

УСПЕХИ ФИЗИЧЕСКИХ НАУКФИЗИКА НАШИХ ДНЕЙ

523.11

**ТЕОРИЯ ВАКУУМА, БЫТЬ МОЖЕТ, РЕШАЕТ
ЗАГАДКУ КОСМОЛОГИИ****Я. Б. Зельдович**

Принцип неопределенности и нулевые колебания.— О собственных функциях.— Плотность энергии вакуума.— Пустое искривленное пространство.— Энергодоминантность и сингулярность.— Космология и поляризация вакуума. Решение без сингулярности.

1. ВВЕДЕНИЕ

Читатель вправе спросить: нет ли ошибки в заглавии статьи? Прежде всего: существует ли теория вакуума? Можно ли сказать что-то содержательное о пустоте, т. е. о пространстве, которое ничего не содержит? Всякому ясно, что подлежат экспериментальному и теоретическому исследованию тела и вещества, заполняющие пространство. Современная наука достигла огромных успехов, сводя все многообразие веществ сперва к комбинациям сравнительно небольшого числа химических элементов, а потом в начале XX века к комбинациям трех сортов элементарных частиц — протонов, нейтронов и электронов.

Эти частицы связаны между собой определенными силами: в частности, связь электронов в атомах, химическая связь в молекулах и силы, соединяющие молекулы в твердых и жидких телах,— все эти силы являются проявлениями электромагнитного взаимодействия электронов с ядрами, а также электронов и ядер между собой. Разумеется, для результата существенно, что электроны подчиняются квантовой механике и принципу Паули.

Но теория электромагнетизма приводит к выводу, что наряду со статическим электрическим полем, окружающим заряды, существуют и специфические решения — поле, свободно распространяющееся в пространстве и описывающее электромагнитные волны (радиоволны, свет, рентгеновские лучи, гамма-лучи). Следующий решительный шаг был сделан Эйнштейном в 1905 г.: свет следует рассматривать как поток фотонов, т. е. особых частиц. Этот вывод был сделан на основании анализа опытных данных, ему предшествовало построение Планком теории теплового излучения. В работе Планка впервые в физике появилась квантовая постоянная. Вскоре множество опытов по химическому действию света, по фотоэффекту и по рассеянию света электронами подтвердило фотонную теорию.

Значительно позже, в конце 20-х годов, существование фотонов было доказано теоретически как следствие последовательного применения квантовой теории к электромагнитному полю.

За последние двадцать лет открыто множество новых типов элементарных частиц. Эти частицы «не нужны» для описания объектов и явлений, с которыми мы встречаемся в повседневной жизни, включая самые сложные электронные приборы, ядерную энергетику, биологические и психические явления. Эти частицы можно было бы назвать лишними. Я вспоминаю заголовок статьи, относящейся к этому периоду: «Зачем нужны мю-мезоны». Но природа представляет собой единое целое, всеобъемлющая теория не может основываться на одной части знания, обусловленной мощностью сегодняшних ускорителей.

В последнее время мы начинаем правильно оценивать возможную роль всех различных частиц на начальном этапе развития Вселенной. Для понимания существования протонов и нейтронов нужно знать о всех типах частиц.

Но вернемся к теме, заявленной в заглавии.

Теория вакуума — нуждается ли эта теория в сведениях о частицах?

Вспоминается история человека, которому поручили продавать газированную воду на благотворительном базаре. Ему велели спрашивать: «С каким сиропом Вы желаете...?». Когда покупатель пожелал воды без сиропа, то наш герой спросил: «Без какого сиропа? Без малинового или без вишневого?». Остается добавить, что воду разливал физикохимик, член-корреспондент, а затем почетный член Академии наук Иван Алексеевич Каблуков (1857—1942), о рассеянности которого сложено много легенд.

Итак, вакуум — это пространство, в котором нет протонов, электронов, фотонов, мезонов и т. д., — зачем же нужно знать свойства частиц, которых в данном месте нет. Вкус газированной воды ведь не зависит от состава непримешанного к ней малинового сиропа...

Оказывается, однако, что это простое суждение неверно, и неверно потому, что природа подчиняется квантовой теории.

Применение квантовой теории, притом не только к атомам, плазме, излучению, но и к вакууму, оказывается чрезвычайно важным для астрономии.

Современная богатая и сложная картина вакуума возникает как логическое следствие экспериментов и теорий. Эта картина представляется неизбежной как результат долгой, согласованной работы ученых.

Еще одно замечание, прежде чем мы перейдем к предмету статьи. Концепция вакуума не упрощается простым переопределением слов. Нельзя говорить, что вакуум, т. е. пустое пространство, лишен всяких свойств, «по определению», и все сложности связаны с чем-то, что мы не должны называть «вакуум».

Мы должны определить вакуум как пространство без каких-либо частиц. Такое определение совпадает с условием минимума плотности энергии в данном объеме пространства. Если энергия какой-то области пространства E больше минимального значения энергии этой области E_{\min} , то величину E можно представить как сумму, $E_{\min} + \Delta$ и добавку Δ можно считать энергией поля или частиц, находящихся в данном объеме. Значит, состояние с $E > E_{\min}$ уже не следует называть «вакуум». Но действительные свойства того «минимального» состояния, которое называется «вакуумом», диктуются законами физики, и мы не можем настаивать, чтобы этот минимум был нулем и чтобы простейшая возможная ситуация была бы столь простой, как нам хочется.

Наряду с квантовой теорией нужно учитывать общую теорию относительности (ОТО). Основное положение ОТО состоит в том, что сама геометрия пространства в общем случае зависит от физических условий как в данной точке, так и в окружающих ее областях. В свете специальной

теории относительности правильнее говорить о комплексе «пространство-время». Не только геометрия пространства (углы, длины), но и ход часов, бег времени зависит от физических условий.

На стыке квантовой теории и ОТО лежит область исследований трудная, но необычайно увлекательная и принципиально важная. Некоторые вопросы, возникающие в этой области, я постараюсь осветить в данной статье хотя бы качественно, с минимумом простых формул. Хочу также предупредить читателей, что статья по необходимости оказалась очень неровной по трудности материала.

Физики-теоретики могут без ущерба пропустить гл. 2, гл. 3 и первую половину гл. 4.

Экспериментаторам и студентам эти параграфы могут быть полезны, хотя бы как напоминание о том, что входит в курс физики.

Наконец последняя глава содержит идеи, появившиеся в 1979—1980 гг., не устоявшиеся, не только трудные, но и не достаточно надежные. Тем не менее хотелось бы дать представление об актуальных нерешенных вопросах и привлечь к этим вопросам внимание, даже ценой неполноты и невнятности изложения.

2. ПРИНЦИП НЕОПРЕДЕЛЕННОСТИ И НУЛЕВЫЕ КОЛЕБАНИЯ

XIX столетие дало нам замечательное достижение: экспериментальный гений и глубина понимания Фарадея и теоретическая проницательность Максвелла привели к теории электрических и магнитных полей. Как уже отмечалось выше, следствием этой теории было предсказание электромагнитных волн, свободно распространяющихся в пустом пространстве, т. е. в вакууме. Одна часть спектра, видимый свет, была эмпирически известна с незапамятных времен.

Длинноволновая часть — радиоволны были открыты или, лучше сказать, специально созданы и использованы человечеством благодаря Герцу, Попову, Маркони. Затем пришла квантовая теория. Мы перешагнем через историю развития. Современный взгляд на электромагнитные волны подчеркивает их подобие механическому осциллятору, т. е. массе на пружине.

Если записать соответствующие уравнения (чего мы не будем делать здесь), то оказывается, что именно магнитное поле играет роль пружины, т. е. энергия магнитного поля аналогична энергии деформации пружины, зависящей от ее отклонения от положения равновесия.

Энергия электрического поля — аналог кинетической энергии движущегося шарика.

Таким образом, каждый определенный тип колебаний электромагнитного поля аналогичен механическому колебанию шарика на пружине. Понятие «тип» здесь пока не разъяснено, ниже мы вернемся к нему.

Здесь без детализации воспользуемся аналогией с механическими колебаниями. Что нам известно о колебаниях шарика на пружине? В классической теории масса шарика и упругость пружины определяют частоту колебаний; в ходе колебаний в среднем по времени кинетическая энергия равна потенциальной. И та и другая форма энергии пропорциональна квадрату амплитуды. Главное свойство классических колебаний заключается в том, что амплитуда их может быть любой. Амплитуда не определяется уравнениями движения; амплитуда зависит от начальных условий, которые в классической теории можно задать по произволу. В частности, возможен (т. е. является решением уравнений) случай шарика, покоящегося в состоянии равновесия, т. е. случай равной нулю энергии колебаний.

Далее, уточняя аналогию между шариком на пружине и электромагнитной волной, можно предвидеть основное свойство классических уравнений электромагнитного поля: возможно решение, в котором электрическое поле и магнитное поле везде в точности равны нулю. Соответственно равна нулю и плотность энергии поля. Подразумевается, что мы имеем дело с пространством, в котором нет электрических зарядов. Таким образом, в классической (не квантовой!) теории действительно понятие вакуума достаточно просто — нет полей и нет энергии.

Теперь на сцену выступает квантовая механика. Импульс (т. е. скорость, помноженная на массу) шарика и его координаты не могут быть измерены одновременно. Более того, импульс и координата не могут одновременно иметь определенные значения. Существует знаменитое соотношение неопределенности Гейзенберга. Применительно к электромагнитному полю не могут одновременно обращаться в нуль магнитное и электрическое поля.

Но квантовая механика не просто разрушает картину детерминированного движения. Это не просто негативная теория, она обладает положительным содержанием и предсказательной силой. Она предсказывает, что возможные значения полной энергии осциллятора есть $E_n = (n + 1/2) h\nu$ с произвольным целым n , где h — постоянная Планка, ν — частота осциллятора. Итак, возможны только состояния с тем или иным значением энергии из ряда

$$n=0, \quad E_0 = \frac{1}{2} h\nu, \quad n=1, \quad E_1 = \frac{3}{2} h\nu, \quad n=2, \quad E_2 = \frac{5}{2} h\nu, \dots$$

Если осциллятор может обмениваться энергией с какими-то другими телами, то он отдает или получает энергию только определенными порциями, кратными $h\nu$.

При переходе $n = 1 \rightarrow n = 0$ осциллятор отдает $h\nu$, при переходе $n = 0 \rightarrow n = 2$ приобретает $2h\nu$ и т. д.

В начале развития квантовой теории именно этот факт считался главным.

Но здесь мы хотим обратить внимание на загадочные «половинки», т. е. на величину $(1/2) h\nu$ энергии нижнего уровня осциллятора.

Это не ошибка, эксперименты с атомами и молекулами подтверждают наличие «половинок». Даже при самой низкой температуре, отдав всю энергию, которую он только мог отдать, шарик продолжает осциллировать с определенной энергией и амплитудой. Вы не можете пользоваться квантовой механикой и избавиться от этого результата. Нельзя представить себе шарик в покое в состоянии равновесия: это означало бы точно нулевую скорость во вполне определенном положении строгого равновесия — чудовищное нарушение принципа неопределенности, противоречащее современной теории.

По аналогии легко представить себе, что применение квантовой теории к электромагнитному полю неизбежно приведет к сходному результату: окажется, что не может одновременно обращаться в нуль электрическое поле и магнитное поле, окажется, что не может равняться нулю плотность электромагнитной энергии. Можно ставить вопрос о минимуме энергии, как мы говорим о наименьшем (основном) состоянии осциллятора. Ясно, однако, что этот минимум не равен нулю.

Теперь, чтобы продвинуться, пора уточнить вопрос о типах электромагнитных волн, поставить вопрос о том, какие величины входят в формулы, относящиеся к электромагнитным волнам, которые все-таки не являются шариками на пружинах. Существенно то, что в качестве переменных, т. е. в качестве аналога положения и скорости одного шарика, нельзя

взять магнитное и электрическое поле в одной точке пространства: в уравнения Максвелла входят производные по пространственным координатам, эволюция полей в данной точке зависит от значений полей в других точках пространства.

Это обстоятельство и вызывает необходимость рассмотрения отдельных волн, которые друг от друга независимы.

Однако перед тем, как мы пойдем дальше, нужно обсудить один математический вопрос — понятие собственных функций и собственных значений. Но кто может провести границу между математикой и физикой? Вопрос, к которому мы переходим, имеет огромное значение для физики. Читатели-профессиональные теоретики пусть пропустят нижеследующий параграф как тривиальный.

3. О СОБСТВЕННЫХ ФУНКЦИЯХ

Вопрос, которому посвящен этот параграф, возник задолго до квантовой теории, ОТО и даже до теории электромагнетизма.

Мы и начнем его изложение с рассмотрения механической системы, например струны. Для натянутой струны характерно взаимодействие соседних участков струны: оттянув ее или ударив в одном месте мы вызовем движение, которое с течением времени захватит и другие участки струны, которые находились в равновесии и не подвергались внешнему воздействию.

Поэтому нелегко решить задачу, например, о тепловом («броуновском») движении струны: на вопрос о том, какова вероятность определенного отклонения от равновесия данной частицы струны при данной температуре, последует контрвопрос: спрашиваем ли мы об отклонении данной частицы при известных или произвольных положениях других частиц в настоящее время или в прошлом.

Частицы взаимодействуют, и в этом причина сложности ситуации. Конструктивное решение вопроса состоит в том, чтобы найти простейшие невзаимодействующие между собой типы колебаний. Конкретно, для струны такие колебания имеют вид

$$y_n = a_n \cos(\omega_n t + \varphi_n) \sin \frac{\pi n x}{l}, \quad (2.1)$$

где y — отклонение струны от положения равновесия, x — координата вдоль струны, закрепленной в точках $x = 0$ и $x = l$, так что $y = 0$ в этих точках. Величина y_n есть амплитуда данного n -го колебания, т. е. максимальное по времени и по пространству отклонение от положения равновесия ($y = |a_n|$ там, где $|\sin(\pi n x/l)| = 1$, т. е. в точках $x = l(k+1/2)/\pi n$, $k = 1, 2, \dots, n-1$, в момент, когда $|\cos(\omega_n t + \varphi_n)| = 1$), величина φ_n — его фаза, ω_n — частота. Уравнение движения струны оставляет амплитуду и фазу неопределенной. Однако для частоты ω_n уравнение дает вполне определенную величину: $\omega_n = \pi n a/l$, где a — скорость распространения возмущения по волне, $a^2 = q/\mu$. Здесь q — натяжение струны, т. е. сила [г·см/с²], μ — плотность, т. е. масса на единицу длины [г/см]; $[q/\mu] = [\text{см}^2/\text{с}^2]$, как и должно быть.

Как получены эти результаты? Мы ищем решения вида $y(x, t) = A(t) \varphi(x)$. Замечаем, что такой вид решения — с разделенными переменными — возможен лишь для определенного набора функций $\varphi(x)$. Этот набор можно перенумеровать $\varphi_n(x) = \sin(\pi n x/l)$, каждому номеру n соответствует своя функция, с различной длиной волны, с различным (равным $n-1$) числом нулей между закрепленными концами.

Пора выписать уравнение движения струны:

$$\frac{\partial^2 y}{\partial t^2} = a^2 \frac{\partial^2 y}{\partial x^2}. \quad (2.2)$$

После подстановки решения с разделенными переменными получим

$$\frac{d^2 \psi_n}{dt^2} = -\omega_n^2 \psi_n. \quad (2.3)$$

Данный n -й вид колебаний ведет себя как осциллятор, как маятник с определенной частотой; см. формулу, приведенную выше.

Здесь следует отметить, что в том случае, когда плотность (масса на единицу длины) струны переменна, пространственные функции имеют более сложный вид. Еще сложнее ситуация, когда рассматриваются колебания пластины или трехмерного тела (например, сплошного упругого шара или колокола). Однако несмотря на более сложный вид функции пространственных координат x , y , z зависимость каждого вида колебаний от времени остается гармонической, т. е. описывается дифференциальным уравнением (2.3), выписанным выше. Итак, сплошное тело, так же как и электромагнитное излучение в определенном объеме, эквивалентно системе осцилляторов *).

Почему так важна возможность записать решение в форме системы независимых уравнений для отдельных осцилляторов? Одна сторона дела, которую сразу увидели классики XIX века, заключается в том, что, зная набор частных решений, можно построить решение задачи с произвольными начальными условиями. Мы имеем дело с линейным уравнением, любая сумма частных решений также является решением.

Различные начальные условия дают различный набор величин a_n и φ_n в общем выражении, приведенном ниже:

$$y = \sum a_n \cos(\omega_n t + \varphi_n) \sin \frac{\pi n x}{l}.$$

Есть, однако, и более глубокая причина, побуждающая пользоваться именно решениями такого типа.

Дело в том, что эти решения можно перенумеровать и упорядочить. Их можно расположить в ряд в последовательности возрастающего значения частоты. Можно найти число решений с частотой меньше определенного значения или в данном интервале частот.

В частности, для электромагнитного излучения в объеме V число таких решений равно

$$dN = V \frac{8\pi \nu^2 d\nu}{c^3}.$$

При этом подразумевается, что частота ν такая, что соответствующая длина волны $\lambda = c/\nu$ меньше размера сосуда $d \sim V^{1/3}$, и мы рассматриваем интервал $d\nu$ не слишком узкий, так что

$$dN = 8\pi \left(\frac{V}{\lambda^3} \right) \frac{d\nu}{\nu} \gg 1$$

(несмотря на $d\nu/\nu \ll 1$).

*) Струна с постоянной плотностью обладает тем специальным свойством, что все частоты кратны одной, наименьшей частоте, $\omega_n = n\omega_1$. Поэтому движение струны как целого при произвольных начальных условиях является строго периодическим. Через время $T_1 = 2\pi/\omega_1$ начальные условия точно воспроизводятся. В произвольном теле отношения периодов являются трансцендентными числами и движение в целом не является периодическим. Однако те отдельные типы колебаний, которые выделены выше, являются не только периодическими, но и гармоническими. Для дальнейшего только это существенно. Напомним, что рассматривается система без трения, в линейном приближении, что существенно для возможности рассмотрения невзаимодействующих между собой колебаний.

Соответственно, общее число решений с частотой меньше данной ν (отнесенное к единице объема) равно

$$n = \frac{8\pi}{3} \left(\frac{\nu}{c} \right)^3 = \frac{8\pi}{3\lambda^3}.$$

Для струны, колокола и т. п. есть физическое ограничение: минимальная длина волны колебаний не может быть меньше расстояния между атомами. Но в вакууме нет никакой определенной минимальной длины волны! Экспериментально в опытах на ускорителях излучаются фотоны с энергией около 10^{10} электрон-вольт, их длина волны $\lambda \approx 10^{-14}$ см.

В космическом излучении наблюдаются фотоны еще большей энергии и меньшей длины волны. Но важнее соображения релятивистской инвариантности: нет и не может быть никакого предела энергии фотона или длины волны, потому что это величины, которые зависят от движения наблюдателя. Для встречного наблюдателя энергия больше, длина волны меньше.

Вакуум имеет бесконечное число типов колебаний, точнее бесконечное число колебаний приходится на единицу объема вакуума. Теория должна учитывать этот факт, должна уметь преодолеть те трудности — вычислительные и концептуальные, т. е. «физические», — которые с этим фактом связаны.

4. ПЛОТНОСТЬ ЭНЕРГИИ ВАКУУМА

Вернемся теперь к высказанному выше утверждению ($E_0 = 1/2 h\nu$), следующему из квантовой теории.

Допустив скромные $0,5 h\nu$ для каждой отдельной волны, мы немедленно с ужасом обнаруживаем, что все волны вместе дают бесконечную плотность энергии. Если бы мы ограничились определенной максимальной

частотой ν_m , то получили бы ответ вида $\varepsilon = a \int_0^{\nu_m} \frac{1}{2} h\nu \cdot \nu^2 d\nu = (ah/8) \nu_m^4$,

где ε — плотность энергии, a — константа ($a = kc^{-3}$, c — скорость света, k — число порядка единицы). При $\nu_m \rightarrow \infty$ величина ε также стремится к бесконечности. Положив сразу $\nu_m = \infty$, получаем расходящийся интеграл.

В этом состоит известная проблема расходимостей, так называемая «коротковолновая катастрофа» квантовой электродинамики или, точнее, часть этой проблемы *). И здесь нет простого пути спасения, нельзя игнорировать или просто выкинуть эту проблему. Ненулевые поля в отсутствие фотонов (поля, отвечающие «половинкам» $1/2 h\nu$ для всех возможных ν) наблюдаются — «половинки» модифицируют движение электронов в атомах. Знаменитый эксперимент Лэмба — Ризерфорда подтверждает это. Дело в том, что наиболее многочисленные коротковолновые высокочастотные «половинки» относительно слабо действуют на электроны: при действии

*) Расходящиеся интегралы появляются не только при вычислении энергии вакуума, но и в других задачах, например, при вычислении поправок к массе элементарных частиц, зависящих от взаимодействия этих частиц с электромагнитным и с другими полями. Мы не останавливаемся здесь на этих вопросах, несмотря на огромную их важность. Развитие теоретической физики в последнее десятилетие было подчинено условию создания перенормируемой теории, т. е. такой, которая не дает бесконечных ответов. Этот принцип сыграл определяющую роль при формировании теории, объединяющей слабое и электромагнитное взаимодействия. Об этом в своей Нобелевской лекции (к которой мы еще вернемся позже) говорит один из создателей теории С. Вайнберг¹. Будем надеяться, что принцип согласия теории с экспериментальными данными, касающимися полной плотности энергии вакуума, также сыграет свою роль в развитии общей теории, охватывающей все поля.

быстроменяющейся силы за короткий период электроны сдвигаются незначительно. Теория с расходящейся энергией вакуума дает сходящиеся конечные результаты для движения электронов, дает поправки в теории спектров атома, и эксперимент это подтверждает. Решение загадки расходящейся энергии не может заключаться в простом отрицании нулевых колебаний. Такое отрицание привело бы к противоречию с современными точными экспериментами.

При внимательном рассмотрении роли нулевых колебаний («половинок») в лабораторной физике оказывается, что эта роль меньше, чем могло бы показаться на первый взгляд.

Электромагнитная энергия вакуума бесконечна, но при рассмотрении любого процесса, например, испускания фотона атомом, $A^* = A + \gamma$, эта бесконечность сокращается. Если процесс происходит в объеме V и плотность энергии вакуума обозначена ε_V , то уравнение сохранения (полной) энергии E_{tot} имеет вид

$$E_{\text{tot}} = \text{const} = \varepsilon_V V + E_{A^*} = \varepsilon_V V + E_A + E_\gamma.$$

Независимо от того, равна нулю или конечна, или бесконечна величина ε_V , получим закон сохранения энергии в обычной форме $E_{A^*} = E_A + E_\gamma$.

Очевидно также, что энергия вакуума по той же причине не может быть использована практически, для вращения электромоторов или для освещения. Огромным достижением конца 40-х и начала 50-х годов была разработка последовательного способа вычисления влияния «половинок» на движение электрона в атоме и на магнитный момент электрона. Речь идет о теории перенормировки в квантовой электродинамике. Фейнман, Швингер и Томонага в 1965 г. получили Нобелевскую премию за это. Теория великолепно согласуется с экспериментом. Поправки составляют около 10^{-3} и меньше от основного эффекта, точность эксперимента около 10^{-10} , т. е. 10^{-7} от поправок.

Наряду с Нобелевскими лекциями творцов перенормировки ²⁻⁴ можно рекомендовать наглядную, апеллирующую к интуиции читателя статью Вайскопфа ⁵. Новейшие измерения, полностью подтверждающие теорию, см. в статье ⁶. Отметим еще одно явление, связанное с идеей нулевой энергии,— эффект Казимира ⁷. Пусть в вакууме как-то расположены металлические, т. е. электропроводящие тела или диэлектрики. Их присутствие определенным образом влияет на спектр электромагнитных колебаний, а следовательно, и на нулевую энергию. Речь идет теперь уже о нулевой энергии пространства с телами.

Чем выше энергия кванта, тем меньше на его распространение влияет наличие тел в пространстве. Поэтому изменение нулевой энергии электромагнитных колебаний, связанное с наличием тел (металлов или диэлектриков), обращается в бесконечность слабее, чем сама нулевая энергия.

Но Казимир вычисляет еще более тонкий эффект: он находит зависимость нулевой энергии от взаимного расположения тел, например, от расстояния между пластинами незаряженного конденсатора. Но производная от энергии по перемещению есть сила, действующая в направлении перемещения. Эта величина оказывается конечной, соответствующие интегралы сходятся.

Физически ясно, что для коротких волн (во много раз короче расстояния между пластинами в случае конденсатора) расположение тел не существенно, короткие волны не вносят вклад в интеграл, дающий силу.

Таким образом, реально наблюдаемая сила, доступная измерению с помощью весов, оказывается, по существу, зависящей от нулевой энергии электромагнитных колебаний в вакууме. При этом бесконечная величина

этой энергии из расчета выпадает, сокращается, и теория дает результат, совпадающий с опытом.

Однако не во всех явлениях имеет место такая благоприятная ситуация, не везде плотность нулевой энергии выпадает. Не случайно выше была сделана оговорка насчет лабораторной физики.

Важнейшим проявлением ненулевой плотности энергии вакуума могло бы стать ее влияние на поле силы тяготения и на гравитационный потенциал.

В теорию тяготения входит сама плотность энергии тела, включающая и плотность энергии вакуума внутри тела и в окружающем пространстве.

Здесь мы не имеем дело с разностями энергии, которые могли бы сократиться. На первый взгляд возникает непоправимое противоречие. Выйти из этого противоречия, в принципе, возможно, учитывая вклад других частиц. Подчеркнем только (мы вернемся к этому позже), что эта принципиальная возможность современной наукой еще количественно и точно не реализована!

Обратимся к теории электронов. В настоящее время школьникам (по крайней мере интересующимся школьникам) известно, что энергия электрона и его импульс связаны релятивистским уравнением $E^2 = c^2 p^2 + c^4 m^2$, так что $E = \pm \sqrt{c^2 p^2 + c^4 m^2}$, и к двум знакам перед квадратным корнем нужно относиться серьезно.

В классической теории можно было утешаться тем, что импульс и энергия электрона могут меняться только плавно. Если в начальном состоянии все электроны имеют энергию $E = +\sqrt{\quad}$ (подкоренное выражение не выписываем), то так будет всегда и о состояниях с $E = -\sqrt{\quad}$ можно не заботиться. Однако в квантовой теории состояния с $E = -\sqrt{\quad}$ исключить нельзя.

Квантовые законы движения не исключают такой прыжок электрона вниз, например от $E = +m_e c^2$ к $E = -m_e c^2$, с испусканием двух или трех фотонов.

Несколько лет назад отмечалось семидесятилетие Дирака. Одновременно физики праздновали пятидесятилетие замечательной идеи Дирака: прыжки электронов вниз запрещены принципом Паули!

Для этого Дирак ввел представление о бесконечном числе электронов, занимающих все без исключения состояния с отрицательной энергией («море Дирака»). Вакансии, т. е. незанятые состояния в этом море наблюдаются как положительный заряд (отсутствие отрицательного есть нечто положительное). Эта теория была блестяще подтверждена открытием позитронов: все их свойства совпадали с предсказаниями для вакансий в море Дирака.

Электрон с положительной энергией может упасть в незанятое состояние, испуская, например, два фотона. Так описывается аннигиляция электрона и позитрона.

Теория Дирака в действительности симметрична: можно было бы считать «элементарными» позитроны, ввести понятие моря позитронов с отрицательной энергией, а электрон считать вакансией в этом море.

Наконец, в современном изложении вводят операторы рождения и уничтожения электронов и позитронов и формулируют теорию сразу симметрично, не прибегая к наглядному понятию заполненного моря состояний с отрицательной энергией *).

*) Выше мы говорили о симметрии в ином смысле — о возможности выбора между двумя альтернативными теориями, каждая из которых в отдельности (море электронов или море позитронов) несимметрична.

Такая современная симметричная теория снимает вопрос о плотности электрического заряда вакуума. В альтернативных теориях эта величина равнялась $+\infty$ или $-\infty$. В симметричной теории плотность заряда равна нулю уже в силу симметрии.

Для дальнейшего изложения, для интересующего нас вопроса существенна другая особенность теории Дирака.

Поскольку рассматриваются заполненные состояния с отрицательной энергией («моря»), то естественно получается и отрицательная суммарная плотность энергии. Это свойство теории сохраняется полностью и в симметричной современной формулировке.

Таким образом, появляется возможность компенсировать положительный вклад в плотность энергии вакуума, связанный с нулевой энергией фотонов, отрицательным вкладом электронов.

В более общей формулировке: положительный вклад бозонов (частиц с целым спином), в принципе, может быть компенсирован отрицательным вкладом фермионов (частиц с полуцелым спином).

Обнадеживающим является тот факт, что главные, расходящиеся члены в интегралах имеют одинаковый порядок расходимости для положительных и отрицательных интегралов, для бозонов и фермионов.

Однако это отнюдь не дает основания для успокоения. Массы различных частиц не одинаковы. Надо учитывать взаимодействие разных сортов частиц между собой и при рассмотрении энергии вакуума — теперь это уже не удивляет нас.

Поэтому даже после сокращения бесконечностей, т. е. расходящейся части интеграла, вполне возможно было бы получение конечной, отличной от нуля величины.

Астрономические наблюдения говорят нам (см. ниже), что

$$|\rho_{vac}| < 10^{-29} \text{ г/см}^3, \quad |\varepsilon_{vac}| < 10^{-8} \text{ эрг/см}^3.$$

Сокращение между различными полями работает чудесным образом, поскольку простая оценка по порядку величины из соображений размерности давала бы

$$\rho_{vac} = m \left(\frac{mc}{h} \right)^3.$$

Если m — масса протона, мы получаем $m = 1,6 \cdot 10^{-24}$ г, $h/mc = 10^{-13}$ см, $\rho_{vac} = 2 \cdot 10^{14}$ г/см³.

Таким образом, можно утверждать, что плотность энергии вакуума не превосходит 10^{-43} доли простой оценки. Налицо замечательный, не объясненный до сих пор факт.

Вопрос возник около 50 лет назад. Общий прогресс физики элементарных частиц и теории поля за этот период поражает воображение. Трудно найти в истории другое пятидесятилетие, столь же полное открытий.

Тем не менее оказывается, что важнейший теоретический вопрос — о плотности энергии вакуума — до сих пор остается без ответа. Только астрономия дает определенные очень сильные ограничения.

В доказательство того, что вопрос осознан и что он не решен, приведем отрывок из Нобелевской лекции Вайнберга ¹ «В этом [в том, что массы частиц не сверхвелики. — Я. З.] нет ничего невозможного, но я не смог придумать никакой причины, почему это могло бы случиться. Проблема может быть связана со старой загадкой, почему квантовые поправки не создают огромную космологическую постоянную *). В обоих случаях

*) Отличную от нуля плотность энергии вакуума в теории тяготения, т. е. в общей теории относительности и в астрономии, принято называть «космологическая постоянная» или «космологический член в уравнениях». Подробнее об этом см. ниже.

в эффективном лагранжиане мы имеем дело с аномально малым «сверхренормализованным» членом, который подогнан к нулю. В случае космологической постоянной подгонка должна иметь точность около пятидесяти десятичных знаков».

Отметим в этой связи работу Хокинга с интригующим названием «Пространственно-временная пена»⁸.

Развивая идеи, которые давно высказывал Уилер в качественной форме, Хокинг обращается к самым малым масштабам, в которых необходимо учитывать сильные флуктуации метрики, не сводящиеся к нулевым колебаниям гравитационных волн (о метрике пространства подробнее см. начало гл. 5). Хокинг показывает, что возможны такие флуктуации, которые меняют топологию пространства-времени. Наглядным двумерным примером таких флуктуаций является превращение — последовательный переход от плоской пленки мыльного раствора к искривленной пленке (топология еще не изменилась!) и к пене, в которой соприкасаются частицы, первоначально удаленные друг от друга на конечное расстояние и разрываются частицы, ранее соседствовавшие (новая топология!). Отсюда и название статьи. В беседе с автором в июле 1980 г. Хокинг говорил, что при учете пенистой структуры эффективная плотность энергии вакуума, усредненная по большому масштабу, может обратиться в нуль. В статье 1978 г. и в последующих препринтах эта идея не отражена и нет количественных оценок остаточной энергии вакуума. Заманчивая идея саморегулировки, приводящей к $\varepsilon_{vac} = 0$, сегодня не реализована.

Как могла бы проявляться конечная плотность энергии? В релятивистской теории нужно, чтобы эта плотность энергии была одинаковой для любого наблюдателя. Это приводит к условию, что давление (натяжение) одинаково во всех направлениях и равно $p = -\varepsilon$, где p — давление, ε — плотность энергии вакуума. Еще в 1917 г. Эйнштейн рассматривал возможность того, что плотность энергии вакуума может быть отлична от нуля. Он пользовался другой терминологией и называл пропорциональную ε величину «космологической постоянной» Λ . Это название подчеркивало, что наиболее сильное влияние такая плотность энергии окажет на космологические явления.

В статье автора в «Успехах физических наук»⁹ «Космологическая постоянная и теория элементарных частиц» дан подробный обзор состояния вопроса, поставленного десятки лет назад. Недавно вопрос об $\varepsilon_{vac} \neq 0$ снова возник (Зельдович, Сюняев,¹⁰) в связи с сообщениями о массе нейтрино.

За время между⁹ и¹⁰ не только сделаны замечательные открытия и построены глубокие теории. Под влиянием конкретных достижений произошел и общий психологический сдвиг в семье физиков.

Критерий простоты природы сменился критерием единства и симметрии природы. Физика, в которой существовали бы только протоны, нейтроны, электроны, безмассовые фотоны и безмассовые нейтрино, была бы максимально проста: помните «простое как мычание» — термин Маяковского.

Казалось, что если нет прямых указаний на то, что масса нейтрино отлична от нуля, то естественно предполагать, что $m_\nu \equiv 0$ — такая теория проще. Сегодня наряду с экспериментальными указаниями на то, что $m_\nu \neq 0$, появляются теоретические работы с различными предсказаниями о возможной массе нейтрино.

Мы не обсуждаем здесь эти конкретные работы. Хочется только обратить внимание на общее настроение: если (или пока) не доказано, что масса нейтрино равна нулю, то в настоящее время естественно предположить, что она не равна нулю. Стабильность протона доказана эксперимен-

тально лишь в определенных пределах — естественно предположить, что с малой вероятностью, не противоречащей опыту, протон все же может распадаться.

Точно также, быть может, мы вернемся и к вопросу о плотности энергии вакуума. Сегодня нельзя исключить, что теория и наблюдения дадут некое, очень малое, но отличное от нуля значение ε_{vac} и Λ (см. выше).

Психологические этюды, приведенные выше, можно дополнить более техническими соображениями.

Физики считали, что разумные теории не дают безразмерных чисел, которые бы слишком сильно отличались от единицы. Единственным эстетически приемлемым исключением был нуль. При наличии порядковой оценки ε_{vac} , построенной с помощью массы протона (см. выше, $m^4 c^5 / \hbar^3$), казалось, что $\varepsilon_{\text{vac}} \equiv 0$ есть единственное приемлемое решение.

Но в действительности объединение всех сил природы, в том числе и гравитации, неизбежно приводит к появлению чисел, сильно отличающихся от единицы. Первым примером такого числа является $Gm_p^2/\hbar c = 10^{-37}$, где G — ньютоновская постоянная тяготения. В современных теориях объединения сильного и слабого взаимодействия фигурирует Х-бозон, который в 10^{15} раз тяжелее протона. Располагая такими числами, нетрудно построить формулу для ε_{vac} , не равной нулю, правильную по размерности и не противоречащую опыту. Например,

$$\varepsilon_{\text{vac}} = \frac{Gm_e^4 c^4}{\hbar^4} = 10^{-40} \text{ г/см}^3.$$

В последнее время появляются теории, устанавливающие определенную симметрию между бозонами и фермионами.

Это теория супергравитации и суперсимметрии. При этом число сортов бозонов и фермионов может оказаться таким, что космологическая постоянная автоматически будет равной нулю, что и является определенной заслугой соответствующего варианта теории. Однако наблюдаемые массы частиц очень различны. Нет сомнения, что в природе симметрия имеет место лишь асимптотически, при высоких энергиях. Поэтому вывод из суперсимметричных теорий правильно формулировать так, что может автоматически исчезнуть самый опасный член типа $\int k^3 dk$, расходящийся

как k^4 . Однако это не исключает членов, расходящихся как $m^2 \int k dk$ и конечных типа m^4 . К тому же в вакууме возможны флуктуации полей типа подбарьерных переходов, не описываемые теорией малых возмущений, — так называемые инстантоны.

Это обстоятельство также не учитывается в суперсимметричных теориях, когда эти теории дают $\varepsilon_{\text{vac}} = 0$.

Таким образом, несмотря на большую ценность понятия суперсимметрии, приходится констатировать, что остается загадка малости ε_{vac} , малости Λ .

Последнее замечание: ε_{vac} , получающееся в результате почти полного сокращения больших положительных и отрицательных величин, может иметь или положительное, или отрицательное значение. В этом заключается различие с обычным веществом и обычными полями (возбуждениями вакуума), которые всегда вносят только положительный вклад.

5. ПУСТОЕ ИСКРИВЛЕННОЕ ПРОСТРАНСТВО

Читатели знают об общей теории относительности, если не об ее математических сторонах, то по крайней мере о ее идеях. Лобачевский в России и Больяи в Венгрии первыми указали на то, что пространство

не обязательно должно удовлетворять законам школьной евклидовой геометрии (Гаусс развивал аналогичные идеи, но побоялся опубликовать их).

Следующей идеей, разработанной Риманом, была идея геометрии, меняющейся от места к месту. Тогда возникли вопросы о том, что заставляет меняться геометрию и какое влияние оказывает такая геометрия на движение тел, распространение света и т. п. Однако во времена Римана физика еще не была готова к ответу на эти вопросы.

Прежде чем была реализована эта программа, было изучено электромагнитное поле, разработана атомная теория, была введена специальная теория относительности, связывающая пространство и время. Только после разработки релятивистской теории электромагнитного поля возникли предпосылки для создания релятивистской теории тяготения.

В конце этого пути в 1915 г. Эйнштейн создал общую теорию относительности. В основу этой теории положено: 1) плотность и давление вещества делают пространство-время искривленным и 2) движение в искривленном пространстве описывает движение под действием гравитационных сил.

Общая теория относительности и есть теория гравитации, лучше сказать, — англичане передают этот оттенок артиклем the вполне определенная теория гравитации, притом логически замкнутая и великолепно согласующаяся со всеми опытами. Вместо кривизны пространства, обсуждавшейся в XIX веке (время до начала XX века полагалось абсолютным и ни от чего не зависящим), мы говорим теперь о кривизне комплекса пространства-времени. Сам темп времени — тикание часов или колебания в атомах, или старение человека — оказывается зависящим от гравитационного потенциала. В простейшем случае именно разница в течении времени и является мерой гравитационного потенциала. Гравитационное красное смещение, т. е. потеря энергии фотоном, покидающим Солнце и преодолевающим гравитационный барьер, зависит от потенциала. Здесь мы используем соображения квантовой теории. Фотон использует часть своей энергии E_γ , чтобы покинуть Солнце — и энергия фотона связана с его частотой $E_\gamma = h\nu$. Если E_γ на Земле меньше, чем E_γ на Солнце в начале пути, тогда ν_{abs} (наблюдаемая на Земле частота) меньше, чем ν_{em} (испускаемая на Солнце частота). Но длина волны равна $\lambda = c/\nu$, длина волны растет, спектральные линии сдвигаются из голубой части спектра в красную — отсюда название «красное смещение».

Частота уменьшается, и это означает, что наблюдатель на Земле может с полным правом сказать, что все процессы на Солнце протекают немного медленнее, течение времени иное. Сказанное поясняет, почему вместо простого термина «пространственная кривизна» используют термин «пространство-временная кривизна».

Пространство-время искривлено в областях, занимаемых материей, но то же происходит и в окружающих областях — гравитация является дальнедействующей силой, упругость пространства-времени перестраивает его и вне тел.

Вернемся к теме нашей статьи — к поведению вакуума.

Следует повторить все наши предшествовавшие рассуждения об энергии нулевых колебаний частиц и волн, о море Дирака и т. п. Но волны (например, электромагнитные) и волновые функции частиц (электронов, например) теперь искажены.

Мы должны повторить все расчеты с искаженными волнами. С самого начала ясно, что на малых масштабах, для самых коротких волн вызываемое кривизной искажение малое. Это означает, что самые неприятные бесконечности, связанные с бесконечным числом коротких волн, сокра-

щаются для каждого отдельного поля, если мы интересуемся только различием случаев искривленного и плоского пространства-времени *).

С другой стороны, это различие для фермионов и бозонов одного и того же знака. Поэтому сокращение плотности энергии в плоском пространстве не дает указаний на сокращение различия для искривленного и плоского пространства. Наше представление состоит в том, что эта разность ненулевая и является наблюдаемой. Давайте проанализируем эту разницу.

Одна часть эффекта маскируется полностью (об этом см. статью Гинзбурга, Киржница и Любушина¹¹⁾). Речь идет о той части плотности энергии и давления, возникающей в нулевых колебаниях и в море Дирака, которая строго пропорциональна плотности энергии и давлению обычного вещества, вызывающего искривление пространства. В пустом пространстве эта часть равна нулю. Ясно, что наше общее представление о гравитации не меняется из-за этой части поляризации вакуума; остается справедливым закон Ньютона для слабых полей и уравнения ОТО. Для наглядности предположим, что на каждый грамм вещества приходится, скажем, 0,01 г массового эквивалента плотности энергии, обусловленной указанными эффектами, т. е. поляризацией вакуума.

Подставим этот вклад в уравнения гравитации — мы получим $(Gm + 0,01Gm)/r^2$ для ньютоновского ускорения. Но полученный ответ можно интерпретировать как изменение гравитационной постоянной

$$Gm + 0,01Gm = G \cdot 1,01m = G'm.$$

Именно G' мы наблюдаем и измеряем, и поэтому для всей макроскопической физики оказывается несущественной «старая» ненаблюдаемая величина G , и (эти проблемы взаимосвязаны) не возникает вопрос о том, каков действительный вклад вакуума, взятый выше в тексте произвольно, для удобства, равным 0,01.

Подобная процедура была впервые использована в 50-е годы в квантовой электродинамике применительно к электрическому заряду: свободный заряд вызывает появление заряда поляризации вакуума, возмущая движение заряженных частиц (электронов и т. п.) в дираковском море состояний с отрицательной энергией, которое есть повсюду в вакууме. С каждым электрическим зарядом связано преобразование $e \rightarrow e'$ подобно преобразованию $G \rightarrow G'$ для гравитационной константы. Эта процедура называется «перенормировкой заряда». Ее можно провести даже при бесконечном отношении e/e' . Теоретики разработали схемы, как рассчитать все наблюдаемые эффекты, пользуясь всегда только наблюдаемой величиной e' . Это был большой успех 50-х годов! Надо подчеркнуть, однако, различие между электродинамикой и теорией тяготения. В электродинамике можно изучать взаимодействие двух элементарных частиц на очень близком расстоянии. При этом удастся прощупать распределение облака заряда вакуума, ответственного за различие между e и e' . Это облако на ощутимую величину меняет энергию атомных уровней. Для атома водорода изменение составляет около $2 \cdot 10^{-8}$ энергии связи электрона в атоме, но современные методы дали возможность измерить эту величину.

Гравитационное взаимодействие мы изучаем только на макроскопическом уровне, поэтому экспериментальное исследование гравитационной поляризации вакуума в настоящее время лежит далеко за пределами возможного. Приходится ограничиваться анализом теоретических выводов.

*) Таким образом, рассматривается ситуация, подобная эффекту Казимира; см. выше.

Итак, энергия вакуума в плоском пространстве очень мала; об этом мы знаем из астрофизических данных. Часть энергии вакуума в искривленном пространстве, пропорциональная энергии вещества, вызывающего искривление, проявляется как изменение неизвестной заранее константы тяготения и в этом смысле ненаблюдаема.

Представление о возможном изменении константы тяготения за счет поляризации вакуума не меняет формы уравнений, но меняет их смысл. Сахаров ¹² высказал гипотезу, согласно которой величина гравитационной постоянной может целиком определяться поляризацией вакуума.

Уравнения тяготения можно наглядно интерпретировать как проявление упругости пространства (напомним, что везде, где для краткости говорится «пространство», подразумевается четырехмерный комплекс «пространство-время»).

Первая половина ОТО состоит в рассмотрении движения частиц в кривом пространстве-времени. Кривизна влияет на движение частиц. Математически развивая эти соображения, мы находим уравнения движения частиц, уравнения движения жидкости, уравнения Максвелла.

По третьему закону Ньютона, — по закону равенства действия и противодействия, — естественно можно представить себе, что есть и обратное влияние частиц и полей на пространство. Когда рельсы действуют на вагон, поворачивая его траекторию, то и вагон с определенной силой действует на рельсы.

Вторая часть ОТО подобна рассмотрению поведения рельсов: кроме сил со стороны вагона надо учесть упругость рельсов и их связь со шпалами, с насыпью. Можно сказать, что уравнения Эйнштейна описывают упругость пространства. Поставленный в ¹² вопрос состоит в том, что, может быть, эта упругость целиком определяется эффектами поляризации вакуума, т. е. подобна эффекту Казимира. Уравнения Эйнштейна запишем так:

$$T_i^h = \frac{c^4}{8\pi G} \left(R_i^h - \frac{1}{2} \delta_i^h R \right).$$

Действие вещества на метрику определяется тензором энергии и импульса T_i^h (в частности, компонента T_0^0 есть просто плотность энергии, размерность эрг/см³).

В правой части в скобках стоит кривизна, имеющая размерность см⁻². Коэффициент $c^4/8\pi G$ как раз и представляет собой упругость вакуума.

Этот коэффициент велик в системе единиц см — г — с, $c^4/8\pi G = 5 \cdot 10^{45}$ г·см/с⁴. Но это еще ничего не значит. Важно, однако, что если мы упругость будем выражать через величины, характеризующие элементарные частицы, — постоянную Планка \hbar , скорость света c и массу протона, — то получим величину, во много раз меньше приведенной выше. Нужную размерность имеет выражение $m^2 c^3 / \hbar = 10^{11}$ (в системе см — г — с), почти в 10^{35} раз меньше, чем нужно. Казалось бы, что гипотеза не реальна, однако в настоящее время серьезно рассматриваются элементарные частицы, масса которых в 10^{15} раз больше массы протона. Гипотеза, сводящая упругость вакуума, а тем самым и теорию тяготения, к поляризации вакуума, снова привлекает внимание теоретиков. Свидетельством этого является доклад известного американского физика Адлера в Физическом институте имени Лебедева в октябре 1980 г.

Однако независимо от объяснения уравнений ОТО существуют вопросы, связанные с их изменением в ситуации, когда компоненты тензора кривизны R_{ikm}^i велики. Напомним, что полная характеристика кривизны дается именно этим тензором с четырьмя индексами. Величины R_{ik} и R — это

определенным образом составленные суммы величин R_{ikh}^i . Обращение в нуль или малость сумм R_{ih}^i и R еще не означает малости слагаемых R_{ikh}^i .

Итак, дело в том, что существуют другие части поляризации вакуума, которые зависят от пространственно-временной кривизны иначе, они не пропорциональны комбинации $R_{ih} - \frac{1}{2} g_{ih} R$, которая, в свою очередь согласно ОТО, пропорциональна плотности и давлению вещества *). Эти члены мы будем именовать дальше истинной поляризацией вакуума — ИПВ. Характерной чертой ИПВ является то, что она остается ненулевой даже в полном вакууме, т. е. в пространстве, лишенном обычного вещества. Пионерские работы в этом направлении делали в США Де Витт и Паркер. При этом свойства плотности энергии вакуума и давления вакуума, возникающих в процессе ИПВ, радикально отличаются от энергии и давления обычного вещества.

Насколько мне известно, именно в совместной работе Л. П. Питаевского и моей ¹⁴, опубликованной в 1971 г., эта особенность ИПВ была указана явно.

Дело в том, что в ИПВ *плотность энергии не обязана быть больше, чем давление*.

На языке специалистов это называется «нарушением принципа энергодоминантности для ИПВ». В следующей главе статьи мы обсуждаем важные следствия этого свойства.

6. ЭНЕРГОДОМИНАНТНОСТЬ И СИНГУЛЯРНОСТЬ

Начнем с простых примеров. Давление идеального газа равно $p_g = nkT$, где n — число частиц в единичном объеме, k — постоянная Больцмана, T — температура, индекс «g» означает газ. Средняя кинетическая энергия атомов составляет $3kT/2$, поэтому плотность кинетической (тепловой) энергии составляет $\varepsilon_{th, g} = 3nkT/2 > p_g$.

Но в нашем рассмотрении, связанном с гравитацией, мы должны также учесть массу покоя каждого атома и ее энергетический эквивалент mc^2 . Поэтому полная плотность энергии составляет

$$\varepsilon_{tot, g} = n \left(mc^2 + \frac{3}{2} kT \right).$$

В нерелятивистском газе $kT \ll mc^2$ и $\varepsilon_{tot} \gg p$, т. е. значительно больше, чем p . Для водорода при 1 000 000 К давление составляет $2 \cdot 10^{-7}$ от плотности энергии (включающей массу покоя).

Тепловое излучение дает другой пример с гораздо большим отношением давление/энергия: хорошо известно, что (индекс г — radiation, излучение) $\varepsilon_r = 7,5 \cdot 10^{-15} T^4$ (T в градусах Кельвина, ε — в эрг/см³) и $p = \varepsilon/3$, но все же p меньше, чем ε . Предельным случаем является излучение, распространяющееся в обоих направлениях только вдоль одной линии (вместо хаотического трехмерного теплового излучения). В этом случае давление p_{xx} на поверхность, перпендикулярную оси x , равно ε , но даже в этом случае оно не превышает ε . Энергодоминантность все равно не нарушается!

По-видимому, следует ожидать, что энергодоминантность остается справедливой в самых экзотических ситуациях, в сверхплотной ядерной материи, в нейтронных звездах и т. п.

Почему так много внимания уделяется энергодоминантности?

*) Здесь величины g_{ih} , R_{ih} , R , зависящие от координат и времени, описывают метрику и кривизну пространства-времени. Мы отсылаем читателя к «Теории поля» Ландау и Лифшица ¹³, где подробно и педагогично изложены основы ОТО.

Английская школа — Хокинг, Пенроуз, Эллис и многие другие блестящие ученые доказали, что сингулярность, т. е. бесконечность, неизбежна в классической общей теории относительности. Это должно означать для космологии, что классическая теория Эйнштейна неприменима при описании самого начала расширения Вселенной, характерного для современной теории горячего взрыва. Сингулярность не приводит к абсурду всю теорию, но это есть пограничный столб на границе применимости теории.

В самом грубом приближении сингулярность можно представить себе как бесконечную плотность вещества, что согласно ОТО связано с бесконечной кривизной. С такой ситуацией можно встретиться либо при неограниченном сжатии, например при коллапсе, либо при расширении от начального состояния бесконечной плотности. Уравнения ОТО симметричны относительно изменения знака времени, поэтому математически задачи о коллапсе и расширении похожи. Сингулярность при коллапсе звезды не очень опасна: она скрыта черной дырой. Если в звезде происходит гравитационный коллапс, последние испускаемые электромагнитные волны и нейтрино излучаются веществом задолго до образования сингулярности. Точнее, надо было бы сказать, что коллапс опасен для того, кто падает в черную дыру, но безопасен для далекого «постороннего» наблюдателя.

Но в космологии сингулярность становится проблемой: она появляется не в конце, а в самом начале эволюции Вселенной. Бесконечная плотность в начале эволюции есть общий удел всего вещества, заполняющего в настоящее время Вселенную. Иногда говорят о Вселенной как о гигантской черной (или, точнее, белой) дыре, через которую прошло все сущее.

Однако одним из условий — аксиом, на которые опираются теоремы о неизбежности сингулярности, является условие энергодоминантности, $\epsilon > p$, или, в более общем виде, $\epsilon \geq T_{\alpha}^{\beta}$, где α и β пространственные индексы. Величина T_{α}^{β} есть обобщение понятия давления на тот случай, когда напряжения неізотропны: T_{α}^{β} есть компонента силы, направленной по оси α на единицу поверхности s^{β} , такой, что нормаль к поверхности направлена по оси β . Понятие T_{α}^{β} включает и наличие сдвиговых напряжений наряду с давлением. В изотропном случае, когда имеет место закон Паскаля, формула $T_{\alpha}^{\beta} = \delta_{\alpha}^{\beta} p$ есть определение давления.

В принципе, легко понять, что без каких-то условий, относящихся к веществу, заполняющему пространство, теоремы о сингулярности невозможны. В самом деле, в рамках ОТО можно перевернуть задачу. Напишем метрику, описывающую сжатие мира, плавно сменяющееся расширением, например

$$ds^2 = dt^2 - a^2(t) [dx^2 + dy^2 + dz^2],$$

или

$$ds^2 = dt^2 - a^2(t) [dr^2 + \sin^2 r (d\theta^2 + \sin^2 \theta d\varphi^2)],$$

с функцией $a(t)$ вида $a(t) = \sqrt{k^2 t^2 + r_0^2}$, или $a(t) = r_0 \operatorname{ch} kt$. Функции $a(t)$ здесь выбраны так, что $a(t) \rightarrow \infty$ при $t \rightarrow -\infty$ и $a(t) \rightarrow \infty$ при $t \rightarrow +\infty$, а при $t = 0$ функция $a(t)$ имеет определенное значение r_0 . Уравнения ОТО позволяют элементарно определить плотность энергии и давление (как функции времени) вещества, заполняющего Вселенную. Более того, можно быть уверенным, что энергия и давление в ходе эволюции будут удовлетворять первому началу термодинамики, $d(a^3 \epsilon) = -p d(a^3)$, которая соответствует известному $dE = -p dV$ при учете того, что элемент объема пропорционален величине a^3 .

Известно (см., например, ¹³ или ²⁵), что общие уравнения для выписанных выше вариантов метрики упрощаются и имеют вид

$$\frac{1}{a} \frac{d^2 a}{dt^2} = -\frac{4\pi G}{3c^2} (\varepsilon + 3p),$$

$$\left(\frac{1}{a} \frac{da}{dt} \right)^2 = \frac{8\pi G}{3c^2} \varepsilon + \frac{kc^2}{a^2},$$

где $k = 0$ для первой (плоской) метрики и $k = -1$ для второй метрики (для замкнутого мира). Существуют, однако, и решения с $k = +1$ (гиперболический мир). Здесь различие между тремя вариантами несущественно. Во всех случаях элементарно для произвольной плавной функции $a(t)$ можно найти ε и p , которые остаются конечными на всем интервале $-\infty < t < +\infty$.

Заметим, что плавный переход от сжатия к расширению обязательно требует того, чтобы где-то в промежутке находилась точка максимального сжатия. В этой точке $a(t)$ имеет минимум $da/dt = 0$. Само по себе это не определяет еще ε , однако существенно, что в минимуме a вторая производная положительна! Значит, сумма $\varepsilon + 3p$ обязана быть отрицательной.

Соответственно, если положить, что сумма $\varepsilon + 3p \geq 0$ всегда положительна, то плавный переход от сжатия к расширению невозможен, сингулярность $a \rightarrow 0$ неизбежна. Это и есть простейшая запись условия энергодоминантности для изотропной однородной Вселенной.

В условие естественно входят ε и p , являющиеся суммой вклада вещества и вакуума, в том числе ИПВ, т. е. вклад искривленного пространства (вакуума). Для вещества естественно $\varepsilon > 0$ и $p > 0$. Если бы давление вещества стало уменьшаться при сжатии, вещество было бы неустойчиво и распалось бы на две фазы. Поэтому надежды на построение реалистического решения без сингулярности связаны с предположениями о поведении ИПВ. ИПВ — «истинная поляризация вакуума» — не связана условием энергодоминантности. Это не противоречит тому, что в конечном счете ИПВ составляется из вкладов различных известных полей (фотонного, электронного и т. п.), каждое из которых подчиняется энергодоминантности.

В ходе вычисления ИПВ нам приходится вычитать одни вклады из других, например, энергию моря электронов отрицательной энергии из нулевой энергии фотонов. Необходимость вычитания ясно видна из того факта, что космологическая постоянная, т. е. ИПВ плоского пространства мала или даже равна нулю. Для любой совокупности реальных фотонов условие энергодоминантности выполняется, как и для реальных электронов и позитронов.

Нарушение энергодоминантности легко увидеть в самом первом примере ИПВ плоского пространства-времени: вполне возможно, что плотность энергии вакуума ε_{vac} отрицательна, в этом случае $p_{vac} = -\varepsilon_{vac}$ положительно и поэтому давление больше, чем плотность энергии, что и требовалось доказать. Но в случае плоского пространства-времени мы знаем, что абсолютная величина ε_{vac} и p_{vac} мала. Космологическая постоянная заведомо не повлияет на ход решения на ранней стадии эволюции Вселенной, когда плотности были значительно больше сегодняшних $\rho = 10^{-29} - 10^{-31}$ г/см³. Нас интересует именно ИПВ (с ударением на «истинная») в сильно искривленном пространстве-времени, вблизи сингулярности, притом такая, которая позволит избежать сингулярности. Теперь уже никакое формальное несоответствие, никакая строгая теорема, никакие фундаментальные возражения не могут воспрепятствовать нашим надеждам на существование решения, свободного от сингулярности. Этому вопросу посвящена гл. 7. Здесь же мы только кратко упомянем еще об

одном фундаментальном эффекте, связанном с нарушением энергодоминантности в ИПВ: о возможности спонтанного образования обычных частиц гравитационными силами.

На этот счет также существовала теорема Хокинга¹⁵, запрещающая образование частиц гравитационным полем; теорема использовала условие энергодоминантности. Суть доказательства теоремы заключалась в следующем. Гравитационные силы выделяют энергию в данном объеме, совершая механическую работу при сжатии объема, когда сжатию противодействует давление, или при расширении объема, чему противодействует натяжение; возможно и сочетание обоих эффектов — сжатие в одном направлении и расширение в другом *).

Но представьте себе, что мы начинаем с вакуума — с нулевой плотности энергии и с нулевого натяжения и давления. Если есть энергодоминантность, то давление и натяжение остаются нулевыми, пока равна нулю плотность энергии.

Выделение энергии невозможно — так же как в случае с бароном Мюнхгаузеном из немецкого фольклора, который пытался вытащить себя из воды за волосы...

Но если энергодоминантность нарушается, то пространственно-временная кривизна может сначала создать давление и натяжение, а затем давление и натяжение могут вызвать выделение энергии. На этом и строилась линия рассуждений Питаевского и Зельдовича¹⁴ (дальнейшие примеры приводились Зельдовичем в сборнике «Волшебство без волшебства», изданном к 60-летию Уилера¹⁶). Пусть искажение пространства-времени мало и характеризуется малым параметром δ . Тогда давление и натяжение ИПВ порядка величины δ . Деформация также пропорциональна δ . Работа есть произведение давления и/или натяжения на деформацию, так что она пропорциональна δ^2 . Нарушив ИПВ, можно свести концы с концами: прямой расчет энергии образующихся частиц показывает, что она пропорциональна δ^2 в полном соответствии с вышеприведенными простыми соображениями. Здесь в ходе образования частиц мы прямо воспользовались нарушением принципа энергодоминантности, поскольку при малом δ малое p пропорционально δ и, следовательно, больше, чем ϵ , которое пропорционально δ^2 .

Возможность образования частиц гравитационным полем, т. е. кривизной пространства-времени, сейчас не вызывает сомнений. Прекрасная теория Хокинга, описывающая образование частиц малыми черными дырами, — так называемое «испарение черных дыр», лучший пример тому.

7. КОСМОЛОГИЯ И ПОЛЯРИЗАЦИЯ ВАКУУМА. РЕШЕНИЕ БЕЗ СИНГУЛЯРНОСТИ

Мы подходим к концу нашей истории. Все, о чем мы выше узнали, подводит к возможности понять предлагаемые в настоящее время решения увлекательнейшей загадки начальной стадии эволюции Вселенной.

Есть чудесное космологическое решение — закон, определяющий структуру и скорость расширения Вселенной, с исключительно высокой симметрией. Это — решение, предложенное голландским ученым де Ситтером очень давно, в 1917 г.

Одним из вариантов модели де Ситтера является плоская Вселенная с евклидовой геометрией трехмерного пространства, т. е. сечения $t = \text{const}$

*) Рассмотрим магнитное поле — оно имеет натяжение вдоль силовой линии и давление поперек. Легко получить движение, которое накачивает энергию и увеличивает поле, если мы растягиваем плазму вдоль поля, сжимая ее поперек поля.

четырёхмерного пространства-времени. Но масштаб этого пространства растёт экспоненциально:

$$ds^2 = c^2 dt^2 - dl^2, \quad dl^2 = e^{2Ht} (dx^2 + dy^2 + dz^2) = \\ = a^2(t) (dx^2 + dy^2 + dz^2), \quad a = e^{Ht}, \quad H = \text{const.}$$

Поэтому пространственно-временное многообразие не является плоским.

Решение де Ситтера обладает очень высокой симметрией. Оно отвечает расширению, но закон расширения даёт равное относительное приращение всех пространственных расстояний за каждый равный малый промежуток времени, т. е. $\Delta a/a = H \Delta t$, причём H постоянно, так что вся картина не меняется со временем. Абсолютное значение a не имеет физического смысла для плоского бесконечного трёхмерного пространства, так что изменение a не изменяет свойства решения.

Используя выписанные в предыдущей главе формулы, легко убедиться, что решение де Ситтера имеет место тогда, когда в правой части уравнения $\varepsilon = -p = 3H^2 c^2 / 8\pi G$.

Решение де Ситтера хочется использовать для описания начала Вселенной. Такое желание связано с тем, что если масштабную функцию $a = e^{Ht}$ экстраполировать к $t = -\infty$, она будет описывать бесконечное прошлое. Ранее предполагалось использовать ее для описания ранней фазы¹⁷, подразумевая, что в этот период Вселенная была заполнена горячей и плотной плазмой. Однако идея того, что именно плазма, т. е. горячее вещество, могла бы иметь отрицательное давление, выглядит странно. В последние несколько лет гипотеза о решении де Ситтера получила новое звучание. Первоначально Гурович и Старобинский в 1979 г. неявно¹⁸, а затем в 1979 и 1980 г. Старобинский^{19,20} явно и с рассмотрением всех этапов перехода от решения де Ситтера к обычному решению Фридмана, предложили модель де Ситтера с правой частью уравнений Эйнштейна, соответствующей ИПВ.

Важно отметить: 1) в решении де Ситтера симметрия метрики обуславливает как раз нужную симметрию плотности энергии и давления ИПВ, 2) «реальное», «нормальное» вещество или излучение должно отсутствовать на начальной стадии — оно бы испортило решение. ИПВ в решении Старобинского не является космологической постоянной плоского пространства-времени, хотя она и обладает тем же свойством, что и космологическая постоянная, $p/\varepsilon = -1$, благодаря симметрии кривизны в решении де Ситтера.

ИПВ пропорциональна четвертой степени величины H , характеризующей скорость изменения шкалы на начальном этапе расширения; ИПВ также пропорциональна числу сортов N элементарных частиц.

Поэтому уравнения Эйнштейна имеют вид *)

$$R_{ik} - \frac{1}{2} g_{ik} R = \frac{8\pi G}{c^4} T_{ik(\text{ИПВ})} \approx GNH^4 \frac{\hbar}{c^7} g_{ik}.$$

*) Для критически-математического читателя отметим, что в метрике де Ситтера все компоненты тензора кривизны четвертого ранга R^i_{klm} симметрично выражаются через локальные значения g_{ik} и пропорциональны H^2 . Поэтому R_{ik} просто равны $k_1 H^2 g_{ik}$, скаляр кривизны R равен $k_2 H^2$, и, следовательно, левая часть уравнения Эйнштейна $R_{ik} - \frac{1}{2} g_{ik} R$ равна $k_3 H^2 g_{ik}$, где k_1, k_2, k_3 — известные безразмерные числа.

Эти соотношения для данной метрики являются инвариантными, они не зависят от того, как задана координатная сетка в пространстве-времени. Метрика де Ситтера, приведенная выше, описывает часть поверхности четырёхмерного гиперболоида, вложенного в плоское пятимерное пространство Минковского. Метрика де Ситтера является четырёхмерным (псевдо) аналогом пространства Лобачевского постоянной кривизны.

Они имеют нетривиальное решение (кроме $H = 0$) при определенном

$$H \sim \frac{1}{\sqrt{GN\hbar/c^5}}.$$

При этом мы систематически опускали безразмерные множители типа квадратов и кубов числа π , но оставили N , которое, по-видимому, как отмечалось выше, составляет величину порядка сотни.

Для вышеприведенного H при N во много раз больше единицы мы вправе пользоваться выписанными выше уравнениями, уравнения и их решение корректны.

Применимость классического рассмотрения общей теории относительности ограничена величиной

$$H_{Pl} = \frac{1}{t_{Pl}} = \frac{1}{\sqrt{G\hbar/c^5}};$$

здесь t_{Pl} называется планковской единицей времени и составляет по порядку величины 10^{-43} с.

Сразу же после введения \hbar , т. е. постоянной Планка, ее автор понял фундаментальное значение этой постоянной для всей физики, выходящее далеко за рамки теории теплового излучения.

Две великие постоянные — ньютоновская гравитационная постоянная $G = 6,7 \cdot 10^{-8}$ см³/с²г и скорость света $c = 3 \cdot 10^{10}$ см/с были известны уже за несколько столетий до работы Планка.

Планк выдвинул далеко идущую идею о том, что трех констант G , c , \hbar достаточно для того, чтобы определить естественные единицы массы, расстояния и времени. С сегодняшней точки зрения классическая ОТО применима лишь к метрике, мало меняющейся за время t_{Pl} и на расстоянии $l_{Pl} = ct_{Pl}$. Применительно к метрике де Ситтера это дает условие $H < 1/t_{Pl}$.

Возвращаясь к решению Старобинского, отметим, что его можно рассматривать на уровне классической общей теории относительности, используя классические идеи о времени и пространстве, потому что мы квантуем все остальные поля, электромагнитное, электронное и т. п. Большое число N квантованных полей позволяет сохранить классические представления о времени и пространстве. Более точно, можно добавить, что и маленькая рябь на пространственно-временной метрике, отождествляемая с гравитационными волнами, может быть *квантована*, но это не испортит *классическую* картину усредненной метрики. Важно отметить, что при большом N окажется $H < 1/t_{Pl}$, а это и есть условие применимости классической теории.

Вернемся к гравитационным волнам в предлагаемой метрике без сингулярности. Старобинский¹⁹ рассматривает судьбу этих гравитационных волн. Первоначально они находятся на уровне нулевых колебаний, но

Слово «псевдо» связано с тем, что один из дифференциалов (cdt) имеет знак, противоположный остальным трем.

Из соображений размерности следует, что при большой кривизне, когда можно пренебречь массой частиц, ИПВ выражается через квадраты тензора кривизны и вторые производные этого тензора. Расчет ИПВ не зависит от гравитационной постоянной, ИПВ вычисляется в заданной метрике. Метрика характеризуется единственной размерной величиной H , и отсюда с точностью до численного множителя однозначно следует вид выражения в правой части, приведенной ниже в тексте. При этом симметрия решения де Ситтера приводит к выводу, что T_{ik} пропорционально g_{ik} . Следовательно, левая и правая части уравнений Эйнштейна пропорциональны g_{ik} . Решая одно из уравнений, мы автоматически обеспечиваем решение всех 10 уравнений. Можно добавить еще, что метрика де Ситтера является конформно плоской, ее можно записать в виде $\eta^{-2} [d\eta^2 - dx^2 - dy^2 - dz^2]$. Поэтому рождение реальных частиц не происходит до тех пор, пока не начинаются отклонения метрики от де Ситтеровской.

затем они растут по амплитуде, образуются реальные гравитоны, а потом и классические гравитационные волны. Возможно, эти волны будут наблюдаться на спутниках еще до конца нынешнего столетия. Очень хочется дождаться экспериментальных данных о гравитационных волнах космологического происхождения, так же как хочется узнать, как распадается протон и какая масса разных сортов нейтрино. Ради этого стоит прожить еще 20—30 лет!

Решение де Ситтера нестабильно, и Старобинский²⁰ подробно исследует, как закон $a(t) \sim e^{Ht}$ расширения пустого (кроме ИПВ) пространства с течением времени нарушается. Он показывает, что после периода пульсаций $a(t) \propto t^{2/3} [b + r \cos \varphi(t)]$ из-за образования реальных частиц возникает нейтральная очень горячая плазма и закон расширения сменяется на нормальный фридмановский закон для случая преобладания излучения:

$$a(t) \propto \sqrt{t}.$$

Затем следует период, подробно рассмотренный в обзоре Долгова и автора²¹. За счет несохранения барионов и малой барион-антибарионной асимметрии теории получается избыток барионов. Эта идея сейчас очень популярна. Здесь необходимо отметить, что возможность несохранения барионов была впервые отмечена Вайнбергом в 1964 г., в лекциях²². Он основывался на том, что нет безмассового векторного поля, которое было бы связано с барионным зарядом так, как электромагнитное поле связано с электрическим зарядом. Там же Вайнберг пишет о том, что барионная асимметрия Вселенной, возможно, связана с несохранением барионов. Однако при этом Вайнберг имел в виду стационарную Вселенную (Хойл, Бонди) с непрерывным рождением вещества. В известной книге Вайнберга «Первые три минуты»²³ закон сохранения барионов фигурирует среди фундаментальных законов природы, идея 1964 г. не упоминается. Сахарову²⁴ принадлежит заслуга соединения идеи несохранения барионов с теорией горячей Вселенной. Протоны и нейтроны практически устойчивы (время жизни $> 10^{30}$ лет) при низкой температуре, но процесс изменения барионного заряда может идти достаточно быстро (за время меньше 10^{-8} с) при высокой температуре за счет диссоциации барионов на кварки. Современное состояние вопроса освещено в упомянутом обзоре²¹. Наконец, после падения температуры ниже $1 \text{ ГэВ} = 10^{13} \text{ К}$ начинается широко известный сценарий эволюции Вселенной. Напомним в форме тезисов основные события: 1) нуклеосинтез гелия-4 и дейтерия в первичной плазме; 2) эра радиационно-доминированной плазмы, состоящей из фотонов и нейтрино с небольшой примесью разреженного, ионизованного газа; 3) отщепление излучения от вещества после объединения электронов и протонов в нейтральные атомы водорода; 4) рост возмущений, приводящий к образованию галактик, звезд и всего остального; см. книгу²³ или более подробные и математические^{25,26}. Все эти части сценария известны сейчас так же хорошо, как то, что Земля — круглая, как то, что Земля вместе с другими планетами вращается вокруг Солнца. Веские предположения о возможной массе покоя нейтрино меняют количественно теорию. В свете этих предположений происходит изменение картины образования структуры Вселенной (выделение скоплений галактик и т. п.). Слова, написанные выше об уровне наших знаний в применении ко всему сценарию эволюции звучат слишком самонадеянно. Не об этом ли сказано, что в космологии (или вообще в астрофизике) «часто ошибаются, но никогда не сомневаются»?

Однако теории нуклеосинтеза и образования равновесного фонового радиоизлучения и в самом деле остаются непоколебимым фундаментом

теории горячей Вселенной. Вместе с тем до последнего времени оставалось ощущение неполноты картины.

До того, как возникли идеи, связанные с решением де Ситтера и поляризацией вакуума, все время возникали мучительные вопросы. Что же представляет собой начало? Что было прежде, чем 20 миллиардов лет назад началось расширение?

Теперь у нас есть *некое* решение без резкого начала, простирающееся от $t = -\infty$.

Я написал «некое» решение, поскольку оно возможно не есть подлинное решение, возможно, найдутся другие предпочтительные варианты. Но даже *некое* решение является огромным шагом вперед. Теория горячей Вселенной в этом решении более не связана с произвольным актом творения, происшедшим в сингулярности. Один важный качественный контраргумент против теории расширяющейся Вселенной, быть может, уже разрушен. Существенной деталью новой концепции является то обстоятельство, что закон расширения де Ситтера походя решает вопрос о причинности. Любые две точки или частицы (далеко отстоящие друг от друга в настоящее время) в далеком де Ситтеровском прошлом находились на очень малом, экспоненциально малом расстоянии. Они могли быть причинно связаны в прошлом, и это делает возможным, по крайней мере, в принципе, объяснение однородности Вселенной в больших масштабах.

Конечно, остается еще много самых разнообразных вопросов о различных стадиях эволюции Вселенной. Наиболее труден вопрос о нестабильности решения де Ситтера. Можно сказать, что эта нестабильность на новом этапе, в других условиях воспроизводит нестабильность, которую имело бы обычное вещество, если бы давление в нем было отрицательным. Если в решении де Ситтера есть конечная вероятность (отнесенная к единице объема и единице времени) перехода в другое состояние, то можно ли продолжать это решение в области $t = -\infty$, решает ли оно задачу о мире, существующем вечно, но в разных формах? Далее, решение де Ситтера не одно: есть три решения, соответствующие плоскому, замкнутому и открытому миру (см. любую книгу по космологии). Выбор между этими вариантами по данным астрономических наблюдений был любимой задачей космологов много лет. Тип Вселенной, соответствующий этому выбору, остается потом в ходе расширения неизменным. Однако задача выбора варианта оказалась трудной и до сих пор не решена.

Не может ли фундаментальная теория обосновать выбор плоского варианта, который, весьма вероятно, соответствует современным наблюдательным данным с учетом массы нейтрино и косвенных соображений о росте возмущений?

Было бы чрезвычайно важно разработать теорию возмущений плотности в модели де Ситтера с ИПВ. Это еще один тип возмущений начальной метрики, помимо гравитационных волн. Дальнейшее поведение этих двух типов возмущений существенно отличается.

Гравитационные волны на поздней стадии затухают, их трудно наблюдать. Возмущения плотности растут вследствие гравитационной неустойчивости, и именно из-за этих возмущений из слабо-неоднородного газа образуются галактики.

Не могли ли эти возмущения возникнуть вследствие квантовых нулевых колебаний начальной метрики де Ситтера подобно гравитационным волнам? Какой спектр и какую амплитуду волн плотности можно ожидать? Согласуется ли теория с тем, что известно о современной пространственной структуре Вселенной и об амплитуде флуктуаций температуры фонового радиоизлучения?

Другие вопросы относятся к физике высокотемпературного периода. Был период, когда температура значительно превышала значения, отвечающие массам покоя известных частиц. Точной теории такой плазмы пока не существует. Был ли фазовый переход, о чем пишут советские физики Киржниц и Линде²⁷ (см. также²⁸), образовывались ли при этом нити и стенки аналогично структурам, возникающим при кристаллизации жидкости? Что представляет собой плазма со свободными кварками?

Период, наиболее близкий к нам, характеризуется точным знанием фундаментальных законов, определяющих поведение рассматриваемых частиц. Но задачи, относящиеся к этому периоду, сложны математически. Трехмерная гидродинамика и перенос тепла излучением — вот те проблемы, которые должны быть решены для того, чтобы изучать в деталях образование галактик и звезд. Ответы, относящиеся к структуре Вселенной, т. е. к ее неоднородности, являются статистическими, что затрудняет их сопоставление с наблюдениями. К этому добавляется необходимость знать гораздо больше о массе нейтрино, по сравнению с тем, что известно в настоящее время. Здесь уместно отметить, что венгерские ученые Маркс и Салаи первые активно начали рассматривать роль массы покоя нейтрино в космологии еще в 60-х годах.

Напомним, что в 1974 г. астрономические указания на существование скрытой массы привели независимо советские астрономы Эйнасто, Каасик и Саар и американцы Пиблс, Острайкер и Яхил.

Маркс и Салаи отстаивали объяснение скрытой (избыточной по сравнению с суммой масс звезд и газа) массы галактик и скоплений галактик тем, что скопления галактик и индивидуальные большие галактики окружены облаками (галó, ореолами), состоящими из нейтрино. Очевидно, что такая картина возможна лишь в том случае, если масса покоя нейтрино отлична от нуля и нейтрино движутся со скоростями во много раз меньше скорости света. В настоящее время в центре внимания роль массивных нейтрино в эволюции возмущений, до и после рекомбинации плазмы.

Этот очень короткий обзор показывает, что для теоретиков, занимающихся астрономическими проблемами, нет угрозы безработицы.

Прикосновение к великой тайне начала Вселенной является, может быть, самым волнующим моментом в развитии естественных наук. Счастье жить в такое время и ощущать драматический момент мужания человеческого познания.

Я благодарен А. Д. Долгову, Л. Б. Окуню, А. А. Старобинскому, М. Ю. Хлопову за обсуждение, ценные замечания и помощь. Особо хочу выразить благодарность В. Л. Гинзбургу, прочитавшему первые два варианта статьи, за справедливые и благожелательные критические замечания.

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

1. Weinberg S.— *Rev. Mod. Phys.*, 1980, v. 52, p. 121.
2. Фейнман Р.— *УФН*, 1967, т. 91, с. 29.
3. Швингер Ю.— *Ibid.*, с. 49.
4. Томонага С. И.— *Ibid.*, с. 61.
5. Вайскопф В. Сдвиг уровней атомных электронов.— М.: ИЛ., 1950.— С. 1.
6. Филд Дж., Пикассо Э., Комбли Ф.— *УФН*, 1979, т. 127, с. 553
7. Бараш Ю. С., Гинзбург В. Л.— *УФН*, 1975, т. 116, с. 5.
8. Hawking S.— *Nucl. Phys. Ser. B*, 1978, v. 144, p. 349.
9. Зельдович Я. Б.— *УФН*, 1968, т. 95, с. 209.
10. Зельдович Я. Б., Сюняев Р. А.— *Письма АЖ*, 1980, т. 6, с. 451
11. Гинзбург В. Л., Киржниц Д. А., Любушин А. А.— *ЖЭТФ*, 1971, т. 60, с. 451.
12. Сахаров А. Д.— *ДАН СССР*, 1967, т. 117, с. 70.

13. Ландау Л. Д., Лифшиц Е. М. Теория поля.— М.: Наука, 1973.
14. Zel'dovich Ya. B., Pitaevskij L. P.— *Comm. Math. Phys.*, 1971, v. 23, p. 185.
15. Hawking S.— *Comm. Math. Phys.*, 1970, v. 18, p. 301.
16. Zel'dovich Ya. B.— In: *Magic without magic: John Archibald Wheeler / Ed. G. R. Klauder.*— San Francisco: W. H. Freeman, 1972.— P. 277.
17. Gurevich L. E.— *Astrophys. Space and Sci.*, 1975, v. 38, p. 67.
18. Гурович В. Ц., Старобинский А. А.— *ЖЭТФ*, 1979, т. 77, с. 1683.
19. Старобинский А. А.— *Письма ЖЭТФ*, 1979, т. 30, с. 719.
20. Starobinskii A. A.— *Phys. Lett. Ser. B*, 1980, v. 91, p. 99.
21. Долгов А. Д., Зельдович Я. Б.— *УФН*, 1980, т. 130, с. 559; *Rev. Mod. Phys.*, 1981, v. 53, No. 1.
22. Weinberg S.— In: *Lectures on Particles and Fields / Ed. S. Deser, K. Ford.*— N.Y., 1964.— P. 482.
23. Weinberg S. *The First Three Minutes.*— N.Y.: Basic Books Publ., 1977.
24. Сахаров А. Д.— *Письма ЖЭТФ*, 1967, т. 5, с. 32.
25. Зельдович Я. Б., Новиков И. Д. *Строение и эволюция Вселенной.* М.: Москва, 1975.
26. Пиблс П. *Физическая космология.*— М.: Мир, 1975.
27. Kirzhnits D. A., Linde A. D.— *Phys. Lett. Ser. B*, 1972, v. 42, p. 471.
28. Zel'dovich Ya. B.— *Mon. Not. RAS*, 1980, v. 192, p. 246.