к этому масштабу анизотропию можно разбить на две части: крупномасштабную ($\theta > 1^\circ$) и мелкомасштабную ($\theta < 1^\circ$). В первом случае немгновенностью рекомбинации можно пренебречь, а во втором — необходимо учитывать развитие процесса рекомбинации во времени, что приведет к затуханию наблюдаемой анизотропии.

11.2. Приближение мгновенной рекомбинации

Начнем с крупномасштабной анизотропии РИ: $l < 200$.

Сначала нам нужно определить гиперповерхность, на которой произошел отрыв квантов от вещества. В разных точках он происходил в разное время относительно гиперповерхности $t = \text{const}$, где t — фридмановское время.

В крупномасштабном приближении рекомбинацию можно считать мгновенной и совпадающей с гиперповерхностью последнего рассеяния:

$$t_{\text{ls}} = t_{\text{rec}} = t + a\delta\eta, \quad \delta\eta = \delta\eta(\mathbf{x}). \quad (11.2.1)$$

Физически событие последнего рассеяния определяется условием равенства температуры плазмы и температуры рекомбинации, которая близка к 3000 К. При большей температуре электроны были свободные, при меньшей они связаны с протонами в атомах водорода. Через соотношение $\varepsilon_\gamma \propto T^4$ температуру можно перевести в плотность энергии излучения. Тогда постоянство температуры в момент последнего рассеяния означает постоянство плотности гамма-квантов:

$$\varepsilon_\gamma(\eta_{\text{rec}}, \mathbf{x}) = \text{const} = \varepsilon_\gamma(\eta, \mathbf{x}) + \varepsilon'_\gamma \delta\eta = \varepsilon_\gamma(\eta) + \delta\varepsilon_\gamma - 4\bar{N}\varepsilon_\gamma \delta\eta.$$

Отсюда находим функцию $\delta\eta(\mathbf{x})$:

$$\delta\eta = \frac{\delta}{3\bar{N}}, \quad (11.2.2)$$

где

$$\delta = \delta(\eta_{\text{rec}}, \mathbf{x}) \equiv \frac{3}{4} \frac{\delta\varepsilon_\gamma}{\varepsilon_\gamma} = \frac{\delta\varepsilon_b}{\varepsilon_b}$$

— адиабатическое возмущение плотности вещества непосредственно перед рекомбинацией плазмы, все величины взяты в момент времени $z_{\text{rec}} = 1100$.

Итак, при $\eta < \eta_{\text{rec}}$ плазма была идеальной средой: кванты были заперты электронами из-за томсоновского рассеяния, а электроны сцеплены с протонами кулоновским взаимодействием. Поэтому плотность излучения сопутствовала плотности барионов.

При $\eta > \eta_{\text{rec}}$ свет свободно распространяется в пространстве-времени и его необходимо рассматривать в кинетическом приближении.

В кинетическом подходе фотоны описываются функцией распределения $f = f(x^\mu, \mathbf{p})$, задающей их число в элементе фазового объема:

$$dN = f d\mathbf{x} d\mathbf{p}, \quad (11.2.3)$$

где $x^\mu = (\eta, \mathbf{x})$ — пространственно-временные координаты, $p_\mu = (p_0, p_i)$ — ковариантный импульс фотона ($\mathbf{p} \equiv p_i$). В отсутствие столкновений функция распределения фотонов является скаляром в фазовом пространстве — она сохраняется вдоль мировых линий фотонов и описывается бесстолкновительным уравнением Больцмана:

$$\frac{df}{d\eta} = \frac{\partial f}{\partial \eta} + \frac{\partial f}{\partial \mathbf{x}} \mathbf{x}' + \frac{\partial f}{\partial \mathbf{p}} \mathbf{p}' = 0. \quad (11.2.4)$$

Здесь в качестве аффинного параметра взято конформное время η , а $\mathbf{x} = \mathbf{x}(\eta)$ — геодезическая траектория фотона.

Вместо трех компонент импульса \mathbf{p} удобно использовать частоту фотона, измеренную наблюдателем, движущимся вместе с веществом,

$$\omega = p_\mu u^\mu, \quad (11.2.5)$$

и направление на небесной сфере, откуда пришел фотон,

$$\mathbf{e} = \frac{\mathbf{p}}{p}, \quad (11.2.6)$$

где u^μ — скорость вещества, $p \equiv \sqrt{p_1^2 + p_2^2 + p_3^2}$.

В момент последнего рассеяния функция распределения фотонов имеет планковский вид, зависящий только от отношения ω/T . Поскольку далее она не изменяется (см. (11.2.4)), то в любой точке x^μ при $\eta > \eta_{\text{ls}}$ функция распределения зависит от инвариантов ω/T , переносимых вдоль мировых линий фотонов в данную точку x^μ . Таким образом, в момент приема имеем:

$$f = \frac{2}{e^{\omega/T} - 1}, \quad (11.2.7)$$

где $T = T(x^\mu, \mathbf{e})$ — температура фотонов в точке x^μ в направлении \mathbf{e} на небесной сфере.

Мы видим, что анизотропия излучения не зависит от частоты принимаемых фотонов:

$$\frac{\omega}{\omega_{\text{ls}}} = \frac{T}{T_{\text{ls}}}, \quad (11.2.8)$$

однако температура фотонов, приходящих к наблюдателю с разных направлений, разная, поскольку на них влияют разные пространственно-временные области поля первичных возмущений. Наша задача заключается в определении связи анизотропии с космологическими неоднородностями и восстановлении распределения температуры РИ на небесной сфере.

В соответствии с линейной теорией возмущений температуру фотонов можно представить в виде:

$$T = T_0 (1 + \delta_T), \quad \delta_T = \delta_T(x^\mu, \mathbf{e}), \quad (11.2.9)$$

где $T_0 = T_0(\eta) \propto a^{-1}$ — невозмущенная температура в модели Фридмана, а $\delta_T \equiv \delta T/T$ — анизотропия температуры, зависящая от направления прихода фотонов и положения наблюдателя.

Как будут различаться результаты измерений разных наблюдателей, движущихся в окрестности события x^μ ?

При лоренцевом бусте системы отсчета $\tilde{x}^\mu = x^\mu + \xi^\mu$, где $\xi^\mu = (\xi^0, -\boldsymbol{\xi})$ — малые функции, частота принимаемого фотона изменится следующим образом (ср. (11.2.5)):

$$\tilde{\omega} = \tilde{g}_{00}^{-1/2} \tilde{p}_0 = \omega (1 + \mathbf{e}\boldsymbol{\xi}').$$

Учитывая перенормировку фоновой температуры для разных наблюдателей, получаем изменение анизотропии:

$$\tilde{\delta}_T = \delta_T + H_0 \xi^0 + \mathbf{e}\boldsymbol{\xi}'. \quad (11.2.10)$$

Мы видим, что небо наблюдателей отличается лишь в монопольном ($l = 0$) и дипольном ($l = 1$) членах, тогда как все остальные мультиполи выглядят для них одинаково.

Таким образом, разные наблюдатели, находящиеся в одной точке пространства-времени, но движущиеся один относительно другого, увидят одну и ту же картину анизотропии РИ в квадрупольной и высших гармониках $l \geq 2$. Однако карты анизотропии для наблюдателей, находящихся в разных местах пространства-времени, будут различаться.

11.3. Угловая корреляционная функция

Для каждой точки пространства-времени существует своя карта РИ $\delta_T(\mathbf{e})$. Разложим ее по сферическим гармоникам Y_{lm} :

$$\delta_T(\mathbf{e}) = \sum_{l,m} a_{lm} Y_{lm}(\mathbf{e}), \quad a_{lm} = |a_{lm}| e^{i\psi_{lm}}, \quad (11.3.1)$$

где $l = 2, 3, \dots$ — орбитальное число, $m = -l, \dots, l$ — азимутальное число. Функция δ_T действительная, но коэффициенты разложения a_{lm} — комплексные числа (соответствующей симметрией обладают фазы ψ_{lm}).

Определим угловой спектр мощности анизотропии РИ как сумму квадратов модулей коэффициентов разложения (11.3.1) для данной гармоники l :

$$C_l = \sum_{m=-l}^l |a_{lm}|^2. \quad (11.3.2)$$

Величина

$$\Delta_l \equiv \frac{\delta T_l}{T} = \left(\frac{l(l+1) C_l}{2\pi} \right)^{1/2} \quad (11.3.3)$$

имеет смысл спектральной амплитуды анизотропии в угловом масштабе l (см. (11.1.4)): она задает парциальную анизотропию на гармонике l в интервале гармоник $\Delta l \sim l$.

Поскольку поле космологических возмущений подчиняется гауссовой статистике и связь q с δ_T линейна, то и коэффициенты разложения a_{lm} распределены по гауссову закону и их фазы ψ_{lm} — случайные числа в интервале $(0, 2\pi)$. В теории мы можем вычислить только *среднюю* вариацию $\langle C_l \rangle$, усредненную по всем реализациям гауссового поля или по ансамблю наблюдателей данной реализации. Однако конкретный наблюдатель во Вселенной, которая есть продукт *единственной* реализации гауссова поля, находит *свой* спектр мощности C_l , усредняя квадрат измеряемого сигнала (11.3.1) по своей небесной сфере. Очевидно, этот спектр отличается от среднего (теоретического). С этой ситуацией мы сталкиваемся в реальных наблюдениях.

Так как величина C_l собрана из $2l + 1$ некоррелированных чисел, то статистическое отклонение измеряемого спектра мощности от среднего составляет величину:

$$\frac{\Delta C_l}{C_l} \simeq (2l + 1)^{-1/2}. \quad (11.3.4)$$

Эта принципиальная неопределенность в сопоставлении теории с наблюдениями называется *космической вариацией*: она не зависит от качества и точности эксперимента и связана только с тем, что мы имеем дело с *одним* наблюдателем. Для низких гармоник величина космической вариации довольно большая ($\sim 20\%$ при $l = 2$) и медленно спадает с ростом l . К счастью, в области высоких гармоник влияние космической вариации невелико.

Для того чтобы связать спектр мощности с вариацией температуры, определим *угловую корреляционную функцию* — произведение амплитуд (11.3.1) в двух направлениях \mathbf{e}_1 и \mathbf{e}_2 , усредненное по всем направлениям на небе при фиксированном угле θ между \mathbf{e}_1 и \mathbf{e}_2 :

$$C(\theta) \equiv \langle \delta_T(\mathbf{e}_1) \cdot \delta_T(\mathbf{e}_2) \rangle = \frac{1}{4\pi} \sum (2l + 1) C_l P_l(\cos \theta), \quad (11.3.5)$$

где $\cos \theta = \mathbf{e}_1 \mathbf{e}_2$, $P_l(\cos \theta)$ — полиномы Лежандра, суммирование ведется по всем $l \geq 2$. Аналогично можно определить вариацию разности температур между двумя направлениями на небесной сфере:

$$\Delta_T^2 \equiv \langle |\delta_T(\mathbf{e}_1) - \delta_T(\mathbf{e}_2)|^2 \rangle = 2 [C(0) - C(\theta)]. \quad (11.3.6)$$

Для нахождения связи между угловыми и пространственными спектрами неоднородностей нам надо разобраться, каким образом происходит генерация *угловой* анизотропии (11.3.1) в *трехмерном* поле первичных космологических возмущений.

11.4. Карта реликтового излучения

Рассмотрим процесс генерации анизотропии РИ. Для этого нужно решить задачу о переносе излучения после его отрыва от вещества. Обратимся к характеристикам уравнения Больцмана, которые являются геодезическими траекториями фотонов.

Импульс свободного фотона представим в форме эйконала. Раскладывая его в ряд и оставляя лишь члены нулевого и первого порядка малости, получаем:

$$p_\mu = \mathcal{P}_{,\mu} = p(e_\mu + \epsilon_{,\mu}), \quad (11.4.1)$$

$$g_{\mu\nu} = a^2(\eta_{\mu\nu} + \bar{h}_{\mu\nu}), \quad (11.4.2)$$

где $e_\mu \equiv (1, \mathbf{e})$ — единичный вектор ($|\mathbf{e}| = 1$), ϵ и $\bar{h}_{\mu\nu}$ — малые функции координат (η, \mathbf{x}) , $\bar{h}_{\mu\alpha}\eta^{\alpha\nu} = h_{\mu\alpha}g^{\alpha\nu}$ (см. 4.5.1), $\mathcal{P} = p(2u + \epsilon)$ — эйконал. Световые координаты (u, v) определены следующим образом:

$$u = \frac{\eta + r - \eta_0}{2}, \quad v = \frac{\eta - r - \eta_0}{2}, \quad (11.4.3)$$

где радиус $r = |\mathbf{x} - \mathbf{x}_0|$ отсчитывается от мировой линии наблюдателя, находящегося в точке \mathbf{x}_0 . Введение коэффициента $1/2$ связано с тем, что на невозмущенной траектории $u = 0$ фотона, приходящего к наблюдателю в момент времени η_0 , аффинный параметр равен $v = \eta - \eta_0$. Это позволяет в качестве аффинного параметра на световом конусе прошлого вместо v использовать время η .

Поле скоростей вещества имеет вид (ср. (4.5.2)):

$$a u^\mu = (1 - D, \mathbf{v}), \quad (11.4.4)$$

где \mathbf{v} — пространственный вектор скорости барионов. Частота фотона в точке его приема наблюдателем, движущимся вместе с веществом, равна

$$\omega = p_\mu u^\mu = \frac{p}{a} (1 - D + \mathbf{e}\mathbf{v} + \epsilon'). \quad (11.4.5)$$

Частота фотона в точке последнего рассеяния имеет вид:

$$\omega_{\text{ls}} = \frac{p}{a_{\text{ls}}} \left(1 - \frac{1}{3}\delta - D + \mathbf{e}\mathbf{v} + \epsilon' \right)_{\text{ls}}, \quad (11.4.6)$$

где член $\delta/3$ возник при определении момента последнего рассеяния (11.2.2).

Прежде чем связать обе частоты и построить карту анизотропии фотонов, необходимо решить уравнение геодезических и восстановить функцию ϵ' . Как это часто случается, достаточно рассмотреть только первый интеграл уравнения геодезических:

$$p_\mu p^\mu = 0. \quad (11.4.7)$$

Подставляя сюда соотношения (11.4.1) и (11.4.2), получаем:

$$e^\mu \epsilon_{,\mu} = \epsilon' - \epsilon_{,r} = \epsilon_{,v} = \frac{1}{2} e^\mu e^\nu \bar{h}_{\mu\nu} \equiv \mathcal{F}, \quad (11.4.8)$$

откуда находим искомые функции с помощью интегрирования вдоль светового конуса прошлого:

$$\epsilon = \int \mathcal{F} dv, \quad (11.4.9)$$

$$\epsilon' = \int \mathcal{F}' dv = \frac{1}{2} e^\mu e^\nu \int \frac{\partial \bar{h}_{\mu\nu}}{\partial \eta} dv. \quad (11.4.10)$$

Из (11.2.8), (11.4.5) и (11.4.6) получаем карту температуры РИ в точке приема:

$$T = T_{\text{ls}} \frac{\omega}{\omega_{\text{ls}}} = T_0 \left(1 + \frac{1}{3} \delta_{\text{ls}} - D \Big|_{\text{ls}}^0 + \mathbf{e}\mathbf{v} \Big|_{\text{ls}}^0 + \frac{1}{2} \int_{\text{ls}}^0 \bar{h}'_{\mu\nu} e^\mu e^\nu dv \right).$$

Перепишем эту формулу для анизотропии РИ, опуская монополюсный и дипольный члены, так как они связаны с местоположением и скоростью наблюдателя (см. (11.2.10)):

$$\delta_T(\mathbf{e}) = \left(D + \frac{1}{3} \delta - \mathbf{e}\mathbf{v} \right)_{\text{ls}} + \frac{1}{2} \int_{\text{ls}}^0 \bar{h}'_{\mu\nu} e^\mu e^\nu dv. \quad (11.4.11)$$

Три члена в скобках связаны со скалярными неоднородностями на сфере последнего рассеяния и описывают, соответственно, гравитационное смещение фотонов, адиабатическое изменение температуры из-за вариации плотности плазмы и доплеровский сдвиг частоты из-за радиального движения электронов. Последний член в правой части описывает *интегральный эффект Сакса-Вольфа*: суммарное смещение частоты фотона при его распространении в переменных гравитационных полях S-, T- и V-мод возмущений. В статических гравитационных полях интегральный эффект отсутствует.

Формула (11.4.11) справедлива в любой системе отсчета (4.5.1). Хотя каждый из членов в правой части (11.4.11) калибровочно не инвариантен, суммарный сигнал $\delta_T(\mathbf{e})$, по определению, лоренц-инвариантен, поскольку он относится к экспериментально измеряемой величине. Например, в эйлеровых координатах формула (11.4.11) принимает следующий вид (ср. (5.1.12)):

$$\delta_T(\mathbf{e}) = \left(\Phi + \frac{1}{3} \delta_E - \mathbf{e} \mathbf{v}_b \right)_{ls} + \int_{ls}^0 (\Psi' + \Phi' + \mathfrak{S}) d\eta, \quad (11.4.12)$$

где функция

$$\mathfrak{S} \equiv e^i e^j \mathfrak{S}'_{ij} - a^{-2} (\mathbf{e} \nabla) (\mathbf{e} \mathfrak{S})$$

описывает влияние T- и V-мод возмущений на изменение частоты фотона при его распространении от события последнего рассеяния до регистрации¹⁾. Мода S дает два интегральных вклада в δ_T : первый (Ψ') описывает изменение времени, а второй (Φ') — скалярное искривление пространства на пути следования фотона.

Если физическое действие T- и V-мод на анизотропию РИ полностью описывается выражением (11.4.12), то для S-моды требуется пояснение. Пекулярные скорости движения компонент материи принято определять в эйлеровой системе отсчета, однако для плотности более уместна лагранжева формулировка.

Из (5.1.14) и (5.1.15) получаем следующие соотношения для пекулярной скорости \mathbf{v}_b и сопутствующего материи возмущения плотности барионов $\delta_{\epsilon b}$ на сфере последнего рассеяния (см. раздел 6.5):

$$\mathbf{v}_b = -\frac{\nabla}{H} \mathbf{q}_b, \quad (11.4.13)$$

$$\frac{1}{3} \delta_{\epsilon b} \equiv \frac{1}{3} \delta_b - c_M \bar{q} = \frac{1}{3} \delta_E + q, \quad (11.4.14)$$

$$\mathbf{q}_b \equiv \mathbf{q} + c_M \bar{\mathbf{q}}, \quad q \equiv q - \Phi, \quad \bar{q} \equiv H \int \frac{q'}{H} d\eta,$$

$$c_M \equiv \frac{f_M}{f_R} = \frac{3f_M y}{1 + 3f_b y} \simeq \frac{15}{11} \simeq 1,36,$$

где $y = y_{\text{гес}} \simeq 3/4$. Учитывая связь Ψ и Φ с q (см. (5.1.13), (5.2.6)), получаем стандартное представление анизотропии

¹⁾ V-мода влияет также на скорость электронов на сфере последнего рассеяния (доплеровский член в правой части (11.4.12)).

в приближении мгновенной рекомбинации:

$$\delta_T(\mathbf{e}) = \left(q - 2\mathfrak{q} + \frac{1}{3} \delta_{\epsilon b} + \frac{\mathbf{e} \nabla}{H} \mathfrak{q}_b \right)_{\text{гес}} + 2 \int_0^{1100} [(1 + \gamma) \mathfrak{q} - q] \frac{dz}{1+z} \quad (11.4.15)$$

(интеграл берется вдоль светового конуса прошлого, для краткости члены с T- и V-модами опускаем, и $S = 0$).

Все члены в этой формуле калибровочно-инвариантны и имеют ясный физический смысл. Первые два члена в круглых скобках правой части (11.4.15) описывают искривление гиперповерхности рекомбинации, третий член — это возмущение плотности звукового поля плазмы на сфере последнего рассеяния, а четвертый член — доплеровский. Все это осциллирующие по частоте функции с постоянной амплитудой начального S-спектра q_k и с фазами косинуса (сумма первых трех членов) и синуса (четвертый член) соответственно (см. раздел 6.5).

Интеграл описывает изменение частоты фотона при его распространении в гравитационном поле S-моды. После отсоединения излучения от вещества давление материи ничтожно и потенциал кривизны q статичен (см. раздел 6.6):

$$\delta_p \simeq S \simeq 0, \quad q \simeq q_0(\mathbf{x}), \quad \mathfrak{q} \simeq q \frac{H}{a} \int \frac{da}{H}.$$

Большую часть своего пути фотон проводит на материально-доминированной стадии, где $\gamma \simeq 3/2$ и гравитационный потенциал Φ также неизменен: $q \simeq 2q/5$. В этом случае интегральный эффект Сакса–Вольфа отсутствует. Фотон «ныряет» в гравитационные ямы структуры (или преодолевает гравитационные горы космологических пустот), испытывая голубое (красное) смещение, а потом, выходя из них, «краснеет» («синеет») так, что оба фактора смещения частоты в точности компенсируют друг друга. В результате никаких следов от прохождения фотонами неоднородных структур со статическим гравитационным потенциалом на карте РИ не остается¹⁾.

¹⁾ Это утверждение справедливо для пылевидной среды лишь в первом порядке теории возмущений. На нелинейной стадии развития структуры гравитационное поле зависит от времени и возникает эффект декомпенсации, однако вероятность проекции луча зрения на нелинейную область мала (см. раздел 2.8).

Таким образом, вклад в δ_T от интегрального члена в (11.4.15) естественным образом разбивается на две части в зависимости от влияния на $\Phi' \neq 0$ форм материи, обладающих давлением. Так, на начальном участке мировой линии свободного фотона вблизи рекомбинации незначительное влияние на Φ' оказывает остывающее излучение. Напротив, на позднем этапе $z < 1$ линейный гравитационный потенциал структуры начинает затухать под действием темной энергии, что приводит к небольшому (в пределах нескольких процентов) увеличению крупномасштабной анизотропии РИ. Кардинальное влияние темной энергии на динамику возмущений проявится в будущем, приведя к смене процесса разрушения хаббловского потока на его полное восстановление в зонах квазилинейной эволюции (см. главу 12).

11.5. Спектр анизотропии

Отрыв излучения от вещества в нашей Вселенной произошел в тот момент, когда начинала доминировать темная материя. Эра доминирования нерелятивистского вещества длилась около 10 млрд лет¹⁾ и сменилась эпохой доминирования темной энергии, динамическое время развития которой, как мы уже отмечали, составляет 17 млрд лет.

Таким образом, главным источником анизотропии РИ являются геометрические неоднородности в эпоху рекомбинации водорода $z \sim 1100$, а не поздняя структура Вселенной. Рассмотрим более детально спектральные свойства этой анизотропии.

Согласно (11.4.15), поле анизотропии на сфере последнего рассеяния представимо в виде интеграла Фурье:

$$\delta_T(\eta_0, \mathbf{x}, \mathbf{e}) = \int \delta_{\mathbf{k}}^{(T)} e^{i\mathbf{k}\mathbf{x}_{1s}} \frac{d\mathbf{k}}{(2\pi)^{3/2}}, \quad (11.5.1)$$

где амплитуда $\delta_{\mathbf{k}}^{(T)}$ определена в момент $\eta_{1s} = \eta_{\text{гес}}$, а связь между координатами задана на световом конусе прошлого:

$$\mathbf{x}_{1s} = \mathbf{x} + \mathbf{e}(\eta_0 - \eta_{1s}). \quad (11.5.2)$$

¹⁾ Масштабный фактор поздней Вселенной имеет вид:

$$a = \left(\frac{\Omega_m}{\Omega_E} \right)^{1/3} (\text{sh } \tau)^{2/3}, \quad \tau = \frac{3}{2} H_E t,$$

где $H_E = H_0 \sqrt{\Omega_E} \simeq 60 \text{ км}/(\text{с} \cdot \text{Мпк})$, $H_E^{-1} \simeq 17 \text{ млрд лет}$. В момент равенства плотностей t_* имеем: $\rho_m = \rho_E$, $\text{sh } \tau_* = 1$, $z_* \simeq 0,3$, $t_* \simeq 7/12 H_E^{-1} \simeq \simeq 10 \text{ млрд лет}$. В момент начала инфляции t_E имеем: $\gamma = 1$, $\text{sh } \tau_E = 1/\sqrt{2}$, $z_E \simeq 0,7$, $t_E \simeq 7 \text{ млрд лет}$.

Фурье-гармонику представим в виде оператора

$$\delta_{\mathbf{k}}^{(T)} = \delta_{\mathbf{k}}^{(0)} + \delta_{\mathbf{k}}^{(1)} \mathbf{e}\mathbf{n} = \delta_{\mathbf{k}}^{(0)} + \delta_{\mathbf{k}}^{(1)} \frac{\partial}{k \partial \eta_0}, \quad (11.5.3)$$

где произведение \mathbf{e} на единичный вектор

$$\mathbf{n} \equiv i \frac{\mathbf{k}}{k} = \frac{\partial}{k \partial \mathbf{x}_{1s}} \quad (11.5.4)$$

преобразуется с помощью (11.5.2) к частной производной по современному времени η_0 .

Компоненты $\delta_{\mathbf{k}}^{(0,1)}$ связаны с полем возмущений в момент рекомбинации (см. (11.4.15)):

$$\delta_{\mathbf{k}}^{(0)} \equiv C_0 q_{\mathbf{k}} = \left(q - 2q + \frac{1}{3} \delta_{\epsilon b} \right)_{\text{rec}} = \left(\Phi - q_b + \frac{1}{3} \delta_b \right)_{\text{rec}}, \quad (11.5.5)$$

$$\delta_{\mathbf{k}}^{(1)} \equiv C_1 q_{\mathbf{k}} = \left(\frac{k}{H} q_b \right)_{\text{rec}} = (c_1 \mathcal{K} q_b)_{\text{rec}},$$

где $c_1 = \sqrt{3} f_r \simeq 4\sqrt{3}/13 = 0,53$ (см. (6.5.18)), $q_{\mathbf{k}}$ — фурье-амплитуды поля q начальных возмущений. В силу линейности теории переходные функции $C_{0,1} = C_{0,1}(k)$, связывающие начальное поле q с возмущениями в эпоху рекомбинации, зависят только от волнового числа k .

После подстановки этих выражений в (11.3.5) получаем угловую корреляционную функцию в виде предела $\tilde{\eta}_0 \rightarrow \eta_0$ от следующего интегрального соотношения:

$$C(\theta) = \int q_k^2 \left(C_0 + C_1 \frac{\partial}{k \partial \eta_0} \right) \left(C_0 + C_1 \frac{\partial}{k \partial \tilde{\eta}_0} \right) \times \\ \times \frac{\sin(k |\mathbf{e}_1 \eta_0 - \mathbf{e}_2 \tilde{\eta}_0|)}{k |\mathbf{e}_1 \eta_0 - \mathbf{e}_2 \tilde{\eta}_0|} \frac{dk}{k}. \quad (11.5.6)$$

Учитывая разложение по функциям Бесселя и полиномам Лежандра:

$$\frac{\sin(k |\mathbf{e}_1 \eta_0 - \mathbf{e}_2 \tilde{\eta}_0|)}{k |\mathbf{e}_1 \eta_0 - \mathbf{e}_2 \tilde{\eta}_0|} = \sum_{l=0}^{\infty} (2l+1) j_l(k\eta_0) j_l(k\tilde{\eta}_0) P_l(\cos \theta), \quad (11.5.7)$$

и полагая после дифференцирования $\eta_0 = \tilde{\eta}_0$, получаем спектр мощности анизотропии РИ:

$$C_l = \int q_k^2 W_l^2(k\eta_0) \frac{dk}{k}, \quad (11.5.8)$$

где

$$W_l(x) = 2\sqrt{\pi} [C_0 j_l(x) + C_1 j_l'(x)]$$

— переходная функция, связывающая пространственный трехмерный S-спектр начальных возмущений и двумерный спектр анизотропии РИ на сфере последнего рассеяния (здесь штрих означает производную по аргументу).

11.6. Звуковая модуляция спектра мощности

Рассмотрим основные свойства спектра мощности C_l . Используя формулы раздела 6.5, получаем следующие асимптотики для коэффициентов $C_{0,1}$:

$$C_0 \simeq \begin{cases} \frac{1}{5} \left(1 - \frac{k^2}{\gamma H^2} \right) \simeq \frac{1}{5} \left(1 - \frac{3k^2}{k_{\text{eq}}^2} \right), & k < k_{\text{eq}}, \\ \cos \left(\frac{3k}{2k_{\text{eq}}} \right), & k > k_{\text{eq}}, \end{cases} \quad (11.6.1)$$

$$C_1 \simeq \begin{cases} \frac{3\sqrt{2}k}{5k_{\text{eq}}} \simeq \frac{4k}{5k_{\text{eq}}}, & k < k_{\text{eq}}, \\ \frac{1}{2} \sin \left(\frac{3k}{2k_{\text{eq}}} \right), & k > k_{\text{eq}}. \end{cases}$$

Отсюда видно, что крупномасштабную анизотропию РИ формирует начальный потенциал кривизны $q = q(\mathbf{x})$:

$$\delta_T(\eta_0, \mathbf{x}, \mathbf{e}) \simeq \frac{1}{5} q(\mathbf{x}_{\text{rec}}). \quad (11.6.2)$$

Предполагая в области низких частот ($k < k_{\text{eq}}$) плоский спектр Харрисона–Зельдовича:

$$q_k \simeq q_{\text{HZ}} = \text{const},$$

получаем после интегрирования (11.5.8) спектральные амплитуды (11.3.3) с $l \ll 100$:

$$\Delta_l \equiv \left(\frac{l(l+1)C_l}{2\pi} \right)^{1/2} \simeq \frac{1}{5} q_{\text{HZ}} = \text{const}. \quad (11.6.3)$$

Именно это свойство *прямого соответствия* двумерного спектра (11.6.3) крупномасштабно-инвариантных скалярных возмущений позиционному соотношению (11.6.2) температурной анизотропии предопределило выбор и нормировку спектральной амплитуды Δ_l (см. (11.3.3)).

Мелкомасштабная анизотропия определяется звуковыми вариациями плотности и скорости плазмы с временной фазой растущего решения на сфере последнего рассеяния. В отличие

от звуковых модуляций спектра мощности $P(k)$, подавленных малым фактором $f_b \simeq 1/6$ (см. (6.6.8)), в данном случае глубина модуляции функции Δ_l достигает *ста процентов*, поскольку излучение передает *только* распределение барионной плазмы, подверженной звуковым колебаниям.

Пики функции Δ_l соответствуют фазам косинуса πp и локализованы при

$$l_p \simeq \frac{\pi p}{\theta_{\text{sec}}} \simeq 200p, \quad (11.6.4)$$

где p — натуральное число. Согласно (11.6.1) амплитуда звуковой модуляции постоянна, однако в реальности это не так (рис. 10.9). Если первый пик ($p = 1$) в три раза выше крупномасштабной подложки (ср. (11.6.1)), то последующие ($p = 2, 3, \dots$) сильно подавлены эффектами, связанными с немгновенностью рекомбинации, которые необходимо учитывать при расчете мелкомасштабной анизотропии РИ ($l > 100$).

Отметим два основных фактора воздействия процесса рекомбинации на спектр мощности реликтового излучения:

- конечная длительность рекомбинации;
- диссипация звуковых колебаний.

В первом случае стирается информация о местоположении последнего рассеяния фотона и температура приходящего излучения эффективно усредняется по ширине рекомбинации. Во втором случае происходит физическая диссипация звуковых колебаний плазмы (силковское затухание) из-за конечной длины пробега фотонов в период рекомбинации.

Для оценки эффекта конечной толщины рекомбинации найдем вероятность прихода фотона к наблюдателю после его последнего рассеяния в момент времени t_{1s} .

Поскольку пространственная плотность свободных электронов на пути распространения фотона не равна нулю, то искомая вероятность $\mathcal{P}(t_{1s})$ является произведением вероятности последнего рассеяния фотона за малый промежуток времени δt вблизи t_{1s} :

$$\delta\tau_{1s} = \sigma_T n_e(t_{1s}) \delta t,$$

и вероятности его нерассеяния во все последующие времена $t_n = t_{1s} + n\delta t$ вплоть до момента регистрации:

$$\delta\mathcal{P} = \delta\tau_{1s} (1 - \delta\tau_1) (1 - \delta\tau_2) \dots (1 - \delta\tau_N),$$

где $n = 1, 2, \dots, N$, $\delta t = (t_0 - t_{1s})/N$, $\delta\tau_n = \sigma_T n_e(t_n) \delta t$, σ_T — томсоновское сечение рассеяния, n_e — плотность свободных

электронов. Переходя к пределу $N \rightarrow \infty$,

$$d\mathcal{P} = V d\eta_{ls}, \quad (11.6.5)$$

получаем функцию видности в параметрической форме:

$$V = V(\eta_{ls}) = -\tau' e^{-\tau}, \quad (11.6.6)$$

определяющую вероятность рассеяния фотона в момент времени η_{ls} в зависимости от оптической толщи

$$\tau = \tau(\eta_{ls}) = \int_{ls}^0 \sigma_T n_e a d\eta = \int_0^{ls} \frac{\sigma_T n_e dz}{(1+z)H}. \quad (11.6.7)$$

Функция видности имеет двугорбую форму с высоким максимумом при $z_{\text{rec}} \sim 10^3$ и небольшим подъемом в области *вторичной ионизации* водорода первыми квазарами и звездами во Вселенной ($z < z_{\text{ri}} \sim 10$). Область максимума можно аппроксимировать гауссианой

$$V(\eta_{ls}) \simeq \frac{k_r}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} k_r^2 (\eta_{ls} - \eta_{\text{rec}})^2\right) \quad (11.6.8)$$

с центром в точке η_{rec} и полушириной рекомбинации k_r^{-1} . Подставляя в (11.6.7) зависимость темпа рекомбинации от времени $n_e(\eta)$, получаем стандартные параметры рекомбинационной эпохи в ранней Вселенной:

$$\begin{aligned} k_r &\simeq \frac{20}{\eta_{\text{rec}}} \simeq 7 k_{\text{eq}}, & y_{\text{rec}} &\simeq \frac{3}{4}, \\ z_{\text{rec}} &\simeq 1100, & \left(\frac{n_e}{n}\right)_{\text{rec}} &\simeq 0,3. \end{aligned}$$

Для вычисления C_l с учетом ширины рекомбинации необходимо усреднить сигнал (11.5.1) по распределению (11.6.5)–(11.6.8), в результате чего в подынтегральное выражение попадает фурье-образ (11.6.8), также являющийся гауссианой:

$$\delta_T(\eta_0, \mathbf{x}, \mathbf{e}) = \int \delta_{\mathbf{k}}^{(T)} \exp\left(i\mathbf{k}\mathbf{x}_{\text{rec}} - \frac{k^2}{6k_r^2}\right) \frac{d\mathbf{k}}{(2\pi)^{3/2}}.$$

После соответствующих расчетов переходная функция $W_l(k\eta_0)$ в (11.5.8) приобретает мультипликативный демпфирующий фактор $\exp(-k^2/6k_r^2)$, который и обеспечивает затухание мелко-масштабной части спектра мощности при

$$l \gtrsim l_r = \sqrt{3} k_r \eta_0 \simeq \frac{17}{\theta_{\text{srec}}} \simeq 10^3.$$

Похожий результат получается и при расчете силковского затухания, хотя физика здесь иная — сглаживание мелкомасштабных звуковых колебаний плазмы из-за конечной длины пробега фотонов, сопоставимой с длиной волны возмущения. Характерный масштаб диссипации также оказывается порядка десятой доли звукового горизонта¹⁾, что соответствует мультиполю $l_S \simeq l_r$.

Таким образом, результирующий масштаб затухания фоновой анизотропии получается путем перемножения двух экспоненциальных демпфирующих факторов:

$$\exp\left(-\frac{k^2}{6}(k_r^{-2} + k_S^{-2})\right) \simeq \exp\left(-\frac{k^2}{3k_r^2}\right),$$

откуда следует мультипольный индекс затухания:

$$l_d \simeq l_r/\sqrt{2} \simeq 700.$$

Другим фактором, влияющим на амплитуду наблюдаемой фоновой анизотропии, являются свободные электроны, появляющиеся в поздней Вселенной вследствие процесса вторичной ионизации водорода в эпоху образования галактик ($z \lesssim 10$). С одной стороны, это приводит к дополнительному фактору затухания $\exp(-\tau_{\text{HI}})$ первичной фоновой анизотропии на сфере последнего рассеяния, а с другой — создает дополнительный источник анизотропии и поляризации ближнего фона в анизотропной Вселенной.

Имеющиеся наблюдательные данные по анизотропии и поляризации РИ позволяют определить величину томсоновской толщи реионизации $\tau_{\text{HI}} \simeq 0,08$. Она соответствует полной реионизации водорода во Вселенной при $z \sim 7$ (эффект Гана–Петерсона), что также согласуется с наблюдениями полного поглощения излучения от далеких квазаров в линиях L_α и L_β при $z > 6,1$. Эти наблюдения означают, что доля нейтрального водорода во Вселенной возрастает при увеличении красного смещения: если при относительно небольших z доля атомов водорода в межгалактическом газе порядка 10^{-5} – 10^{-4} , то при $z = 6$ она уже оказывается больше ($f_{\text{HI}} \geq 10^{-3}$), а при еще больших z ($z > 6$) доля

¹⁾ Это объясняется тем, что темп рекомбинации и диссипация звука связаны единым источником, ведущим, в конце концов, к полному просветлению Вселенной. Момент времени, когда Вселенная становится прозрачной для фотонов (длина свободного пробега равна хаббловскому радиусу H^{-1}), составляет величину $y_{\text{dec}} \simeq 0,9$, что всего на 20% превышает масштабный фактор рекомбинации $y_{\text{rec}} \simeq 3/4$ (см. раздел 6.5).

нейтрального водорода во Вселенной резко увеличивается. Соответственно и высокий ионизационный ультрафиолетовый фон, создаваемый излучением звезд и квазаров, начинает убывать при возрастании z : он падает приблизительно в 25 раз с ростом z от 3 до 6.

Наблюдения реионизации водорода хорошо согласуются с теоретическими предсказаниями о коллапсе первых систем при $z \sim 20-10$ и образовании первичных объектов-гало темной материи с массами $M \sim 10^6-10^7 M_\odot$ (подробнее см. раздел 12.7). Однако теория образования первых звезд (так называемых звезд третьего поколения) и черных дыр в таких и более массивных системах оставляет пока желать лучшего. С подобными примерами нечеткого понимания нелинейной эволюции барионной компоненты мы столкнемся в дальнейшем при рассмотрении геометрических тестов Вселенной (см. разделы 12.4, 12.8). К счастью, в линейных и квазилинейных системах, таких как РИ или крупномасштабная структура Вселенной, подобные проблемы не возникают и динамика барионов учитывается сравнительно просто (см., напр., раздел 6.6).

11.7. Точная космология

Подведем итог основным результатам исследования РИ.

- Реликтовое излучение — это наиболее точный инструмент исследования геометрии и состава Вселенной.
- Вид спектра анизотропии РИ доказывает, что в ранней Вселенной существовала предсказанная квантово-гравитационной теорией рождения начальных возмущений растущая адиабатическая мода возмущений плотности, послужившая затравкой для образования галактик и структуры Вселенной.
- Измерения поляризации РИ способны привести к обнаружению космологических гравитационных волн, также предсказанных квантово-гравитационной теорией рождения начальных возмущений, и тем самым достроить теорию ранней Вселенной, основанную на инфляционном Большом взрыве ¹⁾.

Спектр анизотропии РИ чрезвычайно чувствителен к космологическим параметрам и процессу вторичной ионизации. Важ-

¹⁾ Исследования поляризации РИ находятся пока в стадии начального развития и по этой причине не освещаются здесь детально.

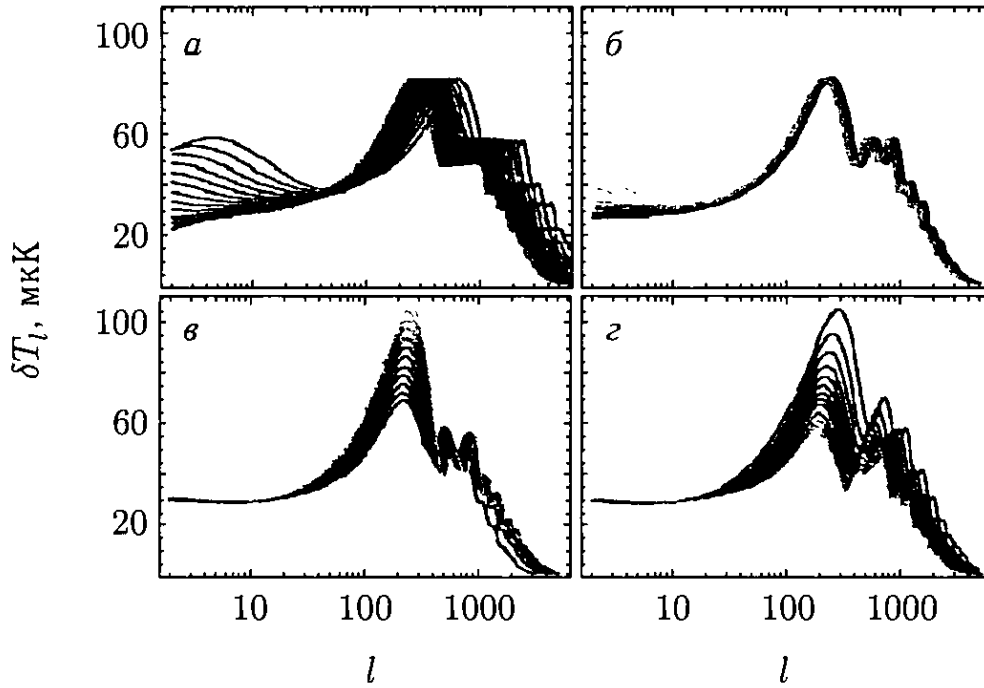


Рис. 11.2. Влияние космологических параметров на спектр анизотропии реликтового излучения $\delta T_l = T \Delta_l$ (Tegmark, <http://space.mit.edu/home/tegmark/cmb/>). Интервалы изменения параметров от синего к зеленому цветам: а) $0,1 < \Omega < 1$; б) $0 < \Omega_E < 0,9$; в) $0,005 < \Omega_b h^2 < 0,06$; г) $0,05 < \Omega_M h^2 < 0,5$.

нейшие пять параметров *первого уровня*, на которых базируется космологическая стандартная модель (КСМ), были надежно установлены с помощью РИ. Перечислим их в порядке возрастания точности определения (табл. 10.1): плоское пространство нулевого порядка, космологические плотности барионов и темной материи, плоский S-спектр начальных возмущений, оптическая толщина вторичной ионизации водорода во Вселенной объектами нелинейной структуры. Рисунок 11.2 иллюстрирует зависимость спектра анизотропии РИ (11.3.3) от отдельных параметров КСМ при их отклонении от соответствующих стандартных значений.

К сожалению, анизотропия РИ малочувствительна к постоянной Хаббла и плотности темной энергии; значения этих параметров проявляются сильнее в поздней Вселенной. Тем не менее, будущие динамические исследования интегрального эффекта Сакса–Вольфа помогут уточнить и эти параметры КСМ (см. подъем крупномасштабной анизотропии на рис. 11.2 а, б). Пока точность определения H_0 и Ω_E ограничивается внешними по отношению к РИ экспериментами и составляет $\sim 10\%$.

С РИ связаны большие надежды на измерение космологических параметров *второго уровня*: массы нейтрино, амплитуды гравитационных волн, показателей наклонов и изменений наклонов спектров мощности возмущений и пр. Эти параметры более

тонко воздействуют на анизотропию и поляризацию РИ, что требует более высокой точности измерений и обработки данных.

Важное значение имеет сочетание и сопоставление с радиоисследованиями РИ результатов других измерений: гравитационного линзирования, пекулярных скоростей материи, барионной звуковой модуляции спектра возмущений плотности, оптических и рентгеновских данных, геометрических методов и др. Следует помнить, что исследования, связанные с *нелинейными* системами (сверхновыми звездами, гамма-всплесками, галактиками, скоплениями галактик и пр.), подвержены систематическим ошибкам из-за сложного взаимодействия барионов с излучением, выявить которые весьма не просто. В то же время исследования РИ и крупномасштабной структуры Вселенной свободны от систематических неточностей, поскольку базируются на *линейной* теории возмущений.

ОБРАЗОВАНИЕ КРУПНОМАСШТАБНОЙ СТРУКТУРЫ ВСЕЛЕННОЙ

Крупномасштабной структурой Вселенной называют наблюдаемое *неоднородное* пространственное распределение материи, возникшее детерминированным образом из начальных малых геометрических возмущений скалярного типа (S-моды космологических возмущений). В современной Вселенной мы наблюдаем развитую нелинейную структуру в малых масштабах ($k > 0,1 \text{ Мпк}^{-1}$) в виде гало галактик, групп и скоплений галактик и более регулярное квазилинейное распределение материи в больших масштабах, вплоть до сотен Мпк — сверхскопления и космологические «пустоты». Переход между нелинейными и квазилинейными структурами многообразен.

Гравитационно-связанные гало распределены во Вселенной неравномерно. Часто они находятся в плоских образованиях — «стенках», поперечный размер которых не превышает 10 Мпк. В этом направлении стенка не расширяется и является нелинейным объектом, а в двух других, более протяженных, она квазилинейна и по ним расширение еще продолжается. Сами стенки неоднородны и состоят из нитеобразных «филаментов», которые могут пересекаться и формировать «узлы» — богатые скопления галактик. В среднем во Вселенной масштаб неоднородности составляет около 10 Мпк ($\sigma_{10} \simeq 1$), однако его значение зависит от местоположения, увеличиваясь в окрестностях скоплений и уменьшаясь вдали от них. Качественно этот масштаб разделяет области нелинейных гало и их распределение в пространстве.

Наблюдаемая структура характерна для развития гравитационной неустойчивости в нерелятивистской бесстолкновительной материи. Начальное квазиизотропное расширение среды в областях, куда вещество притекает, замедляется, останавливается и затем сменяется коллапсом. Начальная стадия коллапса идет преимущественно в одном из трех направлений, приводя к самопересечениям (каустикам) и образованию одномерных встречных потоков. Позднее области повышенной плотности сферизуются, формируя многопоточковые гравитационно-связанные системы — гало темной материи.

Квазилинейные структуры ($\delta_m < 1$) полностью описываются аналитически, поскольку для них имеется малый параметр, позволяющий использовать теорию возмущений. Для описания процессов нелинейного сгущивания имеются две аналитических аппроксимации — Зельдовича (точная в одномерном случае) и Пресса–Шехтера (точная в сферическом случае). Численное моделирование формирования структуры показало, что приближение Зельдовича хорошо описывает крупномасштабное распределение материи, где коллапс только начинается. Напротив, формализм Пресса–Шехтера применим к малым масштабам при описании гравитационно-связанных вириализованных гало.

В этой главе мы рассмотрим процессы генерации линейных и квазилинейных структур во Вселенной, а также образование и характеристики нелинейных гало темной материи. Особое внимание мы уделим темной энергии, поскольку она гравитационно воздействует на темп генерации структуры, и рассмотрим способы ее измерения.

12.1. Неоднородная Вселенная

Физической причиной возникновения нелинейной структуры — образования галактик из малых начальных возмущений кривизны — является гравитационная неустойчивость ТМ, наиболее интенсивно развивающаяся в период ее доминирования во Вселенной (см. разделы 6.5, 6.6).

Поскольку начальное давление у нерелятивистской ТМ отсутствует (относительные скорости частиц равны нулю), холодная материя свободно движется в собственном гравитационном поле квази Хаббловского потока. На линейной стадии эволюции малый контраст плотности ТМ нарастает со временем во всех масштабах с темпом, не зависящим от длины волны возмущения.

На этой стадии неоднородная Вселенная описывается обобщенным уравнением Фридмана (5.2.1). В отсутствие давления скаляр q и, следовательно, пространственная кривизна κ не зависят от времени (см. (5.2.3)), и уравнение сильно упрощается.

Умножая (5.2.1) на $(b/H_E)^2$ и переходя к безразмерным переменным, получаем:

$$\left(\frac{\dot{b}}{H_E}\right)^2 = f^2(b) - \hat{\kappa}(\mathbf{x}), \quad (12.1.1)$$

где

$$f^2(b) \equiv \frac{8\pi G}{3H_E^2} (\rho_m + \rho_E) b^2, \quad (12.1.2)$$

$$\hat{\kappa} = \hat{\kappa}(\mathbf{x}) \equiv \frac{\kappa}{H_E^2} = \Delta \hat{q}, \quad (12.1.3)$$

$$q_0 = q_0(\mathbf{x}) \equiv \frac{3}{2} H_E^2 \hat{q}, \quad (12.1.4)$$

$H_E = H_0 \sqrt{\Omega_E} \simeq (5 \text{ Гпк})^{-1}$ — хаббловская константа темной энергии, $b = b(t, \mathbf{x})$ — масштабный фактор объемного расширения материи, $x^\mu = (t, \mathbf{x})$ — ортогональные лагранжевы координаты (5.1.3) (знак сопутствия «с» опущен).

Рассмотрим сначала упрощенный случай, когда плотность темной энергии постоянна: $\rho_E = \text{const}$. Так как сопутствующая плотность материи $\rho_m \propto b^{-3}$, то функция

$$f(b) \equiv \left(\frac{c_m}{b} + b^2 \right)^{1/2} \gtrsim 1 \quad (12.1.5)$$

имеет минимум $f_{\min} \simeq 1$, достигаемый при $b_{\min}^{-1} \simeq 1,7$. Здесь $c_m \equiv \Omega_m / \Omega_E \simeq 0,4$ — постоянный коэффициент, шестая часть которого связана с барионами, а $5/6$ — с ТМ (в данном приближении обе компоненты движутся вместе).

Произвольная малая функция пространственных координат $\hat{\kappa} = H_E^{-2} \kappa$ описывает локальную нормированную кривизну пространства. Нас интересуют области пространства с положительной правой частью уравнения (12.1.1):

$$\hat{\kappa}(\mathbf{x}) < 1. \quad (12.1.6)$$

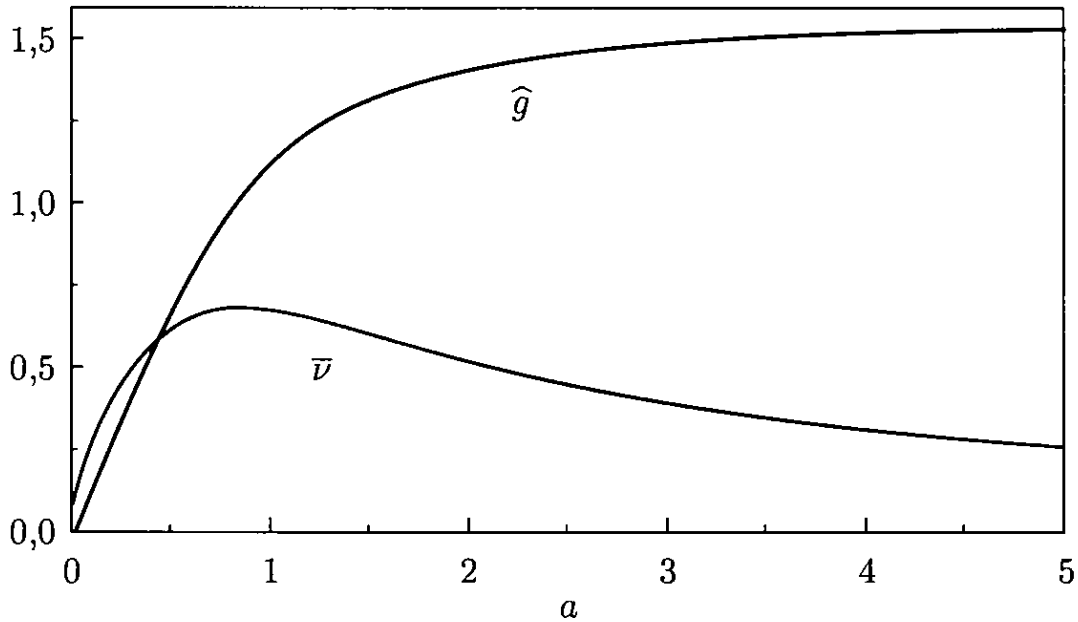
В этих областях плотность материи монотонно убывает со временем (подробнее см. (12.3.2)). Области (12.1.6) включают в себя как сверхскопления ($\kappa > 0$), так и космологические пустоты ($\kappa < 0$).

При $\kappa = 0$ объемный и фоновый масштабные факторы совпадают (при этом анизотропия расширения остается большой, см. (12.2.3)):

$$b = a(t) \equiv \frac{1}{1+z}, \quad H \equiv H_E \frac{f(a)}{a}, \quad (12.1.7)$$

где $f = f(a)$ — фактор роста хаббловской компоненты скорости:

$$\mathbf{V}_H = f(a) \cdot H_E \mathbf{x}. \quad (12.1.8)$$


 Рис. 12.1. Факторы роста плотности $\hat{g}(a)$ и пекулярной скорости $\bar{v}(a)$

В общем случае в линейном порядке по $\hat{\chi}$ получаем:

$$b = a \left(1 - \frac{1}{3} \hat{g} \hat{\chi} \right), \quad (12.1.9)$$

$$\delta_m = \hat{g} \hat{\chi}, \quad (12.1.10)$$

$$H_v \equiv \frac{\dot{b}}{b} = H \left(1 - \frac{1}{3} \bar{h} \hat{\chi} \right), \quad (12.1.11)$$

$$\bar{h} \equiv \frac{\bar{v}}{f} = \frac{\hat{g}}{H}, \quad (12.1.12)$$

где $\delta_m \equiv \delta\rho_m/\rho_m$ — сопутствующее возмущение плотности, $H_v = H_v(t, \mathbf{x})$ — объемная функция Хаббла, $\hat{g} = \hat{g}(a)$ и $\bar{v} = \bar{v}(a)$ — нормированные факторы роста возмущений плотности и пекулярной скорости (ср. (5.5.4), (12.2.5), см. рис. 12.1):

$$\hat{g}(a) \equiv \frac{1}{c_m} \left(a - H \int_0^a \frac{da}{H} \right), \quad (12.1.13)$$

$$\bar{v}(a) = \frac{\hat{g}'}{H_E} \equiv \frac{3H_E}{2a^2} \int_0^a \frac{da}{H}, \quad (12.1.14)$$

где штрих означает производную по времени η . При

$$a \ll 1: \quad \hat{g} \simeq \frac{3a}{2}, \quad \bar{v} \simeq \sqrt{a},$$

$$a \gg 1: \quad \hat{g} = \hat{g}_{\max} \simeq 1,56, \quad \bar{v} \simeq \frac{3}{2a},$$

$$a = 1: \quad \hat{g} \simeq 1,13, \quad \bar{v} \simeq 0,67.$$

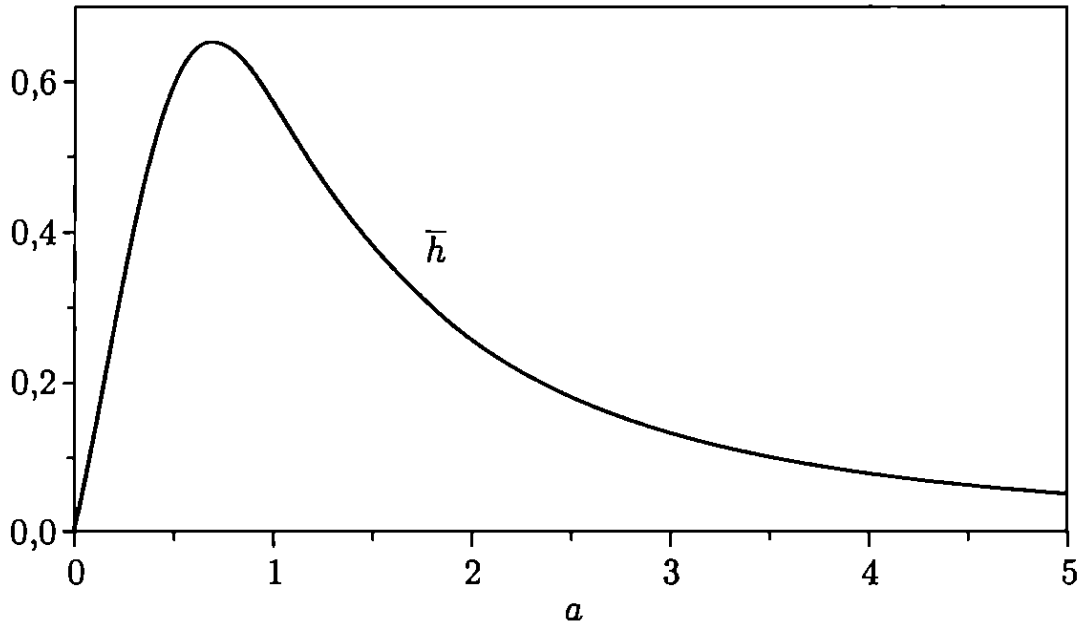


Рис. 12.2. Функция $\bar{h}(a)$, описывающая искажения хаббловского потока

Связь с введенными ранее факторами роста возмущений плотности (см. раздел 3.5) имеет вид:

$$\bar{g}(z) = \frac{g(a)}{g(1)} = \frac{\hat{g}(a)}{\hat{g}(1)}. \quad (12.1.15)$$

Уравнения (12.1.1)–(12.1.14) описывают квазихаббловские анизотропные течения с эффективной хаббловской функцией H_v , зависящей от местоположения наблюдателя. В современную эпоху функция \bar{v} имеет широкий максимум, свидетельствующий о периоде наиболее интенсивной генерации структуры. Положение ее максимума соответствует $z \simeq 0,2$; уровень 90% от максимального значения $\bar{v}_{\max} \simeq 0,68$ достигается при $a \simeq 0,5$ и 1,4, а уровень 50% — при $z \simeq 0,1$ и 4.

Таким образом, современная эпоха — это эпоха *максимальных* пекулярных скоростей, которая продлится еще космологическое время. Функция \bar{v} уменьшится вдвое, лишь когда возраст Вселенной достигнет 35 млрд лет. И только тогда можно будет говорить о наступлении эпохи затухания пекулярных скоростей во всех областях пространства, где $\hat{\kappa} < 1$.

Функция $\bar{h}(a)$, определяющая степень отклонения локальной хаббловской «постоянной» от ее фонового значения, представлена на рис. 12.2. Максимум $\bar{h}_{\max} \simeq 0,65$ достигается при $z \simeq 0,4$, а интервал значений $\bar{h} > \bar{h}_{\max}/2$ лежит в пределах $a \in (0,1, 1,8)$, что соответствует диапазону возраста Вселенной от 0,6 до 22 млрд лет.

Рисунок 12.2 наглядно демонстрирует, что наша Вселенная находится на этапе максимального искажения хаббловского

закона расширения:

$$\bar{h}(a = 1) \simeq 0,87 \cdot \bar{h}_{\max} \simeq 0,57,$$

а восстановление хаббловских течений в квазилинейных областях пространства произойдет лишь через десяток миллиардов лет.

Таким образом, мы можем заключить, что формирование крупномасштабной структуры Вселенной происходит в период от 1 до 20 млрд лет с момента Большого взрыва. Стадия подавления неоднородностей хаббловского потока из-за гравитационного влияния темной энергии еще не наступила, хотя темная энергия и доминирует в плотности материи уже 3,5 млрд лет. Эта задержка связана с тем, что динамическое влияние темной энергии на генерацию структуры Вселенной еще только началось, а характерное время этого воздействия больше современного возраста Вселенной и составляет величину $H_E^{-1} \simeq 17$ млрд лет.

Далее мы исследуем локальные течения вещества и анизотропию хаббловских потоков в нашей Вселенной.

12.2. Анизотропные холодные течения

Для описания поля пекулярных скоростей перейдем к эйлеровым координатам, в которых гравитационное поле локально изотропно в линейном порядке по \mathbf{x} в любой точке пространства (см. раздел 5.1). Следуя (12.1.1), упростим также обозначение ортогональных эйлеровых координат $y^\mu = (\tau, \mathbf{y})$, где $\tau \equiv t_E$, $\mathbf{y} \equiv y^i$ (см. (5.1.10)).

Интересующее нас соотношение между эйлеровыми y^μ и лагранжевыми x^μ координатами имеет вид (см. (5.1.11), (3.5.5)):

$$\tau = t - a\bar{v} H_E \hat{q}, \quad (12.2.1)$$

$$\mathbf{y} = \mathbf{x} + \hat{g} \hat{\mathbf{S}}, \quad (12.2.2)$$

где $\hat{\mathbf{S}} = \hat{\mathbf{S}}(\mathbf{x}) = \mathbf{S}/\hat{g}(1)$ — вектор смещения элемента среды относительно невозмущенного положения:

$$\hat{\mathbf{S}} = -\nabla \hat{q}, \quad \hat{\mathbf{x}} = -\text{div} \hat{\mathbf{S}} = \Delta \hat{q}.$$

Здесь лагранжева координата \mathbf{x} не меняется с течением времени вдоль траектории элемента среды и совпадает с эйлеровой координатой \mathbf{y} при $t \rightarrow 0$, потенциал $\hat{q} = \hat{q}(\mathbf{x})$ не зависит от собственного времени (см. (12.1.3)), фактор роста смещений $\hat{g} = \hat{g}(a)$ является одновременно и фактором роста возмущений плотности (12.1.10). Тот факт, что полное смещение материи

относительно ламинарного хаббловского потока факторизуется на произведение двух функций, одна из которых зависит только от времени, а вторая — от координат, говорит о том, что темп роста возмущений одинаков для всех длин волн.

Смещение элементов среды от невозмущенных хаббловских положений монотонно увеличивается со временем и сегодня составляет величину $\sigma_S \simeq 14$ Мпк (см. раздел 3.5). В будущем, если плотность темной энергии не изменится, среднее полное смещение приблизится к своему предельному значению — порядка 22 Мпк (рис. 12.1).

Из (12.2.1) получаем следующие представления интервала в эйлеровой и лагранжевой сетках (ср. (5.1.12) и (5.1.4)):

$$ds^2 = (1 + 2\Phi) d\tau^2 - a^2 dy^2 = dt^2 - a^2 (\delta_{ij} - 2\hat{g} \hat{q}_{,ij}) dx^i dx^j, \quad (12.2.3)$$

где

$$a \equiv a(t) \cdot (1 - q_0) = a(\tau) \cdot (1 - \Phi)$$

— скалярный инвариантный масштабный фактор (см. (5.1.1)), функция $b(t, \mathbf{x})$ пропорциональна следу пространственной части лагранжева метрического тензора (ср. (12.2.3) и (12.1.9)), а гравитационный потенциал возмущений плотности равен

$$\Phi = \frac{3}{5} \bar{\phi} q_0, \quad (12.2.4)$$

$$\bar{\phi} = \frac{5}{3} \phi(a) = \frac{5}{3} \left(1 - \frac{H}{a} \int_0^a \frac{da}{H} \right).$$

При $a \ll 1$ $\bar{\phi} = 1$ (ср. (5.5.3)). Временной фактор $\bar{\phi} = \bar{\phi}(a)$ распада гравитационного потенциала из-за влияния темной энергии изображен на рис. 12.3. Значение $\bar{\phi}$ может служить мерой динамического воздействия темной энергии на генерацию структуры. В современную эпоху при $a = 1$ $\bar{\phi} \simeq 0,77$.

Интервал (12.2.3) позволяет определить шкалу расстояний от наблюдателя до точек \mathbf{x} на гиперповерхности $t = \text{const}$: $\mathbf{r} = a(t) \cdot \mathbf{y}$. Очевидно, $\mathbf{r}(t, \mathbf{x}) = \mathbf{r}(z, \mathbf{x})$, см. (3.5.5). Дифференцируя \mathbf{r} по собственному времени t , получаем следующее выражение для пекулярной скорости материи:

$$\mathbf{v}_{\text{pec}} \equiv \dot{\mathbf{r}} - H\mathbf{r} = \hat{g}' \hat{\mathbf{S}} = \bar{v} H_E \hat{\mathbf{S}}. \quad (12.2.5)$$

Выражение (12.2.5) совпадает с определением 3-скорости как пространственной компоненты 4-скорости материи в эйлеровой

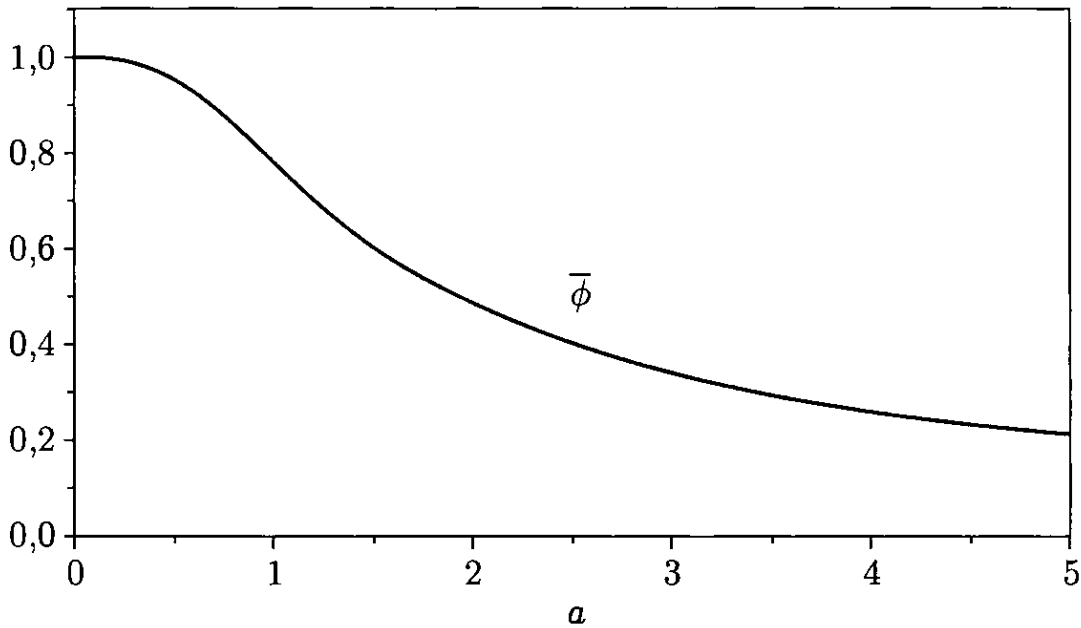


Рис. 12.3. Распад гравитационного потенциала $\bar{\phi}(a)$ крупномасштабной структуры Вселенной под гравитационным воздействием темной энергии

системе отсчета (см. (12.2.1), (5.1.15)):

$$\mathbf{v}_{\text{pec}} = -\frac{\partial t}{a \partial \mathbf{y}} = -\bar{\nu} H_E \nabla \hat{q} = -\frac{\nabla q}{H}.$$

Таким образом, фигурирующая в (12.1.11) величина $\bar{\nu}$ действительно является фактором роста пекулярной скорости.

Отметим, что согласно (12.2.5) полная скорость материи

$$\mathbf{V} = H \mathbf{r} + \mathbf{v}_{\text{pec}}.$$

Первая компонента $H \mathbf{r}$ описывает хаббловскую скорость *неоднородной* Вселенной. Ее тоже можно разбить на две части: одна из них — $\mathbf{V}_H = \bar{H} \mathbf{x}$ (см. (12.1.8)) — связана с однородной компонентой плотности, а вторая — $\hat{g} \bar{H} \hat{\mathbf{S}}$ — с возмущенной. Фактор роста второй компоненты скорости пропорционален фактору роста возмущений плотности \hat{g} , тогда как пекулярная скорость ведет себя $\propto \hat{g}'$.

Обратимся к локальным потокам материи.

В областях (12.1.6) неоднородной Вселенной течения среды описываются тензорным полем $H_{ij} = H_{ij}(t, \mathbf{x})$, обобщающим функцию $H(t)$ для фридмановской модели. Действительно, в фиксированный момент времени t получаем из (12.2.2) координатное расстояние между близкими точками среды в виде:

$$\delta y_i = (\delta_{ij} - \hat{g} \hat{q}_{,ij}) \delta x^j. \quad (12.2.6)$$

Дифференцируя физическое собственное расстояние $\delta r_i = a \delta y_i$ по времени, получаем поле парных скоростей:

$$\delta V_i \equiv \frac{\partial}{\partial t}(\delta r_i) = H_{ij} \delta r^j, \quad (12.2.7)$$

$$H_{ij} = H \delta_{ij} - \hat{g} \hat{q}_{,ij} = H (\delta_{ij} - \bar{h} \hat{q}_{,ij}).$$

След тензорного поля представляет собой объемную функцию Хаббла $H_v = H_{ii}/3$ (см. (12.1.11)), однако сам тензор H_{ij} сильно анизотропен. Анизотропия локального расширения (вариация проекций H_{ij} на радиальные направления от данной точки \mathbf{x}) имеет тот же порядок величины, что и отклонения H_v от среднего значения параметра Хаббла H . На границе квазилинейных областей (12.1.6) эти вариации достигают 100% (вплоть до остановки расширения в некоторых направлениях). Например, в окрестности Местной группы галактик на расстоянии более 2 Мпк от ее барицентра главные значения хаббловского тензора $H_{ij} = \text{diag}(48, 62, 81) \text{ км} \cdot \text{с}^{-1} \cdot \text{Мпк}^{-1}$ (рис. 12.4).

Поле H_{ij} описывает регулярные холодные течения вещества. Следует заметить, что формула (12.2.7) справедлива в приближении малых расстояний между галактиками, не превышающих корреляционный масштаб двухточечной корреляционной функции вектора смещений. Для разных проекций этого вектора относительно направления $\delta \mathbf{u}$ корреляционный радиус варьируется от 15 до 50 Мпк (рис. 3.13). С увеличением расстояния вдоль любого направления случайные отклонения от среднего потока (12.2.7) возрастают. Так устроен спектр космологических возмущений скорости: он падает в сторону коротких длин волн при $k > 10^{-2} \text{ Мпк}^{-1}$ (рис. 3.12). Поэтому вариация (12.2.7) в этой области *растет* с увеличением масштаба.

Напомним, что на расстоянии $\sim 2,5$ Мпк от барицентра Местной группы отклонения от зависимости (12.2.7) составляют величину $\sim 40 \text{ км} \cdot \text{с}^{-1}$, то есть около 25% от средней скорости (рис. 12.4). В то же время полная пекулярная скорость Местной группы относительно реликтового излучения равна $600 \text{ км} \cdot \text{с}^{-1}$. Основные масштабы неоднородностей, ответственные за столь высокую скорость, находятся в интервале от 10 до 70 Мпк.

Мы видим, что стандартная теория образования структуры Вселенной полностью объясняет наблюдаемые движения материи в квазилинейных областях Вселенной ($\hat{\kappa} < 1$). Локальные течения регулярны, гладки и сильно коррелированы. Малость случайных отклонений скоростей галактик от средних скоростей холодных потоков объясняется формой начального спектра про-

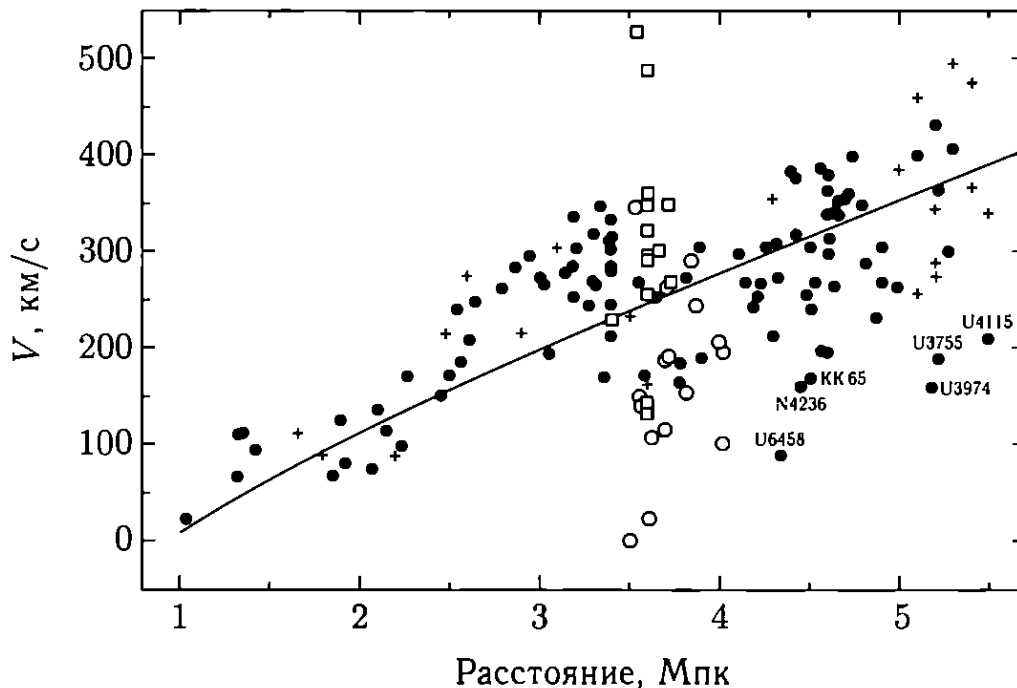


Рис. 12.4. Радиальные скорости галактик $V(r)$ в окрестности Местной группы по данным [Karachentsev et al., 2002]. Значками изображены разные типы галактик, расстояния до галактик даны от барицентра Местной группы

странственных возмущений плотности. На малых расстояниях они носят квазихаббловский характер, сохраняя главные черты хаббловского потока: течения холодные, радиальные и скорость разбегания галактик пропорциональна расстоянию. Однако «постоянная» Хаббла зависит от положения наблюдателя и направления в пространстве. Мы уже приводили в качестве примера окрестность Местной группы. Отношения главных значений хаббловского тензора H_{ij} на расстоянии несколько Мпк от ее барицентра составляют $\sim 3 : 4 : 5$.

С увеличением радиуса отклонения скоростей галактик от средних скоростей коррелированных потоков нарастают, достигая насыщения, начиная с расстояний ~ 10 Мпк, что соответствует минимальному корреляционному радиусу, где отклонения скоростей порядка самой хаббловской скорости. Далее отклонения не растут, тогда как хаббловские скорости продолжают нарастать. На расстояниях, много больших корреляционных размеров, течения вещества приближаются к идеальному закону Хаббла.

12.3. Как измерить темную энергию

Как было показано ранее, темная энергия, сама не являясь составляющей структуры, кардинально влияет на темп ее генерации во Вселенной и историю образования галактик.

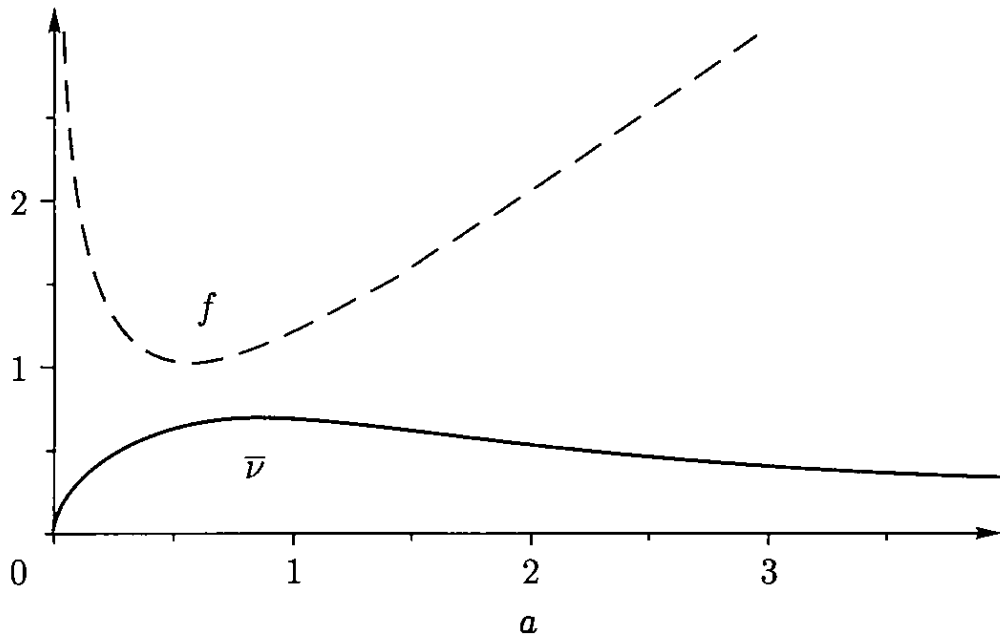


Рис. 12.5. Факторы роста хаббловской $f(a)$ (штриховая линия) и пекулярной $\bar{v}(a)$ (сплошная линия) скоростей материи; современная эпоха соответствует $a = 1$

Рисунок 12.5 иллюстрирует поведение факторов роста хаббловской и пекулярной составляющих скорости материи на квазилинейной стадии эволюции. Со временем расширение областей с $\hat{\kappa} > 0$ может перейти в коллапс, который начинается в отдельных направлениях и далее развивается в сторону формирования гравитационно-обособленных систем гало разных масс. Спектр начальных возмущений устроен так, что все большие массы имеют возможность коллапсировать в более поздние моменты времени (рис. 6.2).

Если бы ТЭ не было, то хаббловская скорость продолжала бы уменьшаться со временем, а пекулярная — нарастать. В результате в любой области пространства, где хотя бы одно из трех главных значений тензора $\hat{q}_{ij} = \text{diag}(\lambda_1, \lambda_2, \lambda_3)$ положительно ($\lambda_1 > 0$, где $\lambda_1 \geq \lambda_2 \geq \lambda_3$, см. (12.2.3)), рано или поздно наступил бы момент, когда отрицательная пекулярная скорость в этом направлении сравнялась бы с хаббловской скоростью и произошла остановка расширения: обращение в нуль главного значения H_{11} (см. (12.2.7)). По этому направлению происходит гравитационное обособление материи и последующий коллапс.

Однако динамическое воздействие ТЭ приводит к тому, что квазилинейные области не успевают эволюционно переходить в стадию коллапса: хаббловская скорость достигает минимума при $z \simeq 0,7$ и затем резко возрастает $\propto a$ (см. (12.1.8)), тогда как пекулярная скорость достигает насыщения (рис. 12.5).

Таким образом, нелинейная структура успевает образовываться в тех пространственных зонах, где нарушено условие (12.1.6), то есть

$$\hat{\kappa} = \lambda_1 + \lambda_2 + \lambda_3 > 1. \quad (12.3.1)$$

Несмотря на пороговый характер этого неравенства, оно не означает, что области, занятые нелинейными структурами, топологически компактны. Действительно, λ_3 (или $\lambda_{2,3}$) может быть отрицательным, и тогда в этом направлении (или направлениях) будет продолжаться расширение.

Как описать переход от квазилинейной стадии в фазу коллапса? Ответ на удивление прост: решение (12.2.2), определяющее эйлеровы координаты потока u в виде функций x и t , можно аналитически продолжить вплоть до первого самопересечения ($\delta u_1 = 0$) и даже далее путем «подшивания» через каустики возникающих многопоточковых течений (аппроксимация Зельдовича).

Качественно эта аналитическая аппроксимация приводит к правильным выводам, а количественные отклонения от реальной эволюции (до первого самопересечения) не превышают 20–30% по разным критериям. Успех столь простого метода связан с тем, что в космологических условиях потенциал q остается малым (образование черных дыр экспоненциально подавлено). Кроме того, решение (12.2.1), (12.2.2) оказывается точным в нелинейном *одномерном* случае, а именно этот вариант начального развития коллапса наиболее типичен во Вселенной.

Таким образом, вычисляя количественные характеристики вновь образующейся нелинейной структуры — функции распределения космологических пустот, свехскоплений, филаментов и пр., средние расстояния между стенками, узлами и пр., корреляционные функции, эволюцию по красному смещению и пр., — можно исходить из линейных возмущений и их гауссовой статистики, поскольку решение (12.2.2) естественно задается потенциальным полем малых начальных неоднородностей плотности. Приведем два примера.

Прежде всего, мы теперь можем более точно обозначить области, где никогда не образуются каустики, то есть идет вечное расширение:

$$\lambda_1 < \hat{g}_{\max}^{-1} \simeq 0,6. \quad (12.3.2)$$

До тех пор, пока возмущения плотности малы, оба условия (12.1.6) и (12.3.2) равнозначны. Однако условие (12.1.6) вовсе не гарантирует отсутствие коллапса: если $\lambda_1 > 0,6$, то расширение в этом направлении сменится на сжатие, пусть даже и в будущем.

Похожим образом мы можем уточнить поле плотности. Из (12.2.2) получаем сопутствующую плотность материи:

$$\rho_m(\eta, \mathbf{x}) = \frac{c_m \rho_E}{a^3 \det(\delta y_i / \delta x_j)} = \frac{c_m \rho_E}{a^3 (1 - \hat{g} \lambda_1) (1 - \hat{g} \lambda_2) (1 - \hat{g} \lambda_3)}.$$

Очевидно, при $\bar{g} \ll 1$ отсюда следует формула (12.1.10).

Мы вернемся к нелинейным структурам в последующих разделах, а сейчас обратимся к вопросу о том, как можно измерить ТЭ.

Очевидно, это можно сделать только с помощью методов наблюдательной космологии. Обнаружить ТЭ в лаборатории (наподобие, например, попыток лабораторного детектирования темной материи) не представляется возможным, поскольку эта всепроникающая субстанция (поле) слабо взаимодействовала с чем бы то ни было на протяжении всей известной нам истории мира, будь то недра звезд, черные дыры, ранняя Вселенная и др. Здесь возможны только не прямые методы измерений. К счастью, они уже доказали свою эффективность и привели к впечатляющим результатам (открытие самой ТЭ).

Что значит измерить темную энергию?

Ответ на этот вопрос подсказывает наблюдательная космология. Имеющееся ограничение основного параметра ТЭ

$$w = \frac{p_E}{\epsilon_E} : \quad |1 + w| < 0,1 \quad (12.3.3)$$

указывает на постоянство функции $w(a, \mathbf{x}) \approx \text{const}$ как во времени, так и в пространстве (см. (5.2.2)). Условие (12.3.3) несколько упрощает постановку задачи по обнаружению ТЭ, поскольку эффекты ее сгущивания ничтожны и пока недоступны измерениям на имеющемся уровне развития наблюдательной техники.

Таким образом, сегодня мы можем обсуждать реальные эксперименты только по временному измерению функции $w(a)$. Учитывая медленную эволюцию, эту функцию можно заменить набором постоянных коэффициентов c_n ($n = 1, 2, \dots$) ряда Тейлора при разложении вблизи значения -1 (см. раздел 10.7):

$$w = w(a) = -1 + c_0 + c_1 \Delta a + c_2 \frac{(\Delta a)^2}{2} + \dots, \quad (12.3.4)$$

где $\Delta a = 1 - a \equiv z(1 + z)^{-1}$. На текущем уровне знаний доступная нам физическая информация о ТЭ заключена в этих коэффициентах. Из (12.3.3) получаем:

$$|c_n| < 0,1.$$

Конечно, в первую очередь на повестке дня стоит вопрос об измерении c_0 и первых коэффициентов ряда. Теоретические предсказания, основанные на связях c_n с физическими параметрами моделей (φ_0 , m и др.), никак не дискриминируют их область значений, поэтому вопрос упирается исключительно в возможность экспериментального определения коэффициентов c_n .

Выделим три главных способа измерения функции $w(a)$:

- структурный;
- динамический;
- геометрический.

Рассмотрим коротко каждый из этих способов.

Структурный способ мы уже обсуждали ранее: ТЭ кардинально влияет на темп генерации галактик во Вселенной. Рисунки 12.1, 12.2, 12.3 и 12.5 демонстрируют кривые роста эволюционирующей структуры для случая, когда все $c_n = 0$. При отличных от нуля значениях c_n эволюционные кривые будут деформироваться, что легко рассчитать теоретически с помощью ОТО. Речь, таким образом, идет о максимально точных наблюдательных измерениях истинных функций роста и об определении с их помощью параметров c_n .

Каким образом можно точно измерить кривые роста?

К сожалению, большой помощи от традиционных астрономических методов наблюдений компактных источников в разных длинах волн (звезд, квазаров, галактик, скоплений галактик и др.) и построений по этим данным количественных характеристик распределения неоднородной материи в пространстве и во времени, ожидать не приходится.

Вспомним, что именно структурный аргумент привел к открытию феномена ТЭ. Вся астрономия XX века обосновала малую долю материи, входящей в крупномасштабную структуру Вселенной, и, следовательно, большую часть массы в форме неструктурированной ТЭ (см. разделы 10.6, 10.7). Однако дальнейшие исследования в этом направлении пока малоперспективны с точки зрения точности. Причина этого — неконтролируемые *нелинейные* эффекты взаимосвязи распределений плотности материи и излучения, которые, в конечном счете, делают невозможным восстановление с требуемой точностью характеристик полной массы материи только по светимости ее барионной компоненты.

Реального прорыва в структурном методе детектирования свойств ТЭ можно ожидать от динамических измерений, чувствительных к градиентам общего гравитационного потенциала. Примером могут служить статистические измерения слабого

гравитационного линзирования галактик фона близлежащими структурами, однако и тут имеются подводные камни. Дело в том, что восстанавливаемая поверхностная плотность полной массы (при заданном красном смещении), являющаяся конечным продуктом данного метода, еще недостаточна для успешного сравнения с теорией, поскольку мы не в состоянии пока рассчитать с нужной точностью вклад барионов в *локальную* плотность материи из-за их сложного взаимодействия с излучением (ударные волны, звездообразование, сверхновые, диссипация, ультрафиолетовый фон, ионизация, остывание и пр.).

Говоря сегодня о предельно точных методах измерений, необходимых для определения физических свойств ТЭ, речь в первую очередь идет об исследованиях *линейных* и *квазилинейных* структур, поскольку здесь имеется хорошо понимаемая, работающая и до конца контролируемая теория, базирующаяся на малых возмущениях. Примером прорыва в этой области являются измерения анизотропии и поляризации реликтового излучения, приведшие к созданию космологической стандартной модели. Такой же прорыв возможен и на пути ее дальнейшего расширения, основанном на точных измерениях космологических параметров, в том числе и параметров ТЭ.

Этот прорыв в состоянии обеспечить *любые* статистические измерения *крупномасштабных* пекулярных скоростей материи, неоднородностей гравитационного потенциала и распределений структур (в том числе с учетом барионной звуковой модуляции спектра) как функций красного смещения. В первом случае речь идет о доплеровских эффектах, связанных с движением материи, например об измерении разностей красных смещений пар галактик или квазаров как функции их поперечного углового расстояния. Во втором случае мы имеем в виду измерения поверхностной деформации гравитационного потенциала, например с помощью эффекта слабого линзирования, но уже в больших масштабах, где нет самопересечений потоков вещества и вклад барионов легко учесть. В третьем случае мы говорим о количественных характеристиках пространственного распределения структур в зависимости от красного смещения.

На этом пути мы приходим к следующему способу измерения ТЭ, который выделен нами в отдельный класс и назван динамическим (хотя и структурный метод во многом свелся к динамическим измерениям). Речь идет об интегральном эффекте Сакса–Вольфа.

Как мы уже отмечали ранее, в отсутствие ТЭ единственным источником фоновой анизотропии реликтового излучения явля-

ются космологические возмущения в ранней Вселенной в эпоху рекомбинации водорода (см. главу 11). ТЭ индуцирует *дополнительный* источник анизотропии, связанный с распадом гравитационного поля в современную эпоху (рис. 12.3). По этому источнику — интегральному эффекту Сакса–Вольфа (см. разделы 11.4, 11.5) — можно таким же *точным* образом определять свойства ТЭ, как это уже произошло с определением космологических параметров ранней Вселенной по измерениям анизотропии реликтового излучения в эпоху рекомбинации. Количеством речь идет об увеличении точности измерений всего лишь на порядок величины. Это и есть реализация целей точной космологии, вполне возможная и ясная, когда мы имеем дело с малыми возмущениями.

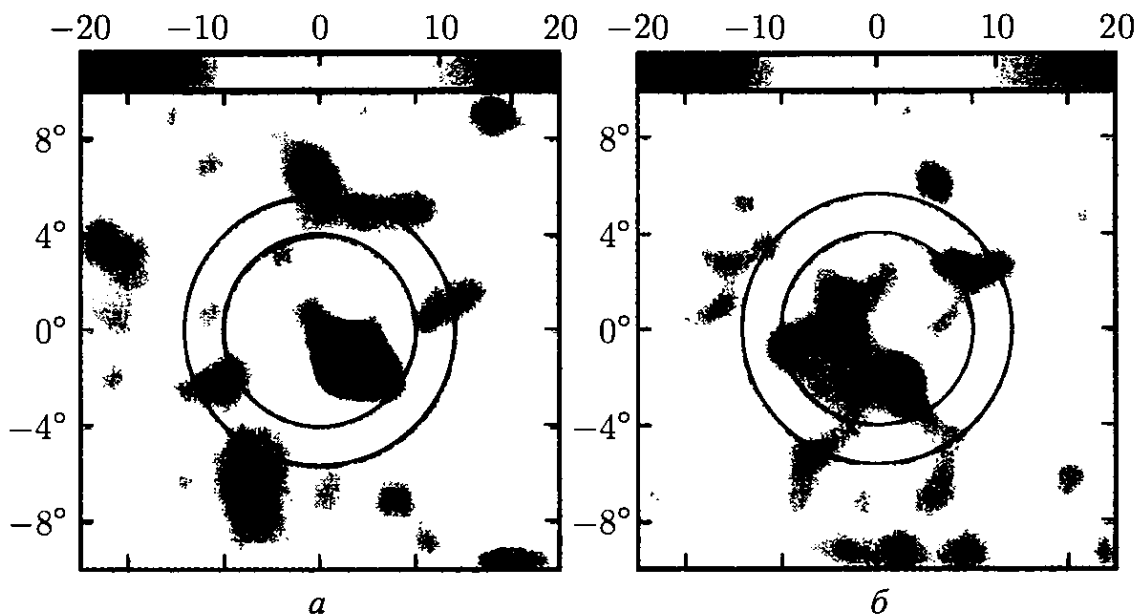


Рис. 12.6. Наложенные друг на друга участки карты реликтового излучения, проецирующиеся на 50 войдов (а) и 50 сверхскоплений (б) размерами ~ 100 Мпк (4°) по данным [Granett et al., 2008]. На верхней панели — шкала анизотропии реликтового излучения, выраженная в мК

Динамический интегральный эффект предсказывает наличие кросс-корреляции между двумя картами неба — анизотропии реликтового излучения (рис. 10.5) и крупномасштабной структуры Вселенной (напр., рис. 10.6). Этот эффект уже измерен и качественно соответствует предсказаниям теории.

На рис. 12.6 представлены наложенные друг на друга участки карт реликтового излучения, проецирующиеся на 50 войдов (рис. 12.6 а) и 50 сверхскоплений (рис. 12.6 б), выделенных по распределению галактик из каталога SDSS. Средний размер войдов/сверхскоплений составляет около 100 Мпк. Суммирование соответствующих областей с пониженной и повы-

шенной плотностями сделано для усиления искомого кросс-корреляционного сигнала на фоне случайного поля первичных возмущений. Мы наблюдаем явный провал температуры реликтового излучения (холодное пятно) в направлениях на войды и горячее пятно соответствующего углового размера в направлениях на сверхскопления, в полном соответствии с предсказаниями теории.

Очевидно, это только первые результаты, дающие соответствие с тем, что мы уже хорошо знаем по структурным исследованиям. Здесь, однако, закодирована *точная* космологическая информация, которая сможет быть расшифрована в будущих наблюдениях для построения расширенной (более точной и с возможными дополнительными параметрами) модели Вселенной, основанной на ОТО. Для реализации этой возможности необходимо иметь две точные карты линейных космологических возмущений: анизотропии реликтового излучения и гравитационного потенциала, восстановленного из наблюдений крупномасштабной структуры Вселенной на разных красных смещениях. Все это возможно уже при имеющемся уровне технических разработок и требует реализации — создания телескопов нового поколения как наземного, так и космического базирования.

12.4. Как измерить Вселенную

Впечатляющие успехи в исследовании Вселенной связаны с разностными и динамическими измерениями градиентов гравитационного потенциала, пекулярных скоростей и возмущений плотности методами наблюдательной космологии. Есть, однако, и прямые астрономические способы измерения геометрии нулевого порядка, минуя структуру. Их для краткости мы будем называть *геометрическими*. Эти способы имеют преимущество при измерениях на больших масштабах, поскольку с увеличением расстояний хаббловские потоки нарастают, а отклонения от них уменьшаются.

Геометрические тесты имеют дело с измерениями расстояний и времен между событиями. В нулевом порядке эти интервалы регулируются функцией $a(t)$ и ее производными. Тем самым, по геометрическим протяженностям можно определить состав Вселенной. По измеряемым космологическим функциям $H(z)$ и $\gamma(z)$ могут быть восстановлены не только свойства ТЭ, но и другие параметры материи.

Мы уже сталкивались с геометрическими тестами, например при подсчете числа галактик внутри сфер заданного радиуса

(рис. 10.2). Напомним два основных классических способа определения геометрических размеров:

- измерение радиальных расстояний как функции z ;
- измерение угловых масштабов как функции z .

В первом случае речь идет о зависимости *видимая звездная величина–красное смещение*:

$$m = m(z), \quad (12.4.1)$$

а во втором — *угловой размер–красное смещение*:

$$\theta = \theta(z). \quad (12.4.2)$$

Соотношение (12.4.1) называется в астрономии *хаббловской диаграммой*. Она позволяет определить изменение масштабного фактора Вселенной $a(t)$, если известна абсолютная светимость объектов, по которым проводят анализ. Действительно, $(a^{-1}-1)$ — это красное смещение источника, а (t_0-t) — расстояние до него в световых годах, измеренное вдоль светового конуса прошлого.

В астрономии часто пользуются *расстоянием по светимости* r , определяющим поток \mathcal{F} принимаемого на Земле излучения от точечного источника со светимостью L_0 , как если бы и источник, и наблюдатель находились в евклидовом пространстве:

$$\mathcal{F} = \frac{L_0}{4\pi r^2}.$$

Вместо \mathcal{F} и L_0 можно использовать *видимую m и абсолютную M звездные величины*, которые соотносятся с расстоянием r между источником и наблюдателем следующим образом:

$$m = M + 5 (\lg r [\text{пк}] - 1). \quad (12.4.3)$$

Чтобы знать абсолютную звездную величину источника, надо понимать его устройство и то, каким образом он светит, то есть иметь предсказательную теорию. Как правило, эта задача не решается с требуемой точностью, поскольку речь идет о сложной газодинамической нелинейной системе со многими параметрами, будь то галактика или скопление галактик, квазар или звезда. Поэтому часто прибегают к эмпирическим соотношениям, позволяющим оценить светимость объекта по другим его наблюдательным характеристикам. Эти феноменологические соотношения строятся и проверяются по ближайшим известным источникам, а затем используются для калибровки далеких объектов данного класса. При этом предполагается схожесть близких и удаленных представителей одного и того же класса источников (гипотеза *стандартной свечи*).

На роль стандартных свеч в разные годы претендовали разные объекты ¹⁾, однако эти исследования не смогут быть использованы в точной космологии до тех пор, пока не будет разработана теория столь сложных систем, опирающаяся на численное моделирование и учитывающая многие пока малоизученные факторы, относящиеся не только к физической природе объектов (например, звездообразование, нестационарность, химический состав, термоядерное горение и др.), но и к их окружению, условиям распространения света и пр. При работе с такими данными всегда будет существовать вероятность того, что мы имеем дело не с эффектами геометрии или состава Вселенной, а с внутренними свойствами источников и их эволюцией (неконтролируемая систематика).

Если для использования хаббловских диаграмм требуется стандартная свеча, то для измерения расстояний по угловому диаметру необходимо знать физический размер наблюдаемого объекта (гипотеза *стандартного размера*).

Пусть физический размер объекта равен d , а его угловой диаметр — θ . Тогда *расстояние по угловому диаметру* D до него определяется следующим образом (ср. (2.7.14)):

$$D = \frac{d}{\theta}.$$

Используя интервал (2.7.11) фридмановской геометрии, получаем связь между введенными выше расстояниями:

$$r = (1 + z)^2 D = (1 + z) R, \quad (12.4.4)$$

где геодезическое сопутствующее расстояние R до объекта определяется следующим выражением (см. (2.7.12)):

$$R \equiv \frac{\bar{R}}{H_0} = \eta_0 - \eta(z), \quad (12.4.5)$$

$$\bar{R} = \bar{R}(z) = H_0 \int_0^z \frac{dz}{H}. \quad (12.4.6)$$

Очевидно, космологическая функция $\bar{R}(z)$ конечна, зависит только от свойств материи и не зависит от H_0 .

¹⁾ Так, в 70–80-х годах XX века в качестве стандартной свечи использовали ярчайшие сD-галактики в центрах богатых скоплений. В настоящее время на эту роль претендуют сверхновые разных классов, гамма-всплески, рентгеновские скопления галактик и другие яркие источники, которые можно наблюдать на больших расстояниях.

С помощью формул (12.4.4) любая диаграмма $m(z)$ может быть формально представлена в виде зависимости $\theta(z)$, и наоборот. В этом нет ничего удивительного, ведь мы исследуем характеристики *единой* геометрии, хотя бы и с помощью разных наблюдательных тестов.

Итак, «рецепт» измерения Вселенной прост: нужно строить диаграммы $\theta(z)$ и $\dot{\theta}(z)$ или $m(z)$ и $\dot{m}(z)$ и др. Если известен физический размер источника или его абсолютная звездная величина, то можно определить и расстояние до него, а следовательно, и размеры Вселенной. Это позволяет восстановить функции $H(z)$, $\gamma(z)$ и другие функции масштабного фактора, а значит, определить состав и свойства материи и эволюцию этих свойств со временем.

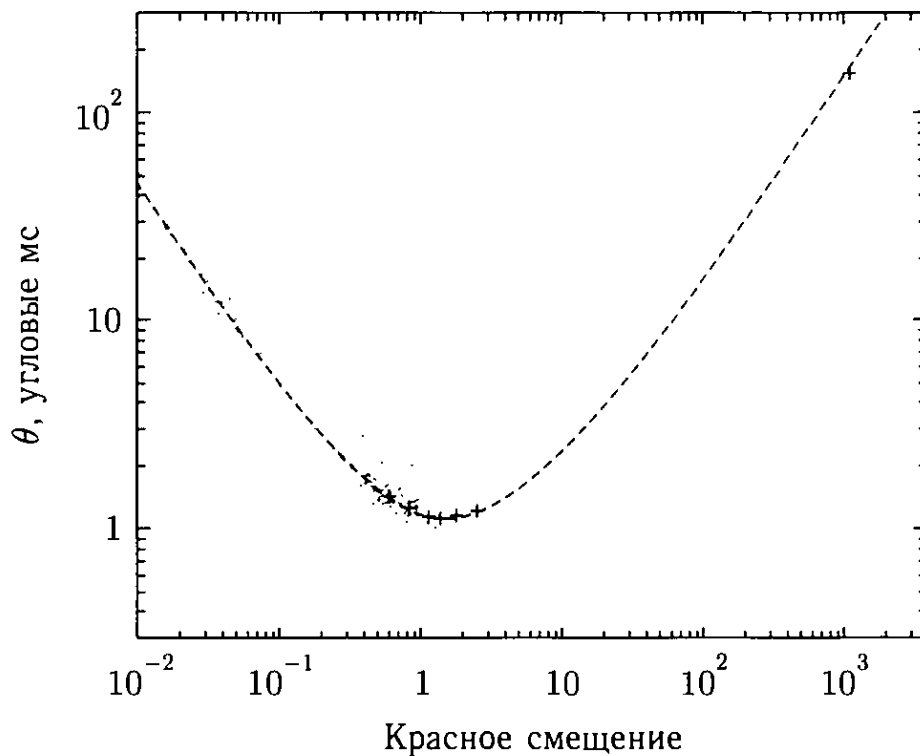


Рис. 12.7. Диаграмма $\theta(z)$, построенная по трем геометрическим тестам, приведенным к стандартному размеру $d = 9,5$ пк [Jackson and Jannetta, 2006]

Реализация этой задачи требует высокой точности измерений и знания природы наблюдаемых объектов. При этом выбранные объекты могут быть весьма разнообразны.

Рисунок 12.7 представляет собой сводную диаграмму $\theta(z)$ для нескольких геометрических тестов, приведенных к стандартному размеру $d = 9,5$ пк. Здесь представлены результаты измерений космологических расстояний по светимостям далеких сверхновых типа Ia (левые голубые точки), угловым размерам ультракомпактных радиисточников (оснований джетов радиоак-

тивных галактических ядер, шесть точек в нижней части кривой) и анизотропии реликтового излучения (звуковой модуляции спектра, правая верхняя точка). Штриховая кривая соответствует космологической стандартной модели.

Наиболее надежной на этой диаграмме является правая точка, определенная по угловому размеру звукового горизонта в момент рекомбинации $z_{\text{rec}} = 1100$ (см. раздел 11.2). Точки, построенные по наблюдениям сверхновых звезд и ультракомпактных радиисточников, опираются соответственно на гипотезы стандартных свеч и размеров, для чего имеются определенные физические основания.

В первом случае таким физическим основанием является природа сверхновой типа Ia — термоядерный взрыв сливающихся белых карликов или аккрецирующего белого карлика с *предельной* чандрасекаровской массой, задающей его максимальную светимость (см. раздел 3.2). Во втором — *предельная* масса центральной аккрецирующей черной дыры, обеспечивающая «стандартный» размер радиоярких квазаров. Бесспорно, класс черных дыр с максимальной массой существует (она, по-видимому, близка к $10^{10} M_{\odot}$). Черные дыры с большими массами просто не успели бы образоваться в нашей Вселенной за конечное время формирования галактических ядер (см. раздел 10.6). В качестве комментария к рис. 12.7 отметим также, что имеющиеся модели взрывов сверхновых звезд и активных квазаров пока не отвечают требованиям, предъявляемым к точным геометрическим измерениям. Скорее наоборот: зная ответ (параметры КСМ), можно попытаться понять собственную космологическую эволюцию подобных типов компактных источников.

Рисунок 12.8 демонстрирует пример ультракомпактного радиисточника с минимальным разрешением изображения порядка парсека, достигнутым с помощью наземной радиоинтерфериметрической сети. Хорошо видно основание джета, однако внутренняя структура самого источника неразрешена. Этот и другие примеры показывают, что исследование нелинейных компактных объектов, составляющих основу для прямых геометрических измерений Вселенной, требует развития новых технологий, реализация которых возможна при использовании средств космической радиоинтерфериметрии.

С помощью космических телескопов база интерферометра увеличивается и, соответственно, повышается его разрешающая способность. Такие инструменты способны не только разрешить удаленные компактные объекты, но и измерить *изменение* их угловых размеров и красных смещений со временем из-за расши-

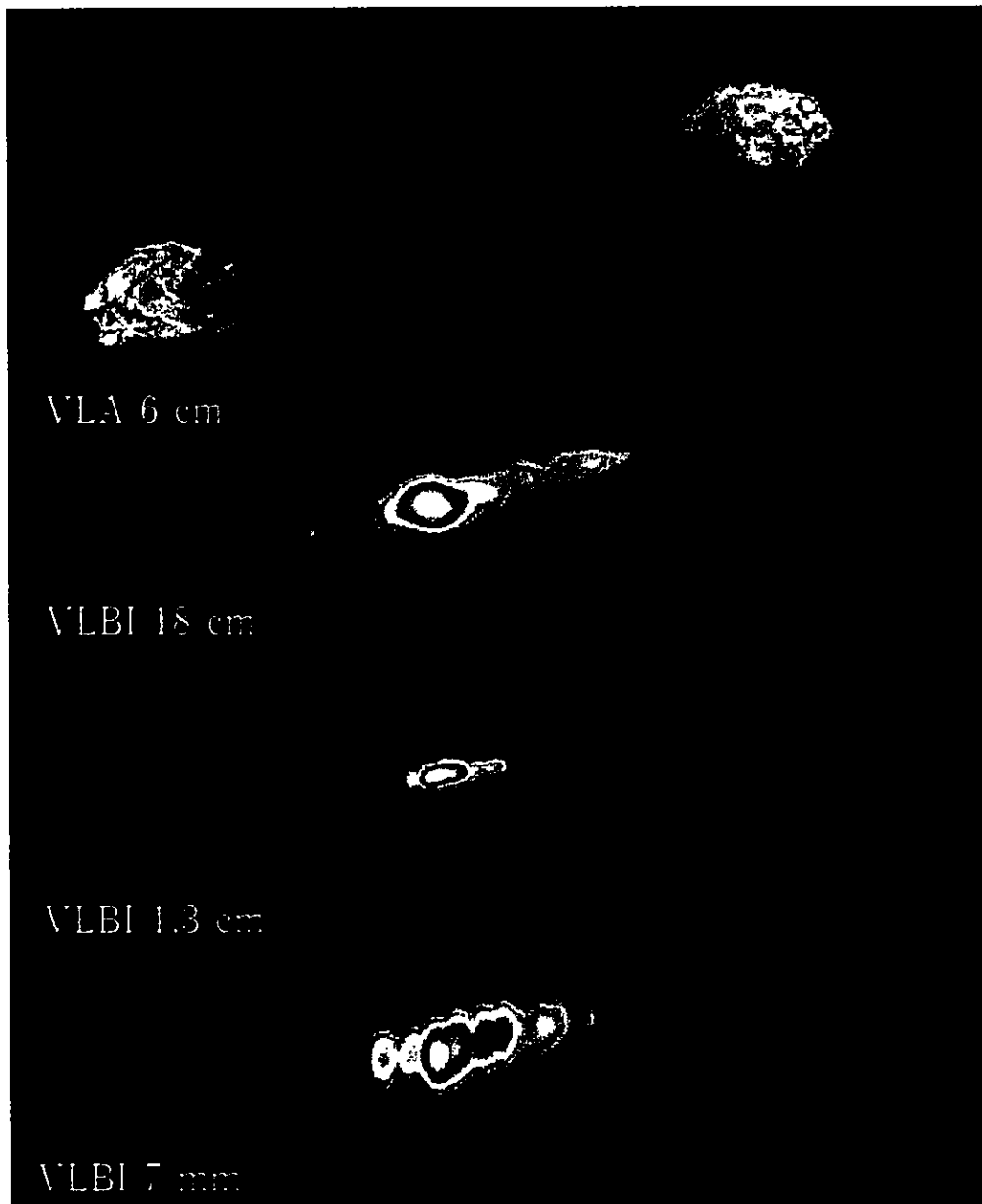


Рис. 12.8. Ультракомпактный центральный радиоисточник в галактике Лебедь А [Krichbaum et al., 1998]

рения Вселенной. Последние два фактора прямо связаны с функцией Хаббла ¹⁾:

$$\frac{\dot{\theta}}{\theta} = -\frac{H(z)}{1+z}, \quad \frac{\dot{z}}{1+z} = H_0 - \frac{H(z)}{1+z}, \quad (12.4.7)$$

что открывает новые возможности для определения H_0 и $H(z)$ на космологических расстояниях.

¹⁾ Эти соотношения получаются после дифференцирования функций

$$\theta = \frac{d}{R a(\eta - R)} \quad \text{и} \quad 1+z = \frac{a(\eta)}{a(\eta - R)}$$

по η при постоянных d и R , где η — время наблюдателя, а $(\eta - R)$ — конформное время источника (рис. 12.9).

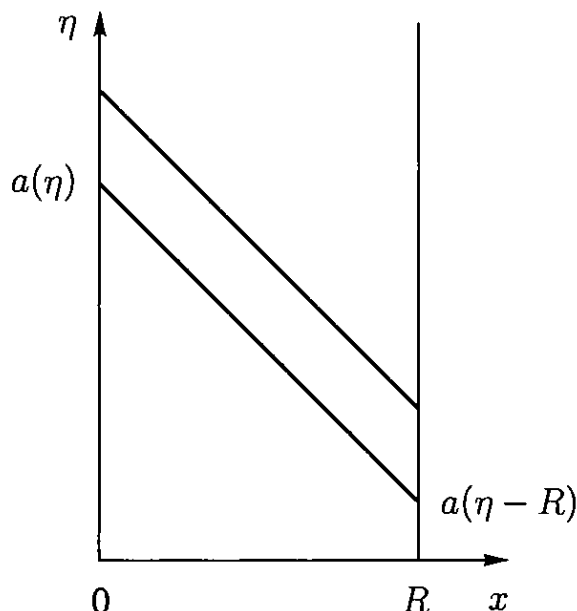


Рис. 12.9. Траектории лучей света в модели Фридмана в конформных координатах (η, x) от источника ($x = R$) к наблюдателю ($x = 0$)

Заметим, что если плотность ТЭ была постоянна, то красное смещение источников хаббловского потока, находящихся при $z_0 \simeq 2,3$, не изменяется в современную эпоху: $\dot{z}(z = z_0) = 0$. При этом смещение более далеких объектов ($z > z_0$) уменьшается с течением времени, а более близких — нарастает: $\dot{z}(z < z_0) > 0$. Эта граница $z = z_0$ весьма чувствительна к эволюции ТЭ. По ее детектированию можно будет ответить на вопрос, каким образом плотность ТЭ изменялась в прошлом.

Определение эволюционных эффектов (12.4.7) на космологических расстояниях потребует нескольких лет мониторинга при угловом разрешении порядка микросекунд дуги, что является довольно сложной технической задачей. Однако уже первые космические интерферометры смогут измерить расстояния до близких галактик по их смещению на фоне далеких источников из-за движения Солнечной системы со скоростью $V_{\odot} \simeq 370 \text{ км} \cdot \text{с}^{-1}$ относительно реликтового излучения (галактический параллакс).

Кроме того, можно измерять геометрические расстояния, наблюдая сверхсветовые движения узлов джетов активных ядер галактик, направленных под малым углом к лучу зрения. Такие измерения уже проводятся на Земле, и их можно существенно улучшить с помощью все тех же космических интерферометров.

Поясним этот эффект с помощью рис. 12.10. Пусть α — малый угол между направлением движения узла джета (конец красной стрелки) и линией, соединяющей наблюдателя «о» с источником «s» (квазаром), η и $x = (R, y, 0)$ — конформные координаты узла (R — геодезическое расстояние от наблюдателя до узла, см. (12.4.5)). Время наблюдения узла джета, движу-

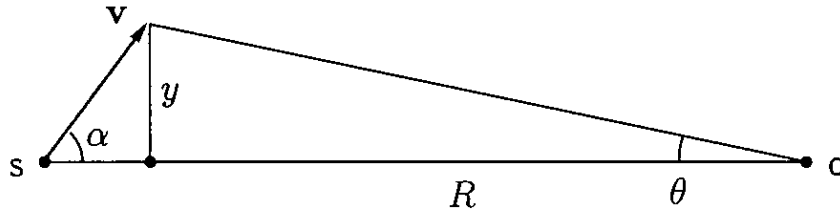


Рис. 12.10. Плоскость (R, y) наблюдателя «о» и узла джета (конец красной стрелки), движущегося со скоростью v относительно источника «s»

щегося с собственной скоростью v , равно

$$d\tau = d\eta - dR = (1 - v \cos \alpha) d\eta,$$

а видимая скорость его перемещения в картинной плоскости наблюдателя составляет величину:

$$v_{\perp} = \frac{dy}{d\tau} = \frac{v \sin \alpha}{1 - v \cos \alpha} \leq v\Gamma. \quad (12.4.8)$$

Эта функция имеет максимум при $\cos \alpha = v$, численно равный гамма-фактору джета $\Gamma \equiv (1 - v^2)^{-1/2}$ при скорости $v = |\mathbf{v}|$, близкой к скорости света.

Таким образом, смещение джета на небесной сфере подчиняется следующей формуле:

$$\dot{\theta}_{\perp} = \frac{v_{\perp}}{R} \leq \frac{0,3}{R} \left(\frac{\Gamma}{20} \right) \text{ угл. мс/год}, \quad (12.4.9)$$

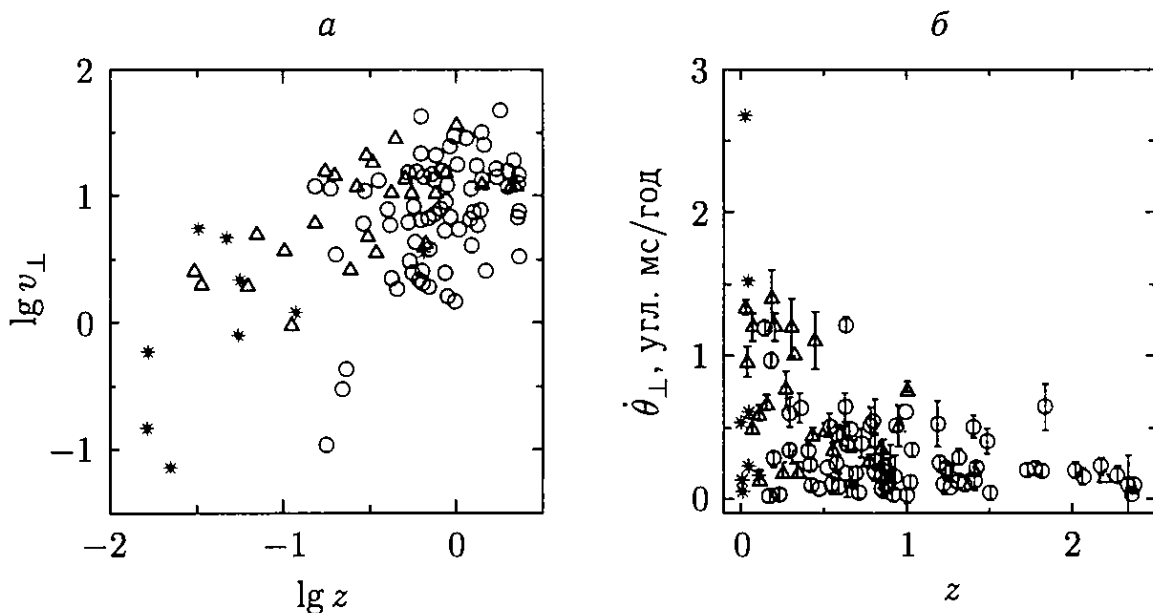


Рис. 12.11. Сверхсветовые выбросы из квазаров [Zhang and Fan, 2008].
 а — поперечная скорость джета $v_{\perp}(z)$, б — угловое смещение $\dot{\theta}_{\perp}(z)$

что составляет величину порядка миллисекунды дуги в год для реально наблюдаемых сверхсветовых скоростей выбросов у квазаров (рис. 12.11 б). Имея большую статистику таких объектов, можно определить геометрическую характеристику $\bar{R}(z)$ (см. (12.4.6)) по верхней огибающей функции (12.4.9). Пример подобного анализа, дающий представление о текущем уровне точности этого метода, приведен на рис. 12.11.

Мы привели несколько примеров, демонстрирующих возможности прямого определения геометрии и состава Вселенной по наблюдениям нелинейных объектов. Этот ряд можно дополнить наблюдениями ярких галактик, компактных групп галактик, богатых рентгеновских скоплений, гамма-всплесков, двойных квазаров и других активных систем. Исследования Вселенной с их помощью пока не столь эффективны, как при использовании квазилинейных систем, базирующихся на измерениях крупномасштабной структуры реликтового излучения и нерелятивистской материи. Можно надеяться, что развитие численного моделирования позволит геометрическим методам стать конкурентоспособными при решении точных задач наблюдательной космологии.

12.5. Образование галактик в неоднородной Вселенной

Рассмотрим, каким образом в развивающихся нестационарных системах возникают гравитационно-связанные гало материи, поддерживаемые гидростатическим равновесием¹⁾, и как распределение этих нелинейных объектов по массам связано с начальным полем возмущений плотности.

В отличие от крупномасштабной структуры, связанной с относительно небольшими вариациями плотности и находящейся в начале своего нелинейного развития, галактики образуются из высоких пиков плотности, типичных для мелкомасштабного (менее нескольких $M_{\text{пк}}$) поля возмущений, и прошли уже большую часть пути своей нелинейной эволюции и процесса релаксации (см. разделы 3.3–3.6). Причиной этого является амплитуда и форма начального спектра возмущений плотности, возрастающего в сторону коротких длин волн (рис. 6.2). Это

¹⁾ Мы также называем равновесные гало *релаксированными* или *вириализованными* системами, понимая, что речь идет о вириализации по скоростям, но не по энергии (см. раздел 3.4).

свойство спектра обеспечивает раннюю фрагментацию вещества на маломассивные «гало-блоки» и их последующий процесс сгущивания в более массивные гало галактик, групп и скоплений.

Иерархическое формирование нелинейных структур от малых масс к большим подтверждается многочисленными наблюдениями и численным моделированием. Важнейшим наблюдательным аргументом в пользу последовательного образования гало является отсутствие большого числа далеких эмиссионных систем, излучающих в линии L_{α} . Это объясняется тем, что массивные системы не переживали в прошлом мощной одиночной вспышки звездообразования, охватывающей одновременно весь газ гало, а формировались путем последовательного слияния многочисленных маломассивных блоков, в которых процесс звездообразования проходил в разное время.

В силу имеющихся свойств пространственного S-спектра видимая структура Вселенной имеет следующие свойства.

- Процесс образования галактик и групп проходил в прошлом и к настоящему времени завершен: основная масса темной материи находится в этих образованиях.
- Гало и их характеристики распределены неравномерно по пространству: их число и массы промодулированы крупномасштабной структурой Вселенной.

Первый пункт дает физические основания для установления простой взаимосвязи возникших во Вселенной нелинейных гало с пиками линейных затравочных возмущений плотности, основанной на модели квазисферического коллапса (метод Пресса–Шехтера). Второй пункт позволяет применить данный метод к большим, но конечным областям распределения материи и, тем самым, связать функцию масс гало с крупномасштабной структурой Вселенной.

Прежде чем перейти к теоретическим расчетам функций масс гало и их проверке с помощью наблюдений и численного эксперимента, напомним основные характеристики и элементы крупномасштабной структуры Вселенной.

Как мы уже подчеркивали, переход к большим масштабам переносит нас к начальным стадиям формирования структуры и к первым фазам коллапса, носящим преимущественно одномерный характер (см. раздел 12.3). Многолетние наблюдения пространственного распределения гравитационно-связанных вирialized гало темной материи на масштабах до 300 Мпк

демонстрируют существование следующих типичных элементов крупномасштабной структуры Вселенной:

- *филаменты* — линейные структуры разного богатства¹⁾ длиной до десятков Мпк и толщиной несколько Мпк;
- *стенки* — плоские образования размерами по протяженным направлениям до сотни Мпк и толщиной несколько Мпк, заполненные филаментами;
- *узлы* — богатые скопления галактик размерами в несколько Мпк, расположенные в местах пересечения филаментов;
- *сверхскопления* — протяженные области пространства размерами от десятка до сотен Мпк, заполненные стенками, филаментами и узлами;
- *космологические «пустоты» (войды)*²⁾ — протяженные области размерами до сотен Мпк, в которых отсутствуют яркие галактики и скопления галактик.

Последние два типа образований квазилинейны из-за их большого размера. Под действием гравитации вещество движется в направлении сверхскоплений, эти области расширяются медленнее среднего хаббловского потока. Из войдов вещество «оттекает» в разных направлениях, их локальный объем увеличивается быстрее среднего. Полное смещение элемента материи за всю историю Вселенной меньше размеров самих структур и составляет в среднем около ~ 14 Мпк (см. (3.5.4)).

Войды занимают наибольший объем Вселенной (более 50%), и их формы по контурам низкой плотности близки к сферическим, тогда как сверхскопления — это более плотные образования, имеющие уплощенные формы. Оба типа структур естественно характеризовать средней плотностью вещества в заданном объеме: для войда она меньше средней плотности вещества во Вселенной, а для сверхскопления — больше.

Наблюдательное определение параметров войдов и сверхскоплений неоднозначно. Если войды — это лакуны в пространственном распределении галактик, то следует указать яркость и класс объектов, по которым они выделяются. Анализ наблюдательных

¹⁾ Богатство элемента крупномасштабной структуры определяется числом входящих в него галактик.

²⁾ Термин космологические «пустоты» отражает тот факт, что на картах неба некоторые участки кажутся пустыми. На самом деле, в космологических «пустотах» может содержаться значительная (но не большая) часть вещества Вселенной, то есть «пустоты» не пустые. По этой причине вместо выражения космологические «пустоты» часто употребляют используемый в устной речи термин «войд» (от англ. «void»).

данных и численный эксперимент показывают, что характерные размеры войдов, определяемые по галактикам невысокой светимости, меньше, чем их размеры, полученные по ярким объектам.

Причина размытости войдов или сверхскоплений, восстанавливаемых по пространственному распределению света, заключена в нетривиальной физической связи между светящимся веществом и темной материей, составляющей основу структуры (см. раздел 10.6). Более правильным и точным является выделение этих образований по пороговому контрасту плотности входящей в них материи относительно среднего уровня.

Далее мы определяем войды (и сверхскопления) как трехмерные пространственные области с заданным отрицательным (положительным) уровнем контраста плотности материи. При уменьшении модуля порогового значения контраста плотности размеры структур растут, а их количество уменьшается. Так, при $\rho < \rho_c$ мы имеем единственный войд, заполняющий всю Вселенную.

Важной характеристикой войдов и сверхскоплений является функция масс населяющих их гало. Наблюдения показывают, что ближайšie к нам космологические пустоты обеднены нормальными и карликовыми галактиками, однако в них встречаются слабые эмиссионные галактики и газовые облака, наблюдаемые как ультрафиолетовые линии поглощения в спектрах расположенных за ними квазаров или ярких галактик, а также линии поглощения нейтрального водорода (21 см) в спектрах удаленных радиоисточников. Галактики в войдах в среднем имеют меньшую светимость, чем галактики в сверхскоплениях, хотя наклоны их функций светимости практически не отличаются друг от друга.

Теория предсказывает наличие в войдах и сверхскоплениях маломассивных слабых галактик, не прошедших стадию активного звездообразования, а также первичных объектов модели иерархического сгущивания с массами вплоть до $10^5 M_\odot$ и менее. Такие объекты не способны удержать газ, а значит звездообразование в них практически невозможно. С увеличением чувствительности измерений наблюдателям удастся обнаружить в Местной группе и ближайшем окружении ранее неизвестные слабые галактики с массами $10^8 - 10^9 M_\odot$, их число согласуется с предсказанием стандартной модели, однако вопрос обнаружения менее массивных галактик остается пока открытым.

Минимальный масштаб, на котором обрывается спектр начальных возмущений плотности, связан с физикой темной материи и пока неизвестен. Область малых масс, поддающихся исследованиям средствами наблюдательной космологии (например,

по наблюдениям «леса» линий поглощения в спектрах далеких квазаров), можно связать с температурой T Вселенной в момент времени, когда сопутствующий размер R данной массы M_R совпадал с хаббловским радиусом Вселенной:

$$R \simeq 10 \left(\frac{10 \text{ кэВ}}{T} \right) \text{ кпк}, \quad M_R \simeq 10^5 \left(\frac{10 \text{ кэВ}}{T} \right)^3 M_\odot.$$

Численное моделирование процессов образования структуры из холодной материи существенно продвинуло исследование функций масс нелинейных гало в разных областях Вселенной. Начиная процесс образования при $z \sim 20$ с затравочными массами $M \sim 10^5 M_\odot$, удается к $z \sim 0$ получить сходное с наблюдаемым пространственное распределение гравитационно-связанных систем с типичными массами $\sim 10^{12} - 10^{13} M_\odot$. При этом вся темная материя оказывается вошедшей в состав вириализованных гало всевозможных масс, а средняя плотность материи внутри гало в момент его образования в ~ 200 раз превышает среднюю плотность во Вселенной, что соответствует предсказаниям квазисферического коллапса.

В последующих разделах мы рассмотрим простые теоретические аппроксимации, описывающие локальные свойства распределений нелинейных гало во Вселенной.

12.6. Функция масс релаксированных гало

Аппроксимация Пресса–Шехтера базируется на двух основных положениях.

- Поле линейных возмущений является гауссовым: доля объема Вселенной со сглаженной плотностью $\delta_R(z, \mathbf{x})$ (3.4.4) выше некоторого порогового значения δ_c равна

$$f(M, z) = \frac{1}{\sqrt{2\pi}} \int_{\nu}^{\infty} e^{-\nu'^2/2} d\nu', \quad (12.6.1)$$

где

$$\nu = \nu(M, z) \equiv \frac{\delta_c}{\sigma_R(z)} \quad (12.6.2)$$

— монотонно растущая функция M (см. (3.4.2), (3.5.3)),

$$\sigma_R(z) = \bar{g}(z) \cdot \sigma_R \quad (12.6.3)$$

— дисперсия контраста плотности в сфере радиуса R (см. (3.4.5)), $M \equiv M_R = 4\pi\rho_m R^3/3$ — масса вещества

внутри этой сферы (см. (3.4.1)), $\bar{g} = \bar{g}(z)$ — фактор роста возмущений плотности (см. (12.1.15)).

- Вещество, находящееся в пиках плотности, коллапсирует и удерживается в гравитационно-связанных объектах: существует такое число δ_c , при котором доля массы вещества во Вселенной ¹⁾ $F(> M, z) = 2f(M, z)$ оказывается к моменту z в составе вириализованных гало с индивидуальными массами выше M :

$$F(> M, z) = 2f(M, z) = \frac{1}{\rho_m} \int_M^{\infty} n dM, \quad (12.6.4)$$

где

$$n = n(M, z) \equiv \frac{\rho_m}{M} f_M \quad (12.6.5)$$

— средняя плотность числа гало с массами M в интервале $dM \sim M$ (дифференциальная функция масс),

$$f_M = f_M(z) \equiv \left| \frac{d \ln \sigma_R}{d \ln M} \right| f(\nu) \quad (12.6.6)$$

— доля вещества, оказавшаяся к красному смещению z в объектах класса M (в гало с массами M).

Сравнение формул (12.6.1) и (12.6.4) позволяет определить среднюю функцию масс нелинейных гало во Вселенной:

$$f(\nu) = f_{PS} \equiv \sqrt{\frac{2}{\pi}} \nu e^{-\nu^2/2}. \quad (12.6.7)$$

Нормировка функции f_M

$$\int_0^{\infty} f_M \frac{dM}{M} = 1 \quad (12.6.8)$$

означает, что при любом z вся темная материя входит в состав нелинейных гало. Речь может идти лишь о распределении этих гало по массам и о типичных для данного z объектах гало, содержащих основную долю материи Вселенной. При больших z вся материя распадается на малые массы, однако с ростом вре-

¹⁾ Коэффициент 2 означает, что в ходе иерархического скучивания *все* холодное вещество оказывается в составе вириализованных гало всевозможных масс: $F(> 0, z) = 1$ при любом z . Аргумент M удобнее, чем R , поскольку M сохраняется как для линейных возмущений, так и для объектов гало. При включении в состав материи компоненты массивных нейтрино необходимо учитывать зависимость фактора роста \bar{g} от пространственного масштаба (см. раздел 12.8).

мени доля крупных гало растет. Масса характерных гало M_* в любую эпоху определяется условием:

$$\nu(M_*, z) = 1.$$

В современной Вселенной $M_* \simeq 10^{13} M_\odot$ (рис. 12.14).

При малых M : $f_M \sim \nu \ll 1$. Это означает, что объекты малой массы захватываются более массивными гало в ходе процесса иерархического сгущивания, то есть они *выбывают* из числа объектов малой массы M и учитываются только во вновь образующихся гало более высокой массы.

Для вычисления δ_c можно воспользоваться приближением сферически-симметричного однородного коллапса¹⁾ (см. (2.5.5)). Аппроксимация линейного контраста плотности КСМ к моменту образования гало дает численное значение порогового контраста $\delta_c = 1,675$, которое очень слабо зависит от вариаций космологических параметров в пределах ошибок измерений.

Многолетние исследования показали, что результаты, полученные с помощью этого метода, хорошо согласуются с численным экспериментом N тел. Тем не менее, базовый формализм не включает в себя механизмы слияния и приливного разрушения объектов, поэтому, как правило, метод Пресса–Шехтера используют при анализе пространственного распределения объектов с достаточно большой массой (таких, как скопления галактик), для которых эти эффекты незначительны.

Для применения данного метода в точной космологии следует учесть поправки на несферичность коллапса. Аналитические оценки показывают, как при этом видоизменяется функция $f(\nu)$. Поэтому для построения точной аналитической аппроксимации необходимо использовать более общую эмпирическую формулу, включающую возможность несферической коррекции (аппрокси-

¹⁾ В рамках данной простой модели момент образования z гало совпадает с моментом коллапса циклоиды $t(z) = 2\pi a_0 / \sqrt{x}$ (см. (2.5.5)). Можно также оценить плотность вириализованного гало в момент его образования: она в 8 раз превышает плотность коллапсирующего шара в момент его остановки, поскольку, согласно теореме вириала, радиус гало — радиус сферы, содержащей массу M однородного вириализованного гало, — вдвое меньше радиуса, охватывающего ту же массу в момент остановки циклоиды. Отсюда, в частности, получаем полезную оценку радиуса гало в момент его образования: он приблизительно в 6 раз меньше радиуса сферы, охватывающей массу M в невозмущенной Вселенной. Образно говоря, гало в момент его рождения — это в 6 раз сжатый по масштабу кусок невозмущенной Вселенной.

мация Шета–Тормена):

$$f(\nu) = f_{ST} \equiv f_0 \sqrt{\frac{2}{\pi}} \tilde{\nu} (1 + \tilde{\nu}^{-p}) e^{-\tilde{\nu}^2/2}, \quad (12.6.9)$$

где

$$\tilde{\nu} = \frac{\nu}{\nu_0},$$

и численно определить поправочные коэффициенты. Они равны:

$$f_0 = 0,32, \quad \nu_0 = 1,2, \quad p = 0,6$$

(очевидно, f_0 — это нормировочный коэффициент, определяемый из условия (12.6.8)).

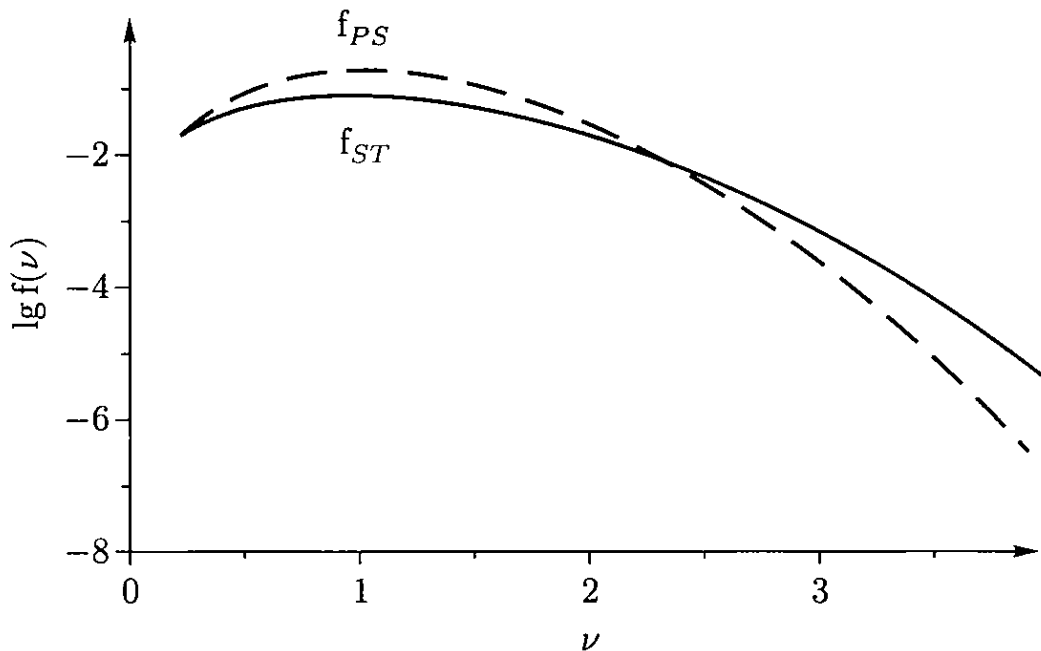


Рис. 12.12. Функции $f(\nu)$ в аппроксимациях функций масс гало Пресса–Шехтера (f_{PS} , штриховая линия) и Шета–Тормена (f_{ST} , сплошная линия)

Такая скорректированная аппроксимация, учитывающая гауссовый характер поля возмущений и гравитационный коллапс пиков плотности, *идеально* описывает результаты численного эксперимента и *широко* применяется в современной космологии. Рисунок 12.12 иллюстрирует функции $f(\nu)$ в космологической стандартной модели. Хорошо видно, что аппроксимация Пресса–Шехтера несколько переоценивает число гравитационно-связанных гало с массами вблизи характерной M_* ($\nu = 1$), но недооценивает число массивных гало с $M > 5 \cdot 10^{14} M_\odot$. Пересечение двух функций происходит при $\nu = 2,3$, что соответствует массе около $10^{14} M_\odot$ в современную эпоху (точное число зависит от нормировки спектра, см. раздел 12.8).

В своей стандартной форме формализм Пресса–Шехтера позволяет рассчитать только среднюю функцию масс вириализованных гало без учета таких крупномасштабных неоднородностей, как войды или сверхскопления, в которых локальная плотность материи ρ_L в масштабе L отличается от фоновой ρ_m . В следующем разделе мы получим функции масс гало, населяющих неоднородную Вселенную.

12.7. Модуляция галактик крупномасштабной структурой

Локальные протяженные области Вселенной удобно характеризовать средней плотностью материи внутри сферы сопутствующего радиуса L с центром в точке \mathbf{r} (ср. (3.4.4)):

$$\rho_L = \rho_L(\mathbf{r}, z) = \int \rho(\mathbf{r}', z) W_L(|\mathbf{r} - \mathbf{r}'|) d\mathbf{r}', \quad (12.7.1)$$

где

$$W_L(r) = \frac{3}{4\pi L^3} \begin{cases} 1, & r \leq L, \\ 0, & r > L. \end{cases}$$

Технический параметр L предполагается больше масштаба неоднородности ($L > 10$ Мпк), так что пространственные области с $\rho_L < \rho_m$ (войды «v») и $\rho_L > \rho_m$ (сверхскопления «s») находятся еще на стадии квазихаббловского расширения. Введем параметр локальной плотности материи:

$$\Omega_L \equiv \Omega_m \frac{\rho_L}{\rho_m}. \quad (12.7.2)$$

Суммируя (12.7.2) с Ω_E , получаем полную плотность энергии в области L . Вычитая отсюда равенство $\Omega_m + \Omega_E = 1$ для плоской Вселенной, получаем параметр Δ_L , характеризующий среднюю кривизну области L :

$$\Delta_L \equiv \Omega_L - \Omega_m \equiv \Omega_m \delta_L, \quad (12.7.3)$$

где

$$\delta_L \equiv \frac{\rho_L}{\rho_m} - 1 \quad (12.7.4)$$

— средний контраст плотности в области L ($|\delta_L| < 1$). При $L \rightarrow \infty$ ρ_L стремится к средней плотности материи во Вселенной $\rho_m = \rho_m(z)$.

Для построения *локальной* функции масс гало следует:

- разделить поле линейных возмущений плотности δ_R на крупномасштабную часть δ_L (подложка) и мелкомасштабную $\delta_{R|L}$, характеризующую затравочные возмущения для

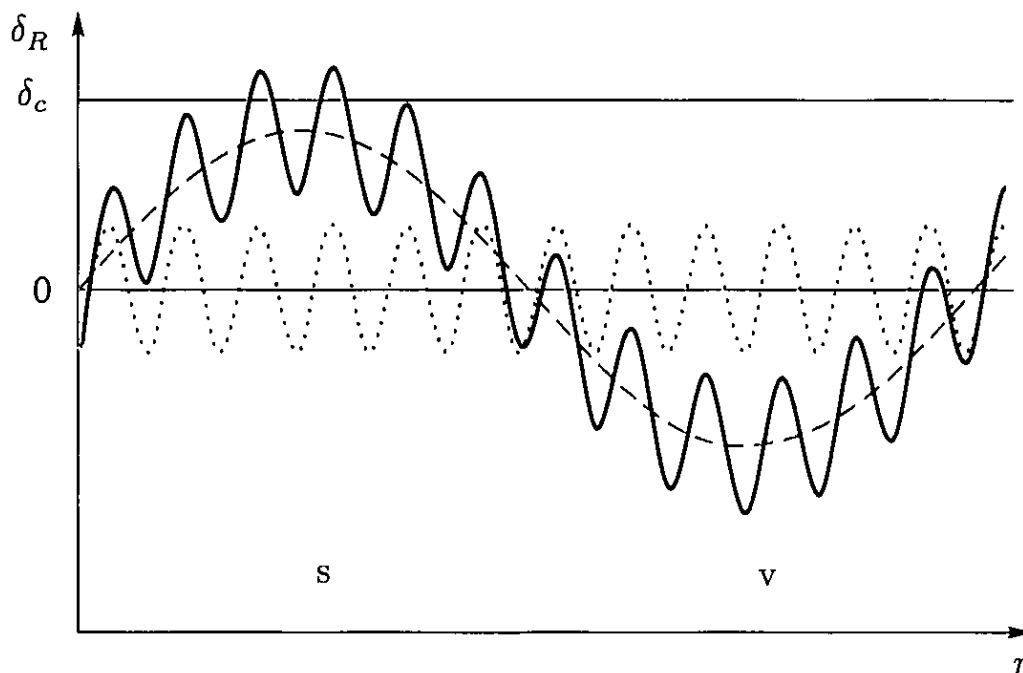


Рис. 12.13. Пространственное поле возмущений плотности $\delta_R(r)$ (сплошная линия) в виде суммы длинноволновой подложки δ_L (штриховая линия) и коротковолнового остатка $\delta_{R|L}$ (пунктирная линия); δ_c — пороговый контраст плотности

образования гало в данной области пространства размером $L > R$ (рис. 12.13):

$$\delta_R = \delta_L + \delta_{R|L}, \quad (12.7.5)$$

- описать квазисферический коллапс на подложке ρ_L , а не на однородном фоне с плотностью ρ_m .

С физической точки зрения речь идет о влиянии местной крупномасштабной структуры Вселенной на локальный процесс иерархического сгущивания. Предполагая масштаб структуры большим ($L > R$), ее влияние на коллапс можно представить в виде ряда по сферическим гармоникам. Монопольный член этого ряда является главным фактором, влияющим на локальную статистику вириализованных гало, что составляет основу данного метода.

Динамика коллапса зависит только от массы вещества, заключенной внутри сферы радиуса R , однако области, где контраст плотности превышает пороговый, реже встречаются в войдах и чаще в сверхскоплениях. Эта информация и заключена в подложке, модулирующей гауссову функцию $\delta_{R|L}$. На этом этапе не учитываются дипольный (градиент плотности), квадрупольный (приливные силы) и другие высшие члены разложения структуры, связанные соответственно с движением гало как целого и несферичностью коллапса. Влияние этих факторов можно

учесть численно в виде поправочных коэффициентов к функции масс гало (см. (12.6.9)).

Дисперсия линейного гауссова поля $\delta_{R|L}$, описывающего локальные свойства возмущений плотности при $R < L$, имеет вид ¹⁾:

$$\sigma_{R|L}^2 = \int_0^{\infty} P(k) |W(kR) - W(kL)|^2 k^2 dk. \quad (12.7.6)$$

Очевидно, $\sigma_{R|L} \rightarrow 0$ при $M \equiv M_R \rightarrow M_L$ (масса в сфере радиуса L , см. (3.4.1)).

Проводя вычисления, аналогичные тем, что были сделаны в предыдущем разделе, получаем функцию масс гало в области L в интервале масс $M \in (0, M_L)$:

$$n_L = n_L(M) \equiv \frac{\rho_L}{M} f_{M|L}, \quad (12.7.7)$$

$$f_{M|L} = \left| \frac{\partial \ln \sigma_{R|L}}{\partial \ln M} \right| f(\nu_L), \quad \int_0^{M_L} f_{M|L} \frac{dM}{M} = 1,$$

где функция $f(\nu)$ совпадает по форме с (12.6.7) (или с (12.6.9) в случае коррекции на несферичность), а величина ν_L следует из (12.6.2) с заменой σ_R на $\sigma_{R|L}$ и δ_c на $\delta_{c|L}$.

Функция $\nu_L = \nu_L(M, z)$ монотонно растет с ростом массы и расходится при $M \rightarrow M_L$. Пороговый контраст плотности $\delta_{c,L}$ зависит от локальной плотности подложки. При расчете $\delta_{c,L}$ необходимо иметь в виду, что области сверхскоплений $\delta_L > 0$ (или войдов $\delta_L < 0$) расширяются медленнее (быстрее) по отношению к однородному фону. Это требует меньшего (большого) значения порогового контраста $\delta_{c,L}$, определенного относительно подложки, для того чтобы данная область повышенной плотности сколлапсировала к моменту z :

$$\delta_{c|L} = \delta_c - \delta_L. \quad (12.7.8)$$

Эта формула справедлива в линейном порядке по δ_L . Для учета нелинейного контраста плотности подложки (но $|\Delta_L| < \Omega_m$, см. (12.7.3)) необходима численная коррекция.

На рис. 12.14 для $z = 0$ представлены функции масс гало $M n_L / \rho_m$ в войде, сверхскоплении и плоской области про-

¹⁾ Здесь мы не учитываем кросс-корреляцию полей $\delta_{R|L}$ и δ_L , которая важна только при $R \sim L$ (см., напр., рис. 3.14).

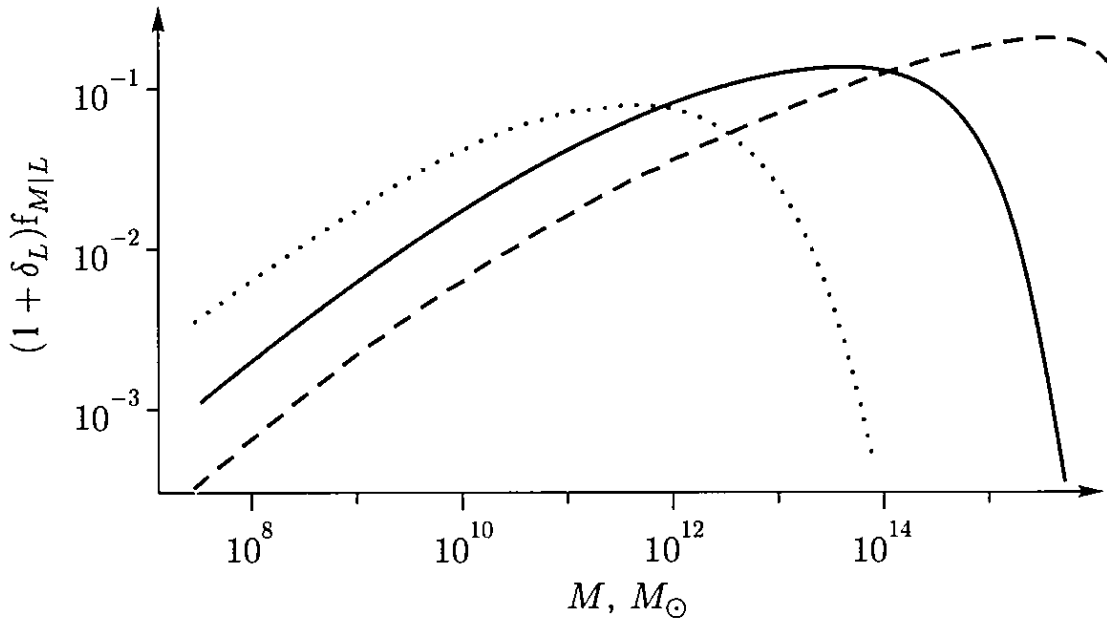


Рис. 12.14. Доля нерелятивистской материи $(1 + \delta_L) f_{M|L}$ в гало класса M (со средней массой M в интервале масс $\delta M = M$) в областях современной Вселенной размером $L = 140$ Мпк и средним контрастом плотности δ_L [Архипова и др., 2007]. Штриховая линия — сверхскопление ($\delta_L = 1/2$), сплошная линия — плоский фон ($\delta_L = 0$), пунктирная — войд ($\delta_L = -1/2$)

пространства размерами $L = 140$ Мпк и различными значениями среднего контраста плотности δ_L . Как видно, основным населением зон пониженной плотности являются *маломассивные* объекты гало. Например, доли вещества в войде с $\delta_L = -1/2$, существующего в виде *первичных изолированных блоков-гало* с массами $10^7 M_\odot$, $10^9 M_\odot$ и $10^{11} M_\odot$, составляют соответственно порядка 1%, 10% и 30%. Поэтому в войдах легче, чем в других частях Вселенной, обнаружить маломассивные первичные объекты темной материи, не вошедшие в более массивные системы гало.

Сплошная кривая на рис. 12.14 для пространственно-плоских областей КСМ характеризует современную Вселенную в целом. Мы видим, что в диапазоне масс гало $5 \cdot 10^{11} - 5 \cdot 10^{13} M_\odot$ (массивные галактики и группы галактик) сосредоточено 30% всей темной материи Вселенной, а в массивные гало с $M > 5 \times 10^{14} M_\odot$ (богатые скопления галактик) входит не более 10% ее нерелятивистского вещества ($f_M < 0,1$).

Итак, процесс иерархического гравитационного сгущивания, собирающий основную долю маломассивных объектов в более массивные системы с $M \lesssim 5 \cdot 10^{13} M_\odot$ и наблюдаемый как фаза активного слияния галактик, в целом во Вселенной уже завершен. Для более массивных образований с $M \gtrsim 10^{14} M_\odot$ этот

процесс сгущивания, наблюдаемый как взаимодействие групп и скоплений, еще продолжается и далек от завершения ¹⁾.

Если рассматривать отдельные области пространства, то в них процессы слияния и взаимодействия галактик проходили и идут по-разному. Так, сгущивание гало в войдах со временем совсем прекращается (см. предельную массу M_L в (12.7.6), (12.7.7)). Напротив, в «плоских» зонах и в областях повышенной плотности процессы сгущивания и слияния прогрессируют и постепенно сдвигаются в сторону больших масс. Можно сказать, что активность в форме слияний, начавшаяся вместе с процессом образования галактик повсеместно во всем объеме Вселенной, со временем вырождается, продолжаясь в уплощенных сверхскоплениях (занимающих меньшую часть объема Вселенной) и деградируя в войдах (занимающих большую часть объема Вселенной). Этот процесс напоминает догорающий костер.

В результате морфология галактик оказывается различной в разных местах Вселенной. Так, в войдах и вблизи их границ более вероятно (чем в сверхскоплениях) найти *изолированные галактики*, лишенные близких партнеров сопоставимых масс. А для сверхскоплений более типичны взаимодействующие галактики, начиная от пар и заканчивая большими ассоциациями или скоплениями.

Подводя итог, можно констатировать, что теория образования нелинейной структуры из темной материи хорошо согласуется с наблюдениями и численным экспериментом и имеет предсказательную силу для дальнейших исследований. Остаются некоторые проблемы, связанные с эволюцией барионной компоненты. Решение этих и других вопросов требует развития вычислительной техники, новых методов наблюдений и обработки данных.

Функции масс гало в войдах и сверхскоплениях значительно разнятся. В частности, пространственная плотность массивных ($M \sim 10^{12} M_\odot$) гало в войдах в несколько раз меньше, чем их плотность в сверхскоплениях, для больших масс различие в функциях масс еще более велико.

В области скоплений галактик статистика ухудшается:

$$n(M) \simeq 2 \cdot 10^{-5} \exp\left(-\frac{M}{M^*}\right) \text{ Мпк}^{-3}, \quad (12.7.9)$$

где $M^* \simeq 4 \cdot 10^{14} M_\odot$ в среднем во Вселенной. Поэтому для сравнения теории с наблюдениями используются *интегральные*

¹⁾ Дополнительным аргументом в пользу этого тезиса является клочковатая структура большинства скоплений галактик в рентгеновской области.

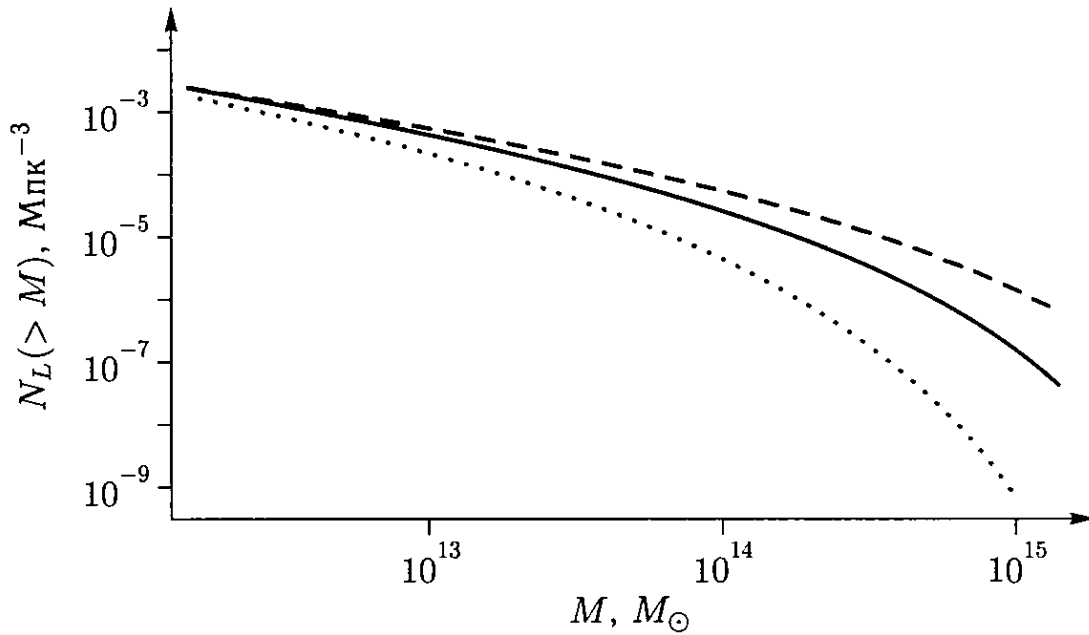


Рис. 12.15. Интегральные функции масс равновесных гало $N_L(>M)$ в интервале $10^{12}-10^{15} M_{\odot}$ для тех же областей Вселенной, что и на рис. 12.14

функции масс гало, обладающие более высокой статистикой данных. Интегральная функция масс описывает полное количество гало N_L в области L (приведенное к единице объема) с массами, больше некоторой ($>M$):

$$N_L(>M) = \int_M^{M_L} n_L(M) \frac{dM}{M}. \quad (12.7.10)$$

Различия между интегральными функциями в войдах и сверхскоплениях экспоненциально расходятся с ростом M , поскольку характерные значения M_L^* масс гало для них различаются (рис. 12.15).

В области малых масс, соответствующих карликовым галактикам, интегральные функции масс малоразличимы; здесь более значимы функции $f_{M|L}$ (рис. 12.14). Реальное различие между населенностями войдов и сверхскоплений усиливается из-за нелинейных эффектов слияния галактик и их приливного разрушения, которые подавляют маломассивную часть функции масс в сверхскоплениях.

Основной популяцией войдов являются маломассивные *первичные объекты* («ро»). Так, доля вещества войдов, находящегося в форме изолированных галактик с массами $< 10^9 M_{\odot}$, достигает 10%, тогда как в других областях Вселенной она не превышает 1–2%.

Шаровые скопления в центральных областях массивных галактик могут являться звездными остатками первичных объектов: они могли образоваться в молекулярных облаках первых гало малых масс и позднее были аккумулярованы в центральные области массивных галактик в ходе гравитационного процесса иерархического сгущивания.

Первичные объекты в войдах могут наблюдаться как слабые карликовые галактики или в виде абсорбционных систем L_α . Действительно, мы можем оценить их пространственную плотность и размер, полагая, что эпоха их образования $z_0 \sim 10$ и средняя масса $M \sim 10^7 M_\odot$:

$$n_{\text{po}} \sim \frac{\rho_m}{M} \sim 10^4 \text{ Мпк}^{-3},$$

$$r_{\text{po}} \sim \frac{1}{6} n_{\text{po}}^{-1/3} (1 + z_0)^{-1} \sim 1 \text{ кпк}.$$

Как видно, эти первичные образования довольно рыхлы, и поэтому эффекты линзирования на них несущественны. Возможность их обнаружения определяется дальнейшей эволюцией. Можно ожидать, что в войдах такие объекты сохранили газовую компоненту, тогда колонная плотность водорода в них может оказаться достаточной для появления слабых линий поглощения L_α . В свою очередь, вероятность попадания таких объектов на луч зрения в войде размером $L \sim 100$ Мпк порядка единицы: $n_{\text{po}} r_{\text{po}}^2 L \sim 1$.

12.8. Нормировка спектра возмущений плотности

Данные по анизотропии реликтового излучения дают точную информацию об амплитуде и форме начального S-спектра в *ранней* Вселенной и о ряде параметров состава (плотностей нерелятивистских компонент и кривизне пространства, см. раздел 10.5). Однако для превращения начальных неоднородностей в реальные наблюдаемые галактики необходимо также знать фактор роста возмущений плотности в *поздней* Вселенной, зависящий от хаббловского радиуса и темной энергии (см. разделы 12.1, 12.2).

Количество и распределение по массам гравитационно-связанных релаксированных гало, возникших из-за развития гравитационной неустойчивости в нерелятивистской материи, экспоненциально зависит от амплитуды начальных скалярных неоднородностей кривизны и фактора роста возмущений плотности (а также, если интересоваться локальными свойствами

распределений гало, от окружающей их крупномасштабной структуры Вселенной, см. разделы 12.4–12.7). Тем самым мы располагаем чувствительным тестом, позволяющим по количественным характеристикам структуры определить амплитуду и форму S-спектра (в меньших масштабах, чем по реликтовому излучению), а также важнейшие дополняющие параметры состава (Ω_E и др.).

Этот тест часто называют нормировкой спектра космологических возмущений плотности. При заданном наблюдаемом распределении по массам гравитационно-связанных гало материи амплитуда S-спектра является функцией остальных космологических параметров (Ω_m , Ω_E и др.) в пределах их возможных вариаций, допускаемых наблюдательной космологией.

По историческим причинам наиболее широко применяется нормировка «на сигму-восемь» — дисперсию контраста плотности в сфере радиуса $8h^{-1} \simeq 11$ Мпк, являющуюся интегральной функцией спектра возмущений плотности¹⁾ (см. (3.4.5)). Данная сфера охватывает массу невозмущенной Вселенной $M_{11} \simeq 2 \times 10^{14} M_\odot$, близкую к массам типичных скоплений галактик. Поэтому этот нормировочный тест идеально ориентирован на использование наблюдательных данных по обилию скоплений галактик, в то же время теоретический расчет функций масс гало в этом диапазоне масс опирается на хорошо разработанный аналитический формализм (см. раздел 12.6).

Значение дисперсии σ_{11} , обеспечившей наблюдаемую пространственную плотность скоплений, в сильной степени зависит от полной плотности материи во Вселенной Ω_m , тогда как изменение других космологических параметров приводят к вариации σ_{11} лишь в пределах 10–20% ее величины. При решении задач точной космологии необходимо учитывать влияние многих космологических параметров на нормировку спектра возмущений плотности, то есть знать величину σ_{11} , требуемую для генерации наблюдаемого обилия скоплений галактик.

Ниже мы продемонстрируем способ нормировки S-спектра по обилию скоплений галактик при $z = 0$, удовлетворяющий текущим требованиям точной космологии.

¹⁾ Мы обозначаем здесь эту дисперсию σ_{11} , имея в виду радиус $R = 8h^{-1} \simeq 11$ Мпк (см. раздел 3.4). Такой выбор размерности приемлем в точной космологии, поскольку радиальные расстояния измеряются в скоростях (красное смещение), а их перевод в Мпк предполагает хаббловскую константу, точность определения которой пока не очень высока (табл. 10.1).

Все космологические параметры можно разбить на два уровня по степени значимости их влияния на величину σ_{11} :

- Ω_m (первый уровень);
- Ω_x , h , n , f_ν , r , Ω_b , τ и др. (второй уровень).

Здесь $\Omega_x \equiv 1 - \Omega_m - \Omega_E$. Среди свободных параметров второго уровня статистически значимыми (в пределах интервалов их изменений) являются первые четыре, остальные для простоты считаем фиксированными ($r = \tau = 0$, $\Omega_b h^2 = 0,023$, $\Omega_E = \text{const}$).

Обозначенный класс моделей представляет собой достаточно полный вариант расширения КСМ на современном уровне знаний: он включает ненулевую пространственную кривизну, горячую, холодную и барионную компоненты вещества, космологическую постоянную и неплоские (но степенные) спектры S-моды. В этом классе моделей принимается в расчет точная зависимость порогового контраста плотности δ_c от Ω_m и Ω_E в искривленной Вселенной.

Дальнейшее исследование проходило в два этапа.

На первом для каждой реализации расширенной модели (со своим набором космологических параметров) путем сравнения теоретической и наблюдательной дифференциальных функций масс скоплений¹⁾ вычислялось оптимальное значение σ_{11} с индивидуальной ошибкой.

На втором этапе на все вычисленные значения σ_{11} (их десятки тысяч и каждое со своей ошибкой) накладывалась аппроксимационная зависимость следующего вида:

$$\sigma_{11} = \Omega_m^{A_1 + A_2 \Omega_m + A_3 \Omega_E} [A_4 + A_5 (\Omega_m - A_6) (1 - A_7 h - A_8 n - A_9 f_\nu)], \quad (12.8.1)$$

и методом минимизации находились наилучшие значения параметров A_i ($i = 1, 2, \dots, 9$). Отметим важное отличие данного этапа от предыдущего. В первом случае функция $n(M)$ являлась функцией одной переменной (массы M). Во втором — функция σ_{11} рассматривалась как функция многих переменных (Ω_m , Ω_E , h , n и f_ν).

¹⁾ Излагаемый подход рассчитан на любые наблюдательные данные по измерениям вириальных масс скоплений галактик. В нашем случае использованы результаты оптических наблюдений скоростей галактик в ближайших 150 скоплениях с медианным значением красного смещения $z \simeq 0,05$. Сравнение с наблюдениями выполнялось с помощью метода Левенберга–Марквардта минимизации χ^2 .

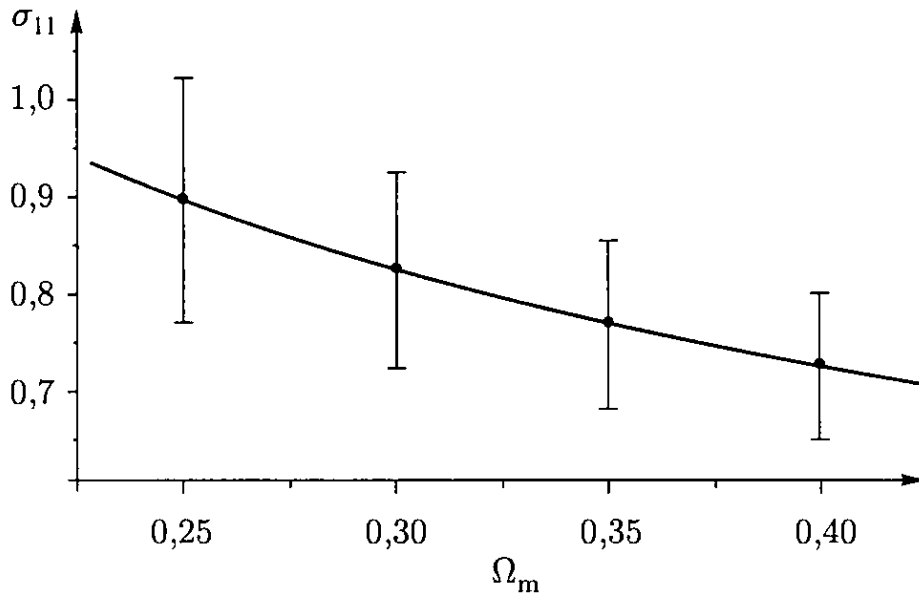


Рис. 12.16. Дисперсия контраста плотности $\sigma_{11}(\Omega_m)$ в сфере радиуса $8h^{-1} \simeq 11$ Мпк для космологической стандартной модели [Малиновский и др., 2008]. Вертикальные линии — точность наблюдательных оптических данных по обилию близких скоплений галактик (уровень достоверности 95 %)

Рисунок 12.16 иллюстрирует второй этап нормировки. На нем изображен «срез» пространства космологических параметров расширенной модели в двумерной плоскости основных параметров $\sigma_{11} - \Omega_m$. Значения остальных космологических параметров взяты в этом случае «стандартными»: $\Omega_E = 1 - \Omega_m$, $n = 1$, $h = 0,7$, $f_\nu = 0$. Рисунок подтверждает сильную ожидаемую зависимость σ_{11} от Ω_m , однако, как показывает анализ, общая амплитуда кривой $\sigma_{11}(\Omega_m)$ включает скрытую систематику — зависимость от выбранных наблюдательных данных.

Как мы уже неоднократно подчеркивали (см., напр., раздел 12.4), это, к сожалению, «обычная история» при работе с нелинейными объектами (в данном случае со скоплениями галактик). Об этом не следует забывать при упоминании о «стандартных значениях» в первую очередь таких параметров, как «сигма-восемь» (σ_{11}), H_0 и Ω_m (или Ω_E) — они в большей степени, чем другие величины, зависят от наблюдений астрономических объектов (звезд, квазаров, галактик, скоплений и пр.). Выход здесь только один: повышение качества наблюдательных данных и контроль систематических эффектов, связанных с эволюцией барионной компоненты материи.

Обратимся теперь к другим параметрам. То, что их вклад в величину σ_{11} носит статистически значимый характер, помимо формул (см. далее (12.8.2)) наглядно показывает рис. 12.17. На нем все пространство моделей спроецировано на плоскость $\sigma_{11} - \Omega_m$. Как видно, высота столбцов σ_{11} для различных зна-

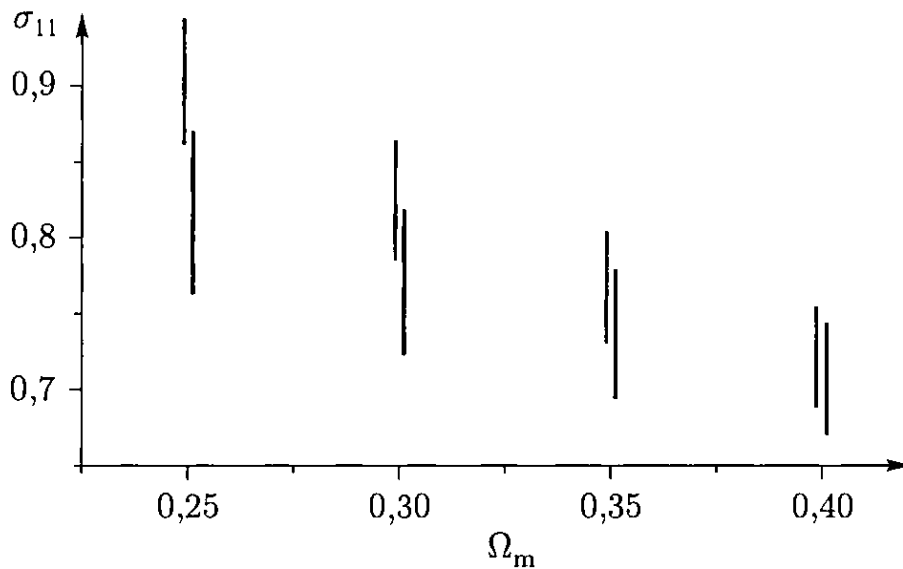


Рис. 12.17. Проекция пространства космологических параметров в расширенной модели Вселенной на плоскость $\sigma_{11} - \Omega_m$ для аппроксимаций Пресса–Шехтера (черные столбцы) и Шета–Тормена (красные столбцы) по данным [Малиновский и др., 2008]

чений параметров Ω_m , h , n и f_ν сопоставима с величиной статистической погрешности, а при некоторых наборах космологических параметров превышает ее. Это обуславливает важность корректного учета вклада всех вышеназванных параметров космологической модели.

Интересным результатом, который также можно увидеть на рис. 12.17, оказался рост влияния параметров второго уровня по мере перехода к меньшим значениям Ω_m . Так как именно в этой области сфокусирован интерес современной космологии, это еще раз подтверждает необходимость учета всех космологических параметров для решения задач точной космологии. Более кучно выглядит разброс значений σ_{11} при вариации параметров второго уровня для аппроксимации Шета–Тормена, что говорит о ее бесспорном преимуществе: наиболее точная нормировка S-спектра мощности получается в случае использования именно этой аппроксимации при работе с функциями масс скоплений галактик.

Приведем окончательный вид формулы (12.8.1) для нормировки S-спектра, полученный в результате сопоставления распространенности оптических скоплений галактик с теоретической функцией масс гало, рассчитанной с помощью аналитически-численной аппроксимации Шета–Тормена в расширенной модели Вселенной:

$$\sigma_{11} \Omega_m^{0,54 + 0,15 \Omega_m - 0,04 \Omega_E} - 0,2 (\Omega_m - 0,75) (1 - 0,2 h - 0,2 n + 0,8 f_\nu) = 0,53 \pm 0,08. \quad (12.8.2)$$

Мы представили этот результат в стандартной форме, где в левой части стоят комбинации космологических параметров, а все ошибки сведены в правую часть и приписаны нормировочной константе (уровень достоверности 95 %).

Данный пример еще раз иллюстрирует специфику космологических исследований при работе с наблюдательными данными:

- необходимость *многомерных* пространств параметров для определения *основных* «стандартных» параметров модели;
- необходимость *многих* разносторонних наблюдательных тестов для снятия проблемы *вырождения* параметров.

На этом примере мы видим, что одной только формулы нормировки (12.8.2) еще недостаточно для нормировки самой амплитуды S-спектра: мы знаем только связь параметров, и для решения этой задачи необходимы дополнительные данные. Отметим также интересную особенность формулы (12.8.2): она имеет вид «плоскости» для параметров второго уровня h , n и f_{ν} . Коэффициенты этой «плоскости» (так же как и другие коэффициенты формулы (12.8.2)) должны быть уточнены более полными данными будущих наблюдений.

Послесловие

Отправной темой книги является квантово-гравитационная теория параметрического рождения скалярных возмущений плотности в однородной модели Фридмана, созданная на основе квантования поля плотности в ОТО одним из авторов (В. Н. Лукашом) в 1980 году. В этой работе было высказано предположение о том, что эффект гравитационного рождения неоднородностей из вакуумных флуктуаций плотности несет ответственность за происхождение затравочных неоднородностей, из которых образовалась крупномасштабная структура Вселенной. Понадобилась четверть века развития астрономии, чтобы убедиться опытным путем, что основные предсказания теории:

- случайная пространственная фаза поля возмущений,
- их гауссова статистика,
- выделенная временная эволюционная фаза решения,

подтверждаются наблюдательными данными.

Кроме уже проверенных фазовых и статистических свойств поля рожденных неоднородностей плотности, теория предсказывает их спектральные характеристики. И здесь проверка еще не закончена.

Существующая теория генерации скалярных возмущений кривизны с необходимой для образования галактик амплитудой отсылает нас к очень ранней Вселенной, когда ее размер почти на 30 порядков мог быть меньше сегодняшнего, а характерные хаббловские частоты могли примыкать к области энергий Великого объединения. Столь беспрецедентная экстраполяция физической теории по шкалам энергий и размеров не является фантастикой и действительно может быть проверена с уже имеющимися средствами наблюдений и обработки данных.

Генерация космологических возмущений в хаббловском потоке идет непрерывно во времени и вызвана двумя фундаментальными причинами:

- невозможностью строго однородного распределения плотности (наличие вакуумных флуктуаций);
- нестационарностью гравитационного поля расширяющейся Вселенной (ненулевая внешняя кривизна).

В силу своего универсального характера и дальнего действия гравитация обладает удивительным свойством. При определенных физических состояниях вещества достаточно общие начальные неоднородные распределения материи и поля под действием гравитации динамически перестраиваются и упорядочиваются. В результате образуются высокосимметричные пространственно-временные гравитационные структуры. Из-за нестационарности гравитации таких структур, по сути, две. Это черные дыры, образующиеся при *коллапсе* (сжатии материи), и хаббловские потоки, образующиеся при *антиколлапсе* (инфляционном расширении материи). В первом случае гравитационное поле характеризуется двумя основными параметрами: массой и угловым моментом черной дыры. Во втором — хаббловской функцией $H(t)$ (внешней кривизной как функцией собственного времени). Ее также можно представить в виде набора параметров в данный момент времени: H , $\gamma \equiv -\dot{H}/H^2$ и других с более высокими производными H .

Теория параметрического усиления изложена в книге. На ее основании мы можем делать выводы о физической природе исследуемого эффекта. Классическая гравитация (при определенных условиях) создает упорядоченные, однородные и существенно нестационарные структуры. Масштаб нестационарности черных дыр — это r_g , а для фридмановской геометрии — хаббловский радиус H^{-1} . В силу квантовой природы материи и взаимодействий, строго однородных распределений не существует: имеются вакуумные флуктуации плотности и полей, которые развиваются во внешней нестационарной метрике. Происходит их спонтанное рождение, энергия реальных (рожденных) возмущений черпается из нестационарного классического гравитационного поля.

Внешняя кривизна H хаббловского потока задает характерную частоту его параметрического воздействия на гармонические осцилляторы полей возмущений метрики, описываемые скалярами q (S-мода) и q_λ (T-мода, $\lambda = \oplus, \otimes$). На частоте $k \simeq \overline{H} \equiv aH$ происходит «раскачка» их вакуумных возбуждений, что приводит к эффекту спонтанного рождения неоднородностей в ходе космологического расширения. Величина эффекта (дисперсия полей) пропорциональна энергии «внешней накачки»¹⁾ H :

$$q_k \simeq \frac{H}{M_P \sqrt{2\gamma}}, \quad q_{\lambda k} \simeq \frac{H}{M_P}.$$

¹⁾ И обратно пропорциональна ее изменению $\sqrt{\gamma}$ (для поля q , более высокие производные H в связь q с внешней метрикой не входят). Здесь для простоты $\gamma < 1$, для оценки поля q также принято $\beta \simeq 1$ и $\dot{\gamma} < \gamma H$ (см. (8.3.14), главы 7–9).

Эти удивительно простые формулы предсказывают эпоху генерации возмущений, близкую по энергии к Великому объединению (см. (10.5.4)). Однако окончательные выводы, подтверждающие применимость данной теории к нашей Вселенной, можно будет сделать только после экспериментального обнаружения Т-моды.

Дело в том, что одной лишь амплитуды S-моды, хорошо известной из наблюдательной космологии, явно не достаточно для определения *двух* параметров ранней Вселенной — H и γ . Только прямое детектирование Т-моды предоставит нам недостающую информацию о дисперсии возмущений и сможет окончательно подтвердить (или опровергнуть) теорию квантово-гравитационного происхождения космологических неоднородностей.

Часто перечисленные выше три свойства поля возмущений связывают с инфляцией. Однако это не так. *Случайная* пространственная и *детерминированная* временная фазы неоднородностей, также как и их гауссова статистика, являются «*визитной карточкой*» квантово-гравитационного эффекта рождения безмассовых полей, к которым принадлежат S- и Т-моды космологических возмущений в модели Фридмана. Эти три свойства присущи *любым* динамическим режимам хаббловского расширения, включая, в частности, и инфляционные.

Существуют ли *прямые* наблюдательные доказательства инфляционного характера Большого взрыва в эпоху генерации возмущений?

Следует заметить, что наблюдаемое верхнее ограничение на амплитуду космологических гравитационных волн, указывающее на малость параметра $\gamma \ll 1$ в ранней Вселенной, можно считать лишь *косвенным* аргументом в пользу инфляционного Большого взрыва (ограничение теории не есть ее доказательство). Окончательное доказательство первичной инфляции станет возможным только после обнаружения Т-моды космологических возмущений и проверки теоретически предсказанной связи показателя наклона Т-спектра с величиной параметра $r \simeq -2n_T$ (см. (10.5.3)). Это может случиться в ближайшие годы при обработке данных космических экспериментов по измерению анизотропии и поляризации реликтового излучения.

Часто в качестве доказательства инфляции приводят наблюдаемую плоскую форму S-спектра ($n \simeq 1$). Такой аргумент некорректен: инфляционные модели позволяют получить *произвольный* наклон S-спектра в зависимости от значения параметров (см. главу 9). Другое дело Т-спектр: он всегда близок к плоскому в инфляционных сценариях, и его наклон жестко

связан с g -параметром. Поэтому обнаружение T -моды явилось бы *прямым* доказательством (или опровержением) первичной инфляции. Несмотря на обнадеживающие предсказания простейших моделей (в частности, модели массивного скалярного поля) о значительной амплитуде космологических гравитационных волн, гарантии детектирования T -моды не существует (некоторые сценарии предсказывают ее незначительную амплитуду, см. главы 8, 9).

Однако надо оставаться оптимистами, и не только в отношении перспектив обнаружения T -моды, но и в вопросах исследования природы и структуры темной материи и темной энергии, барио- и лептогенезиса и др. Человечеству на удивление везет с великими открытиями в области фундаментальной физики последних десятилетий. Сегодняшняя ситуация в космологии напоминает те, что предшествовали великим открытиям: кварков, W - и Z -бозонов, осцилляций нейтрино, анизотропии и поляризации реликтового излучения. Теория исчерпала себя, указав, где и как искать T -моду, темную материю, темную энергию. Дело за экспериментом. Он открывает дорогу к новой физике.

Почему природа добра сегодня к нам и не скрывает от нас свои секреты?

Список литературы

- Архипова Н.А., Комберг Б.В., Лукаш В.Н., Михеева Е.В.* О населении космологических пустот // *Астрон. журн.* 2007. Т. 84. С. 874.
- Верходанов О.В., Парийский Ю.Н.* Радиогалактики и космология. — М.: ФИЗМАТЛИТ, 2009.
- Горбунов Д.С., Рубаков В.А.* Введение в теорию ранней Вселенной: теория горячего Большого взрыва. — М.: ЛКИ, 2008.
- Горбунов Д.С., Рубаков В.А.* Введение в теорию ранней Вселенной: космологические возмущения, инфляционная теория. — М.: УРСС, 2010.
- Гуревич А.В., Зыбин К.П.* Крупномасштабная структура Вселенной. Аналитическая теория // *Успехи физ. наук.* 1995. Т. 165. С. 723–758.
- Гриб А.А.* Основные представления современной космологии. — М.: ФИЗМАТЛИТ, 2008.
- Гриб А.А., Мамаев С.Г., Мостепаненко В.М.* Квантовые эффекты в интенсивных внешних полях. — М.: Атомиздат, 1980.
- Грищук Л.П.* Усиление гравитационных волн в изотропной Вселенной // *ЖЭТФ.* 1974. Т. 67. С. 825.
- Дорошкевич А.Г., Зельдович Я.Б., Сюняев Р.А.* Флуктуации реликтового излучения в адиабатической и энтропийной теориях формирования галактик // *Астрон. журн.* 1978. Т. 55. С. 913–921.
- Захаров А.Ф.* Гравитационные линзы и микролинзы. — М.: Янус-К, 1997.
- Зельдович Я.Б.* Горячая модель Вселенной // *Успехи физ. наук.* 1966. Т. 89. С. 647–668.
- Зельдович Я.Б., Новиков И.Д.* Релятивистская астрофизика. — М.: Наука, 1967.
- Зельдович Я.Б., Новиков И.Д.* Строение и эволюция Вселенной. — М.: Наука, 1975.
- Курт В.Г., Зельдович Я.Б., Сюняев Р.А.* Рекомбинация водорода в горячей модели Вселенной // *ЖЭТФ.* 1968. Т. 55. С. 278–286.
- Ландау Л.Д., Лифшиц Е.М.* Теория поля. — М.: Наука, 1998.
- Линде А.Д.* Физика элементарных частиц и инфляционная космология. — М.: Наука, 1990.

- Лифшиц Е.М.* О гравитационной устойчивости расширяющегося мира // ЖЭТФ. 1946. Т. 16. С. 587–602.
- Лукаш В.Н.* Рождение звуковых волн в ранней Вселенной // Письма в ЖЭТФ. 1980. Т. 31. С. 631–635.
- Лукаш В.Н.* Рождение звуковых волн в изотропной Вселенной // ЖЭТФ. 1980. Т. 79. С. 1601–1616.
- Лукаш В.Н.* О соотношении тензорной и скалярной мод возмущений в космологии Фридмана // Успехи физ. наук. 2006. Т. 176. С. 113–116.
- Лукаш В.Н., Компанеец Д.А.* Возникновение структуры Вселенной из тепловых флуктуаций // Астрон. журн. 1981. Т. 58. С. 482–485.
- Лукаш В.Н., Компанеец Д.А., Новиков И.Д.* Возникновение первичных неоднородностей Вселенной // Астрон. журн. 1982. Т. 59. С. 424–433.
- Лукаш В.Н., Рубаков В.А.* Темная энергия: мифы и реальность // Успехи физ. наук. 2008. Т. 178. С. 301–308.
- Лукаш В.Н., Строков В.Н., Михеева Е.В.* Квантовая Вселенная // Успехи физ. наук [в печати].
- Малиновский А.М., Воеводкин А.А., Лукаш В.Н., Михеева Е.В., Вихлинин А.А.* Космологические ограничения на массу нейтрино по анизотропии реликтового излучения и крупномасштабной структуре Вселенной // Письма в Астрон. журн. 2008. Т. 34. С. 490–495.
- Малиновский А.М., Лукаш В.Н., Михеева Е.В.* Космологические тесты на основе данных по обилию скоплений галактик // Астрон. журн. 2008. Т. 85. С. 675.
- Мизнер Ч., Торн К., Уилер Дж.* Гравитация. — М.: Мир, 1977.
- Михеева Е.В., Лукаш В.Н.* Наблюдательные ограничения на Лямбда-инфляцию // Астрон. журн. 2004. Т. 81. С. 4–10.
- Михеева Е.В., Лукаш В.Н., Архипова Н.А., Малиновский А.М.* Современный статус моделей с «горячим» и «холодным» скрытым веществом // Астрон. журн. 2001. Т. 78. С. 195–204.
- Насельский П.Д., Новиков Д.И., Новиков И.Д.* Реликтовое излучение Вселенной. — М.: Наука, 2003.
- Новиков И.Д., Фролов В.П.* Физика черных дыр. — М.: Наука, 1986.
- Сахаров А.Д.* Начальная стадия расширения Вселенной и возникновение неоднородностей распределения вещества // ЖЭТФ. 1965. Т. 49. С. 345–358.

Старобинский А.А. Спектр реликтового гравитационного излучения и начальное состояние Вселенной // Письма в ЖЭТФ. 1979. Т. 30. С. 719–723.

Строков В.Н. О лагранжевой теории космологических возмущений плотности // Астрон. журн. 2007. Т. 84. С. 483–486.

Халатников И.М. Некоторые вопросы релятивистской гидродинамики // ЖЭТФ. 1954. Т. 27. С. 529.

Afonso C., Albert J.N., Andersen J., Ansari J., Aubourg E. et al. Limits on Galactic Dark Matter with 5 Years of EROS SMC Data // Astron. and Astrophys. 2003. V. 400. P. 951–956.

de Bernardis P., Ade P.A.R., Bock J.J., Bond J.R., Borrill J., Boscaleri A. et al. A Flat Universe from High-Resolution Maps of the Cosmic Microwave Background Radiation // Nature. 2000. V. 404. P. 955–959.

de Blok W.J.G., McGaugh S.S., Bosma A., Rubin V. Mass Density Profiles of Low Surface Brightness Galaxies // Astrophys. J. 2001. V. 552. P. L23.

DeWitt B.S. Quantum Theory of Gravity. I. The Canonical Theory // Phys. Rev. 1967. V. 160. P. 1113–1148.

Bolton A., antwpr.gsfc.nasa.gov/apod/ap080728.html.

Catinella B., Giovanelli R., Haynes M.P. Template Rotation Curves for Disk Galaxies // Astrophys. J. 2006. V. 640. P. 751.

Clowe D., Bradac M., Gonzalez A.H., Markevitch M., Randall S.W., Jones G., Zaritsky D. A Direct Empirical Proof of the Existence of Dark Matter // Astrophys. J. 2006. V. 648. P. L109.

Cosmology: The Physics of the Universe, eds. B.A. Robson et al. — World Scientific, 1996.

Cosmology and Gravitation II, ed. M. Novello. — Editions Frontieres, 1996.

Faber S.M., Jackson R.E. Velocity Dispersions and Mass-to-Light Ratios for Elliptical Galaxies // Astrophys. J. 1976. V. 204. P. 668–683.

Ford L.H., Parker L. Quantized Gravitational Wave Perturbations in Robertson–Walker Universes // Phys. Rev. 1977. V. D16. P. 1601–1608.

Freese K. Review of Observational Evidence for Dark Matter in the Universe and in Upcoming Searches for Dark Matter, arXiv:0812.4005.

Gentile G., Tonini C., Salucci P. Λ CDM Halo Density Profiles: Where do Actual Halos Converge to NFW Ones?, arXiv: astro-ph/0701550.

Gibbons G.W., Hawking S.W. Cosmological Event Horizons, Thermodynamics, and Particle Creation // Phys. Rev. 1977. V. D15. P. 2738–2751.

Graham A.W. Populating the Galaxy Velocity Dispersion — Super-massive Black Hole Mass Diagram: A Catalogue of (M_{bh}, σ) Values, arXiv:0807.2549.

Granett B.R., Neyrinck M.C., Szapudi I. Dark Energy Detected with Supervoids and Supercluster, arXiv:0805.2974.

Hanany S., Ade P., Balbi A., Bock J. et al. MAXIMA-1: A Measurement of the Cosmic Microwave Background Anisotropy on Angular Scale of 10 Arcminutes to 5 Degrees // *Astrophys. J.* 2000. V. 545. P. L5.

Hawking S.W. Quantum Cosmology, eds. S. Hawking and W. Israel. — Cambridge Univ. Press, 1989.

Hinshaw G., Weiland J.L., Hill R.S., Odegard N., Larson D. et al. Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Data Processing, Sky Maps, & Basic Results, arXiv:0803.0732.

Ivanchik A.V., Orlov A.V., Varshalovich D.A. Effects of Possible Deviations of Fundamental Physical Constants on Primordial Nucleosynthesis // *Astron. Lett.* 2001. V. 27. P. 615.

Jackson J.C., Jannetta A.L. Legacy Data and Cosmological Constraints from the Angular-Size/Redshift Relation for Ultracompact Radio Sources // *J. of Cosmology and Astroparticle Phys.* 2006. V. 11. P. 002.

Karachentsev I.D., Makarov D.I., Sharina M.E., Dolphin A.E. et al. Local Galaxy Flows within 5 Mpc, astro-ph/0211011.

Kardashev N. Lemaitre's Universe and Observations // *Astrophys. J.* 1967. V. 150. P. L135–L139.

Kleinert H. Equivalence Principle and Field Quantization in Curved Spacetime // *Electronic Journal of Theoretical Physics.* 2009. V. 6. P. 1–4.

Komatsu E., Dunkley J., Nolta M.R., Bennett C.L. et al. Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation, arXiv:0803.0547.

Komberg B.V., Kravtsov A.V., Lukash V.N. The Search and Investigation of Large Groups of Quasars // *Mon. Not. Roy. Astron. Soc.* 1996. V. 282. P. 713–711.

Komberg B.V., Lukash V.N. Great Attractors at High Redshifts? // *Mon. Not. Roy. Astron. Soc.* 1994. V. 269. P. 277–282.

Krichbaum T.P., Alef W., Witzel A., Zensus J.A., Booth R.S., Greve A., Rogers A.E.E. VLBI Observations of Cygnus A with Sub-Milliarcsecond Resolution // *Astron. and Astrophys.* 1998. V. 329. P. 873.

Lukash V.N., Mikheeva E.V. The Lambda-Inflation and CMB Anisotropy // *Int. J. Mod. Phys.* 2000. V. A15. P. 3783–3801.

Marchesini D., D'Onghia E., Chincarini G., Firmani C., Conconi P., Molinari E., Zacchei A. H_α Rotation Curves: the Soft Core Question // *Astrophys. J.* 2002. V. 575. P. 801.

Mikheeva E.V. Current Observational Constraints on Inflationary Models // *Gravitation & Cosmology.* 2008. V. 14. P. 38–40.

Mikheeva E.V., Doroshkevich A.G., Lukash V.N. A Solution of the Cusp Problem in Relaxed Halos of Dark Matter // *Nuovo Cimento.* 2007. V. 122. P. 1228–1234.

Navarro J.F. The Inner Density Cusp of Cold Dark Matter Halos, astro-ph/0311361.

Nolta M.R., Dunkley J., Hill R.S., Hinshaw G., Komatsu E. et al. Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Angular Power Spectra, arXiv:0803.0593.

Observational and Physical Cosmology, eds. F. Sanchez et al. — Cambridge Univ. Press, 1992.

Peebles P.J.E., Yu J.T. Primeval Adiabatic Perturbation in an Expanding Universe // *Astrophys. J.* 1970. V. 162. P. 815–836.

Percival W.J., Cole S., Eisenstein D.J., Cole R.C., Peacock J.A., Pope A.C., Szalay A.S. Measuring the Baryon Acoustic Oscillation Scale Using the Sloan Digital Sky Survey and 2dF Galaxy Redshift Survey // *Mon. Not. Roy. Astron. Soc.* 2007. V. 381. P. 1053.

Pointecouteau E., Arnaud M., Pratt G.W. The Structural and Scaling Properties of Nearby Galaxy Clusters. I. The Universal Mass Profile // *Astron. and Astrophys.* 2005. V. 435. P. 1–7.

Press W.H., Schechter P. Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation // *Astrophys. J.* 1974. V. 187. P. 425–438.

Rubakov V.A., Sazhin M.V., Veryaskin A.V. Graviton Creation in an Inflationary Universe and the Grand Unification Scale // *Phys. Lett.* 1982. V. B115. P. 189.

Sachs R.K., Wolfe A.M. Perturbations of a Cosmological Model and Angular Variations of the Microwave Background // *Astrophys. J.* 1967. V. 147. P. 73.

Sahni V., Starobinsky A.A. The Case for a Positive Cosmological Λ -Term // *Int. J. Mod. Phys.* 2000. V. D9. P. 373–443.

Schneider P., Ehlers J., Falco E.E. Gravitational Lenses.— Springer, 1999.

Silk J. Cosmic Black-Body Radiation and Galaxy Formation // *Astrophys. J.* 1968. V. 151. P. 459–471.

Sheth R.K., Tormen G. Large-Scale Bias and the Peak Background Split // *Mon. Not. Roy. Astron. Soc.* 1999. V. 308. P. 119–126.

Sunyaev R.A., Zeldovich Ya.B. Small Scale Fluctuation of Relic Radiation // *Astrophys. and Space Sci.* 1970. V. 7. P. 3–19.

Sylos Labini F., Visilyev N.L., Barishev Y.V. Large Scale Fluctuations in the Distribution of Galaxies from the Two Degree Field Galaxy Redshift Survey, arXiv:0902.0229.

Tegmark M., Zaldarriaga M. Separating the Early Universe from the Late Universe: Cosmological Parameter Estimation beyond the Black Box, arXiv:0207047.

Tully R.B., Fisher J.R. A New Method of Determining Distances to Galaxies // *Astron. and Astrophys.* 1977. V. 54. P. 661–673.

Unruh W.G. Notes on Black-Hole Evaporation // *Phys. Rev.* 1976. V. D14. P. 870–892.

Zeldovich Ya.B., Sunyaev R.A. The Interaction of Matter and Radiation in a Hot-Model Universe // *Astrophys. and Space Sci.* 1969. V. 4. P. 301–316.

Zhang Y.-W., Fan J.-H. Statistics of Superluminal Motion in Active Galactic Nuclei // *Chin. J. Astron. and Astrophys.* 2008. V. 8. P. 385–394.